DNA binding to proteolytically activated TLR9 is sequence-independent and enhanced by DNA curvature

Yue Li, Ian C Berke and Yorgo Modis*

Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, CT, USA

Toll-like receptor 9 (TLR9) recognizes microbial DNA in endolysosomal compartments. The ectodomain of TLR9 must be proteolytically cleaved by endosomal proteases to produce the active receptor capable of inducing an innate immune signal. We show that the cleaved TLR9 ectodomain is a monomer in solution and that DNA ligands with phosphodiester backbones induce TLR9 dimerization in a sequence-independent manner. Ligands with phosphorothioate (PS) backbones induce the formation of large TLR9–DNA aggregates, possibly due to the propensity of PS ligands to self-associate. DNA curvature-inducing proteins including high-mobility group box 1 and histones H2A and H2B significantly enhance TLR9 binding, suggesting that TLR9 preferentially recognizes curved DNA backbones. Our work sheds light on the molecular mechanism of TLR9 activation by endogenous protein–nucleic acid complexes, which are associated with autoimmune diseases including systemic lupus erythematosus.

The EMBO Journal (2012) **31**, 919–931. doi:10.1038/emboj.2011.441; Published online 6 December 2011

Subject Categories: signal transduction; immunology

Keywords: CpG; DNA curvature; innate immune receptor; nucleic acid recognition; Toll-like receptor 9

Introduction

Vertebrates rely on their innate immune system to detect microorganisms. Toll-like receptors (TLRs) are the principal family of molecular sentries for the innate immune recognition of microbial patterns outside the cytoplasm (Janeway and Medzhitov, 2002). TLRs recognize broadly conserved structures such as bacterial cell wall components, microbial nucleic acid signatures and certain highly conserved proteins (Kawai and Akira, 2010). This recognition event induces the secretion of antimicrobial and inflammatory cytokines by dendritic cells within minutes of infection through activation of the NF- κ B signalling pathway. A subfamily of TLRs recognizes microbial nucleic acids in endolysosomal compartments exclusively (Ahmad-Nejad *et al*, 2002; Matsumoto *et al*, 2003; Latz *et al*, 2004; Nishiya *et al*, 2005). TLR3

recognizes double-stranded RNA (dsRNA) ligands longer than 40 base pairs (Alexopoulou *et al*, 2001; Leonard *et al*, 2008). TLR7 and TLR8 recognize single-stranded RNA ligands in mouse and humans, respectively (Diebold *et al*, 2004; Heil *et al*, 2004). TLR9 recognizes DNA ligands (Bauer *et al*, 2001). Until recently the prevailing paradigm was that TLR9 recognized unmethylated CpG motifs, which are abundant in bacterial DNA but relatively scarce in mammalian DNA (Krieg *et al*, 1995). Several independent studies have now demonstrated, however, that the dependence on CpG motifs for TLR9 activation is restricted to synthetic phosphorothioate oligodeoxynucleotides (PS-ODNs), and that natural phosphodiester oligodeoxynucleotides (PD-ODNs) bind and activate TLR9 via the 2' deoxyribose backbone in a sequence-independent manner (Latz *et al*, 2004; Yasuda *et al*, 2005; Haas *et al*, 2008; Wagner, 2008; Ashman *et al*, 2011). These studies support the previously proposed hypothesis that the discrimination between microbial and self-DNA is dependent primarily on the colocalization of DNA and TLR9 in endolysosomes rather than on a chemical specificity of the receptor for microbial ligands (Yasuda *et al*, 2005; Barton *et al*, 2006).

Structural studies have shown that binding of the ectodomains of TLRs 1, 2, 3 and 4 to their respective ligands induces homo- or heterodimerization of the receptors (Jin *et al*, 2007; Liu *et al*, 2008; Park *et al*, 2009; reviewed in Botos *et al* (2011)). The apposition of the cytoplasmic Toll/interleukin 1 receptor homology (TIR) domains upon receptor dimerization is thought to activate signalling by recruiting the TIR domains of signalling adaptors such as MyD88 (Kawai and Akira, 2010; Botos *et al*, 2011). In contrast, human TLR9 fused to green or yellow fluorescent protein formed homodimers even in the absence of ligand in Förster resonance energy transfer and chemical crosslinking experiments in live HEK293 cells (Latz *et al*, 2007). However, binding of certain DNA ligands to preformed TLR9 dimers induces a conformational change that brings the TIR domains close together, suggesting a similar mechanism of signalling adaptor recruitment as for the other TLRs (Latz *et al*, 2007).

The ectodomain of TLR9 must be proteolytically activated by endosomal proteases in order for DNA ligand binding to produce an innate immune signal (Ewald *et al*, 2008; Park *et al*, 2008). TLR9 is cleaved between leucine-rich repeats 14 and 15 (near residue 477) by asparagine endopeptidase and cathepsins to generate the functional receptor, thereby restricting receptor activation to endolysosomes (Ewald *et al*, 2011). TLR3 and TLR7 are processed in an analogous manner, suggesting that ectodomain proteolysis is a general regulatory strategy for all the nucleic acid-sensing TLRs (Ewald *et al*, 2011).

During apoptosis, tissue injury or viral infections, endogenous nucleic acid–protein complexes can localize to endolysosomal compartments and induce TLR-dependent signalling (reviewed in Deane and Bolland (2006)).

*Corresponding author. Department of Molecular Biophysics & Biochemistry, Yale University, 266 Whitney Avenue, Bass 430, New Haven, CT 06520, USA. Tel.: +1 203 432 4330; Fax: +1 203 436 4369; E-mail: yorgo.modis@yale.edu

Received: 14 June 2011; accepted: 9 November 2011; published online: 6 December 2011

Mislocalization or overexpression of TLRs can also cause endogenous nucleic acids to trigger an innate immune response (Deane and Bolland, 2006). The resulting TLR-dependent inflammation can stimulate B cells to produce pathogenic antibodies against DNA- and RNA-containing autoantigens (Christensen *et al*, 2005, 2006; Ding *et al*, 2006; Ehlers *et al*, 2006). The production of such autoantibodies is a cardinal feature of systemic lupus erythematosus (SLE). Lupus-prone mice lacking TLR7 or TLR9 have ameliorated or exacerbated autoimmune disease, respectively (Christensen *et al*, 2006).

Various so-called accessory molecules facilitate the delivery of TLR ligands to their cognate receptor. TLR4 requires co-receptor MD2 to bind lipopolysaccharide ligands (Park *et al*, 2009). CD14 enhances ligand recognition by TLR2, TLR3, TLR4, TLR7 and TLR9 (Akashi-Takamura and Miyake, 2008; Baumann *et al*, 2010). High-mobility group box 1 (HMGB1), a nuclear non-histone protein that is released by necrotic cells or during inflammation, enhances the recognition of certain ODNs (with partial PS backbones and 3' poly-G extensions) by TLR9 (Tian *et al*, 2007). HMGB1-DNA complexes bind to the receptor RAGE and are delivered into early endosomes for TLR9 recognition, which results in the activation of pDCs and B cells (Tian *et al*, 2007). HMGB1 induces pronounced bending in its DNA ligands (Paull *et al*, 1993; Pil *et al*, 1993; Stott *et al*, 2006). Notably, core histones, which also induce bending of DNA (Luger *et al*, 1997), can act as autoantigens in SLE (Rekvis and Hannestad, 1980; Hardin and Thomas, 1983).

Previous TLR9 ligand-binding studies have employed constructs with the uncleaved ectodomain. Here, we analyse the ligand-binding properties of the proteolytically activated TLR9 ectodomain. We show that the cleaved TLR9 ectodomain is predominantly a monomer in solution, and that PD-ODNs induce TLR9 dimerization in a sequence-independent manner. We report a marked enhancement of DNA binding to the cleaved TLR9 ectodomain by DNA curvature-enhancing proteins, including HMGB1 and histones H2A and H2B, suggesting that TLR9 preferentially recognizes curved DNA backbones. Our work sheds light on the molecular mechanism of TLR9 activation by endogenous protein-

nucleic acid complexes, which is associated with autoimmune diseases including SLE.

Results

The truncated mouse TLR9 ectodomain is a monomer in solution

Proteolytic cleavage of the TLR9 ectodomain near residue 477 is required to generate a functional receptor (Ewald *et al*, 2008, 2011; Park *et al*, 2008). To study the biophysical and ligand-binding properties of the proteolytically activated form of TLR9, we expressed the C-terminal cleavage fragment of mouse TLR9, mTLR9-cECD (residues 474–824), in a baculovirus-insect cell expression system. Although mTLR9-cECD was fully glycosylated, it was retained in intracellular compartments—presumably endolysosomes (Ewald *et al*, 2008)—instead of being secreted into the cell culture media, and mTLR9-cECD had to be extracted from the cells with detergent. The detergent was removed in the first purification step and all subsequent experiments were performed in the absence of detergent. The elution volume of mTLR9-cECD from a size-exclusion column was consistent with a molecular weight of 52 kDa based on a molecular weight calibration of the column with globular proteins (Figure 1A). Sedimentation velocity analysis of deglycosylated protein was also consistent with a monomer, with 68% of mTLR9-cECD sedimenting at 2.3 S, which corresponds to an apparent molecular weight of 41 kDa (Supplementary Table I). In contrast, full-length human TLR9 forms homodimers even in the absence of ligand (Latz *et al*, 2007; Chen *et al*, 2011). Collectively, these data suggest that the N-terminal region of the ectodomain (residues 1–473) is responsible for dimerization of full-length TLR9 in the absence of ligand. The circular dichroism profile of mTLR9-cECD indicated that the purified protein was folded, with mixed α -helix and β -strand content (Figure 1B).

Phosphodiester oligodeoxynucleotides induce mTLR9-cECD dimerization in a sequence-independent manner

Until recently the prevailing paradigm for TLR9 activation was that the receptor recognized unmethylated CpG motifs

Figure 1 The C-terminal cleavage fragment (residues 474–824) of purified mouse TLR9 (mTLR9-cECD) is folded and monomeric in solution. Size-exclusion chromatograms of mTLR9-cECD (A) and mTLR9-cECD deglycosylated with PNGase F (B). The major peak corresponds to a molecular weight of 78 and 51 kDa for the glycosylated and deglycosylated proteins, respectively, based on a molecular-weight calibration of the size-exclusion column with globular proteins. At least five bands corresponding to different glycosylation states can be distinguished for the glycosylated protein. Insets: Coomassie-stained SDS-PAGE of the size-exclusion fractions from the major peak. (C) The far-UV circular dichroism spectrum of deglycosylated mTLR9-cECD indicates that the purified protein has mixed α -helix and β -strand secondary structure content.

Figure 2 Isothermal titration calorimetry (ITC) of deglycosylated mTLR9-cECD binding to phosphodiester oligodeoxynucleotides (PD-ODNs). Heat release and binding isotherm curves for binding to mTLR9-cECD of (A) LIGPD, a 20-nucleotide PD-ODN with three GACGTT motifs, (B) ranLIGPD, a PD-ODN with the same base composition as LIGPD but with a randomized sequence no CpG motifs, (C) methLIGPD, a cytosine-methylated derivative of LIGPD, (D) INHPD, a 15-nucleotide PD-ODN with a CGGGG sequence. Top panels: enthalpic heat released at 25°C during titrations of PD-ODN into a solution of mTLR9-cECD. Bottom panels: integrated binding isotherms of the titrations and best-fit curves using a single-site binding model. Insets: enthalpic heat released at 25°C during titration of the ligand into buffer without mTLR9-cECD. K_d , dissociation equilibrium constant; N , number of binding sites; ΔH , change in enthalpy during binding, in kcal per mole of injectant; ΔS , change in entropy during binding, in kcal per mole of injectant.

within microbial DNA (Krieg *et al*, 1995; Bauer *et al*, 2001; Latz *et al*, 2007). It has now been established, however, that the CpG motif dependency of TLR9 activation is restricted to PS-ODNs, and that natural PD-ODN ligands activate TLR9 in a sequence-independent manner (Yasuda *et al*, 2005; Haas *et al*, 2008; Wagner, 2008). To determine the extent to which these ligand-binding properties are determined by the proteolytically activated ectodomain fragment of TLR9, we analysed binding of mTLR9-cECD to various PD-ODNs by

isothermal titration calorimetry (ITC; Figure 2 and Table I). A 20-nucleotide PD-ODN with three GACGTT motifs (LIGPD) bound to mTLR9-cECD with a dissociation equilibrium constant (K_d) of 694 ± 87 nM. GACGTT is the optimal murine CpG motif (Bauer *et al*, 2001). A PD-ODN with the same base composition as LIGPD, but a randomized sequence with no CG motifs (ranLIGPD), bound with a K_d of 2.3 ± 0.2 μM. These binding affinities are approximately 25- and 75-fold weaker, respectively, than those reported previously for the

Table 1 DNA ligands used and their mTLR9–cECD-binding properties

ODN	Sequence (5' → 3')	Backbone linkage	Dissoc. const. K_D (method used)	PNGase F treatment?	Oligomeric state of mTLR9–cECD
LIGPS-TAMRA	<u>TCGACGTTGACGTTGACGTT</u>	PS	$0.52 \pm 0.1 \mu\text{M}$ (FP)	No (AUC, FP)	Soluble aggregates
LIGPS	<u>TCGACGTTGACGTTGACGTT</u>	PS	$< 8.6 \mu\text{M}$ (FP)	Yes (AUC), No (AUC, FP)	Soluble aggregates
LIGPD	<u>TCGACGTTGACGTTGACGTT</u>	PD	$694 \pm 87 \text{ nM}$ (ITC)	Yes (AUC, ITC)	Dimeric
INHPS-TAMRA	TCCTGGCGGGGAAGT	PS	$1.43 \pm 0.16 \mu\text{M}$ (FP)	No (AUC, FP)	Soluble aggregates
INHPS	TCCTGGCGGGGAAGT	PS	ND	No (AUC)	Soluble aggregates
INHDP	TCCTGGCGGGGAAGT	PD	$1.25 \pm 0.14 \mu\text{M}$ (ITC)	Yes (AUC, ITC)	Dimeric
14PS	TAATGACGTTAATG	PS	ND	No (AUC)	Soluble aggregates
14PD	TAATGACGTTAATG	PD	No binding (ITC, AUC)	No (AUC, ITC)	NA
14PD _G	TAATGACGTTAATG	PD	No binding (ITC, AUC)	Yes (AUC, ITC)	NA
MethLIGPD	T(m)CGA(m)CGTTGA(m)CGTTGA(m)CGTT	PD	$730 \pm 90 \mu\text{M}$ (ITC)	Yes (AUC, ITC)	Dimeric
RanLIGPD	GTCAGACAGTCTGTGTGTC	PD	$2.3 \pm 0.2 \mu\text{M}$ (ITC)	Yes (AUC, ITC)	Dimeric
RevLIGPD	TGCAGCTTGACCTTGACGTT	PD	ND	Yes (AUC)	Dimeric
50LIGPD	<u>GCTTCTAGATCAATGGTTCGACGTTGACGTT</u> <u>GACGTTCTCCAGGTCAATC</u>	PD	Multiple binding sites (ITC, AUC)	Yes (AUC, ITC)	Mixture of dimeric and higher order oligomeric species
ds50LIGPD	<u>GCTTCTAGATCAATGGTTCGACGTTGACGTT</u> <u>TGACGTTCTCCAGGTCAATC</u>	PD	Multiple binding sites (ITC, AUC)	Yes (AUC, ITC)	Mixture of dimeric and higher order oligomeric species
601H2	GGTGCCGAGGACGTTCAATTGGACGTTGACA	PD	$140 \pm 46 \text{ nM}$ (FP)	Yes (AUC, FP)	Soluble aggregation

TAMRA, DNA is 3'-labelled with tetramethyl rhodamine; ODN, oligonucleotide; PS, phosphorothioate; PD, phosphodiester; (m)C, C5-methylated cytosine; FP, fluorescence anisotropy; ITC, isothermal titration calorimetry; AUC, sedimentation-velocity analytical ultracentrifugation; ND, not determined; NA, not available.

binding of full-length human TLR9 ectodomain to similar PD-ODNs in luminescent-proximity (AlphaScreen) assays (Haas *et al*, 2008). All binding affinities were measured at pH 5.5, which is similar to the pH of endolysosomal compartments. mTLR9–cECD was unstable in buffers with pH > 7. We note that TLR3 only binds tightly to dsRNA ligands at pH values similar to those found inside endolysosomal compartments (de Bouteiller *et al*, 2005; Leonard *et al*, 2008).

CpG motifs are largely methylated at the C5 position of the cytosine in mammalian DNA, but not bacterial DNA. Cytosine methylation abolishes the ability of PS-ODNs to activate TLR9 (Krieg, 2002), suggesting that TLR9 uses methylation to discriminate microbes from self. However, Haas *et al* (2008) have shown that with the addition of 3' poly-G extensions, PD-ODNs with methylated CpG motifs or no canonical CpG motifs at all induced TLR9-dependent IL-6 and IFN- α production by dendritic cells. 3' poly-G extensions reduce nuclease sensitivity (Bishop *et al*, 1996), promote multimerization through G-tetrads (GGGG sequences; Sen and Gilbert, 1988), and hence enhance endosomal uptake of PD-ODNs (Dalpke *et al*, 2002). We analysed binding of mTLR9–cECD to a cytosine-methylated PD-ODN ligand (methLIGPD) by ITC and found that methylation did not affect the binding affinity ($K_d = 730 \pm 90 \text{ nM}$) or stoichiometry of the complex (2:1 TLR9/DNA; Figure 2C).

G-tetrads can inhibit the stimulatory activity of CpG-containing (or non-CpG) PS-ODNs (Gursel *et al*, 2003; Ashman *et al*, 2005; Barrat *et al*, 2005; Duramad *et al*, 2005), but not PD-ODNs (Haas *et al*, 2008). A 15-nucleotide PD-ODN with a CGGGG sequence, INHPD, bound to mTLR9–cECD with a similar affinity ($K_d = 1.25 \pm 0.14 \mu\text{M}$) and stoichiometry (2:1 TLR9/DNA) as PD-ODNs lacking a G-tetrad (Figure 2D). Moreover, a 14-nucleotide PD-ODN with a canonical CpG sequence, 14PD, failed to bind, based on a lack of heat absorption or release in ITC. Thus, the minimum PD-ODN length for binding to mTLR9–cECD is 15 nucleotides. Together, these data support the hypothesis that PD DNA

backbones of at least 15 nucleotides activate TLR9 in a sequence-independent manner.

The ITC measurements show that DNA binding by TLR9 is enthalpy driven (Figure 2). However, in the case of LIGPD, ranLIGPD and methLIGPD, binding is accompanied by significant decreases in entropy, suggesting that the DNA backbone becomes more rigid upon binding to TLR9. Reduced flexibility of TLR9 amino-acid side chains or binding of water molecules or ions upon binding DNA may also contribute to the observed decrease in entropy. Unexpectedly, binding of INHPD to mTLR9–cECD causes an insignificant entropy decrease ($\Delta S = -2.6 \text{ kcal mol}^{-1}$). The enthalpy of binding is also much lower for INHPD than for the other three ligands (Figure 2). The smaller changes in entropy and enthalpy suggest that INHPD has an inherently rigid backbone, possibly due to its G-tetrad, and that this rigidity restricts the number of favourable contacts INHPD can form with mTLR9–cECD. Nevertheless, the combined effects of the lower entropy and enthalpy changes associated with INHPD binding fortuitously result in a similar binding affinity for INHPD as for LIGPD.

In the signal-transduction mechanism proposed for human TLR9, based on studies with PS-ODNs, ligand binding to a preformed TLR9 homodimer induces a conformational rearrangement, which is then transmitted across the membrane to appropriately orient and assemble the TIR domains (Latz *et al*, 2007). To determine how the oligomeric state of mTLR9–cECD is affected by PD-ODNs, we measured the sedimentation coefficient of mTLR9–cECD in the presence and absence of ligand by analytical ultracentrifugation. Binding of mTLR9–cECD to LIGPD or methLIGPD caused the sedimentation coefficient of mTLR9–cECD to increase from 2.3 to 4.1 S (Figure 3A and Supplementary Table I), which is consistent with 41-kD mTLR9–cECD monomers forming dimers as they bind a single PD-ODN molecule. Moreover, the ITC isotherm curves show clearly that mTLR9–cECD binds to PD-ODNs with a 2:1 (TLR9/DNA)

Figure 3 Binding to DNA changes the oligomeric state of mouse TLR9-cECD. Analytical ultracentrifugation analysis of binding of (A) deglycosylated TLR9-cECD (dgTLR9-cECD) to PD-ODN LIGPD and (B) glycosylated TLR9-cECD and dgTLR9-cECD to PS-ODN LIGPS. The sedimentation coefficient distributions $c(S)$ were calculated from sedimentation velocity experiments, in which 6- μ M TLR9-cECD was mixed with 3- μ M DNA and the absorbance at 260 nm was monitored. Peak A (1.7 S) corresponds to monomeric ligand. Peaks B (2.3 S for dgTLR9-cECD and 2.6 S for TLR9-cECD) and C (3.9 S for both proteins) correspond to monomeric and dimeric forms of TLR9-cECD, respectively. Peak D (4.1 S) corresponds to a complex containing two TLR9-cECD molecules and one LIGPD molecule. LIGPS induces the formation of large protein-DNA aggregates with sedimentation coefficients between 7 and 60 S (peak M). The peak at approximately 0.5 S that is visible in most $c(S)$ curves corresponds to prematurely truncated ODNs (or to other small-molecule contaminants) from the ODN synthesis reaction.

molar ratio (Figure 2). Together, our data show that single-stranded PD-ODNs induce dimerization of mTLR9-cECD monomers.

The binding properties of double-stranded PD-ODN ligands to mTLR9-cECD are slightly more complex than for single-stranded PD-ODNs. A 31-base pair double-stranded PD-ODN, 601H2, bound to mTLR9-cECD with a 4- to 16-fold higher affinity ($K_d = 140 \pm 46$ nM) than the single-stranded PD-ODNs (Table I). However, the sequence of 601H2 is based on one of the strongest known histone octamer binding and positioning sequences (Vasudevan *et al*, 2010). The higher binding affinity of 601H2 may therefore be due to a greater inherent propensity of 601H2 to adopt a curved conformation rather than to the double strand (see below). Analysis of 601H2 binding to mTLR9-cECD by ITC suggests a binding scheme with multiple binding sites (Supplementary Figure 1A). Analytical ultracentrifugation sedimentation profiles show that 601H2 induces mTLR9-cECD to form large aggregates (30-150S; Supplementary Figure 1B). The formation of higher order mTLR9-cECD oligomers is not surprising. On the basis of the 2:1 TLR9/DNA stoichiometry of complexes with 15-base pair single-stranded ligands, up to eight mTLR9-cECD might be expected to bind to 601H2, as 601H2 is twice as long as the single-stranded ligands and has two strands.

Phosphorothioate oligodeoxynucleotides cause mTLR9-cECD to form large aggregates

Ligand binding studies and sequence requirements for the activation of TLR9 by ssDNA have primarily been performed with synthetic PS-ODNs (Bauer *et al*, 2001; Krieg, 2002; Latz *et al*, 2007) to improve nuclease resistance (Krieg, 2002) and cellular uptake (Sester *et al*, 2000). Because of their increased biological half-life, PS-ODNs are widely used in therapeutic applications. To clarify the molecular basis of TLR9 activation by PS-ODNs, we measured the affinity and stoichiometry of binding of PS-ODN ligands to mTLR9-cECD using various biophysical approaches. First, the binding affinities for mTLR9-cECD of the CpG-containing PS-ODNs LIGPS and INHPS were measured by fluorescence anisotropy polarization (Table I). The dissociation constants were determined by titrating mTLR9-cECD into a solution containing PS-ODNs labelled at the 3' end with a tetramethylrhodamine (TAMRA) fluorophore. Curves were fitted as described in the Materials and methods, giving surprisingly weak K_d values of 0.52 ± 0.12 and 1.43 ± 0.16 μ M for labelled LIGPS and INHPS binding to mTLR9-cECD, respectively. A control for nonspecific binding using bovine serum albumen instead of TLR9-cECD showed no binding to TAMRA-labelled LIGPS or 601H2 (Supplementary Figure 2). In contrast, previous studies with uncleaved TLR9 ectodomain constructs have reported much greater binding affinities of PS-ODNs with CpG sequences, with K_d in the 0.5–5 nM range (Latz *et al*, 2007; Haas *et al*, 2008; Ashman *et al*, 2011). A recent measurement of TLR9 ligand binding in live cells by fluorescence cross-correlation spectroscopy also yielded tighter binding affinities, with K_d values of 62 and 153 nM for CpG-rich and non-CpG PS-ODNs, respectively (Chen *et al*, 2011).

ITC measurements with unlabeled LIGPS suggest that multiple ligand molecules ($n \geq 3$) bind sequentially to mTLR9-cECD in a complex binding scheme that requires multimerization of the receptor (Supplementary Figure 3A).

Figure 4 Binding of phosphorothioate oligodeoxynucleotides (PS-ODNs) causes loss of secondary structure in mTLR9-cECD. Circular dichroism spectra of mTLR9-cECD titrated with (A) LIGPS or (B) 14PS. The spectra show a decrease of intensity around 210 nm. (C, D) Circular dichroism spectra of mTLR9-cECD are unaffected by titration of LIGPD (C) or 14PD (D). MRE, mean residue molar ellipticity.

To investigate the binding mechanism of PS-ODNs to TLR9, specifically the effect of PS-ODNs on the oligomeric state of mTLR9-cECD, we measured the sedimentation coefficient of mTLR9-cECD in the presence and absence of PS-ODNs by analytical ultracentrifugation. We found that binding of LIGPS or INHPS to mTLR9-cECD induced the formation of very large and highly heterogeneous aggregates with sedimentation coefficients in the 7–40 S range (Figure 3B and Supplementary Table I). The large aggregates of mTLR9-cECD in complex with LIGPS could also be visualized directly by electron microscopy. Surprisingly, although the minimum length for binding of PD-ODNs is 15 nucleotides, a 14-nucleotide PS-ODN with a canonical CpG sequence (14PS) caused some aggregation of mTLR9-cECD (Supplementary Table I and Supplementary Figure 3B).

To complement the sedimentation studies, titrations of mTLR9-cECD with LIGPD or LIGPS were analysed by circular dichroism spectroscopy. Titration with PS-ODNs, including LIGPS and 14PS, reduced the intensity of the spectrum at 210 nm (and to a lesser extent between 202 and 232 nm; Figure 4A and B). This suggests that PS-ODN-induced aggregation of mTLR9-cECD is accompanied by a change in secondary structure, consistent with a study with uncleaved TLR9 (Latz *et al*, 2007). In contrast, PD-ODNs such as LIGPD or 14PD did not induce any changes in circular dichroism (Figure 4C and D). Together, our data suggest that PD-ODN-induced mTLR9-cECD dimerization and PS-ODN-induced mTLR9-cECD aggregation occur by distinct mechanisms.

TLR9 binds to supercoiled plasmid DNA, but the affinity of TLR9 for linearized plasmids is low (Cornelie *et al*, 2004; Latz

et al, 2004; Kindrachuk *et al*, 2007). We qualitatively assessed the interaction between supercoiled and linearized plasmid DNA and mTLR9-cECD using electrophoretic mobility shift assays. mTLR9-cECD bound to both supercoiled and linearized plasmids and binding was abolished by thermal denaturation of the protein. Moreover, PS-ODNs such as INHPS or LIGPS competed with the plasmid for mTLR9-cECD binding (Figure 5). The INHPS/plasmid molar ratio required to displace the plasmid from mTLR9-cECD was 1000:1 for supercoiled plasmid and only 10:1 for linearized plasmid, suggesting that the affinity of mTLR9-cECD for supercoiled plasmid is approximately 100-fold greater than for linearized plasmid.

The C-terminal cleavage fragment of TLR9 activates NF- κ B signalling

To confirm the biological relevance of our *in vitro* ligand binding data, we measured TLR9-dependent NF- κ B signalling responses in HEK293 human embryonic kidney cells stimulated with PS-ODN DNA ligands. TLR9-signalling activation was measured using an NF- κ B-dependent luciferase reporter gene. Cells transfected with the C-terminal cleavage fragment of TLR9 (residues 474–1032) produced specific signalling responses when stimulated with LIGPS or 14PS, demonstrating that the cleaved receptor is capable of producing a response in the absence of the N-terminal region of the ectodomain (residues 1–473). However, the levels of NF- κ B activation were significantly lower than for full-length TLR9, suggesting that residues 1–473 are required for maximal signalling activation (Figure 7; see Discussion).

Figure 5 mTLR9-cECD binds preferentially to supercoiled plasmid DNA. Electrophoretic mobility shift assays (EMSAs) containing 30 μ g of mTLR9-cECD and 300 ng of (A) supercoiled pET21a plasmid or (B) linearized pET21a plasmid. After heating at 95°C for 5 min, mTLR9-cECD no longer binds the plasmid (leftmost lane). INHPS ODN was titrated in at various concentrations until the plasmid was displaced from mTLR9-cECD. Unliganded mTLR9-cECD did not enter the gel (not shown). Most of the mTLR9-cECD/plasmid complexes migrated as a sharp band near the loading wells. A minority of the mTLR9-cECD/plasmid complexes, presumably those that aggregated or had more TLR9 bound, did not enter the gel and can be seen coating the loading wells, especially in the three central wells in B.

High-mobility group box 1 protein enhances mTLR9-cECD binding to DNA

Endogenous DNA in the form of nucleic acid-protein complexes released from apoptotic cells can sometimes stimulate TLR9 signalling that can be associated with autoimmune disease. HMGB1 enhances the recognition of certain ODNs (with partial PS backbones and 3' poly-G extensions) by TLR9 (Tian *et al*, 2007). Notably, histones H2A and H2B can act as autoantigens in SLE (Rekviig and Hannestad, 1980; Hardin and Thomas, 1983), and both HMGB1 and histones induce pronounced bending of DNA (Paull *et al*, 1993; Pil *et al*, 1993; Luger *et al*, 1997; Stott *et al*, 2006). We therefore hypothesized that the DNA curvature-inducing activities of these proteins may enhance DNA recognition by TLR9. In support of this hypothesis, we measured the binding affinity of mTLR9-cECD for preformed HMGB1-ODN complexes by fluorescence spectroscopy using TAMRA-labelled ODNs. The K_d values for mTLR9-cECD binding to HMGB1-LIGPS and HMGB1-INHPS were 23 ± 8 and 10 ± 8 nM, respectively (Figure 6A and B). HMGB1 therefore confers a 23-fold enhancement of the binding affinity of mTLR9-cECD for LIGPS and a 143-fold enhancement of the binding affinity of mTLR9-cECD for INHPS (Table I). For reference, the K_d values for HMGB1 binding to LIGPS and INHPS in the absence of mTLR9-cECD were 68 ± 37 and 60 ± 21 nM, respectively, consistent with previous measurements (Favichio *et al*, 2009). Together, our data suggest that HMGB1 significantly enhances mTLR9-cECD binding to DNA.

Core histones H2A-H2B enhance mTLR9-cECD binding to DNA

In order to test whether core histones, such as HMGB1, enhance TLR9-DNA binding, we measured binding of mTLR9-cECD to a 'quarter-nucleosome' consisting of a pre-formed complex of H2A, H2B and 601H2, a 31-base pair double-stranded DNA fragment with a PD backbone (Table I). To promote histone-DNA complex formation, we based the sequence of the DNA fragment on the H2A-H2B-specific region of '601' DNA, one of the strongest known histone octamer binding and positioning sequences

(Vasudevan *et al*, 2010). To promote TLR9 binding, we introduced into 601H2 two GACGTT motifs in positions predicted to be solvent-accessible based on the crystal structure of the nucleosome (Luger *et al*, 1997; Vasudevan *et al*, 2010). We measured the binding affinity of mTLR9-cECD for a preformed H2A-H2B-601H2 quarter-nucleosome by fluorescence spectroscopy using TAMRA-labelled 601H2. The K_d was 5 ± 3 nM (Figure 6C, D and Table I). The K_d for mTLR9-cECD binding to 601H2 in the absence of the histones was 140 ± 46 nM. The H2A-H2B heterodimer therefore enhances the affinity of mTLR9-cECD for 601H2 by a factor of 28. For reference, the K_d for H2A-H2B binding to 601H2 in the absence of mTLR9-cECD was 31 ± 20 nM. Notably, binding of mTLR9-cECD to both HMGB1-ODN and H2A-H2B-ODN complexes was best described by a competitive relationship between TLR9 and HMGB1 or H2A-H2B for DNA binding (the determination of binding constants is described in Materials and methods).

Discussion

A synthesis of the data we have presented here with previous TLR9 ligand-binding studies allows us to propose a new model of TLR9 activation (Figure 8). In the absence of ligand, the C-terminal region of the cleaved TLR9 ectodomain (TLR9-cECD) is a monomer, whereas uncleaved TLR9 forms homodimers with the TIR domains too far apart to produce a signal (Latz *et al*, 2007). This suggests that TLR9 dimerization in the absence of DNA is dependent on the N-terminal region of the ectodomain. We cannot, however, exclude the possibility that the transmembrane helix, the TIR domain or a cofactor present in HEK293 cells may also contribute to constitutive dimerization. Binding of PD-ODN ligands by TLR9-cECD induces the formation of dimers with the two protein molecules binding a single ligand molecule. The ligand-bound TLR9-cECD dimers presumably adopt a conformation with apposed C termini, as in the activated TLR9 receptor, in which the TIR domains are brought together by a ligand-induced conformational change (Latz *et al*, 2007).

The binding affinities of PD and PS oligonucleotides for TLR9-cECD are 25- to 50-fold weaker than for full-length

Figure 6 Biochemical characterization of the effect of DNA curvature-enhancing proteins HMGB1 and histones H2A-H2B on DNA binding by TLR9. Equilibrium binding of mTLR9-cECD to preformed complexes of recombinant mouse HMGB1 and (A) 3'-TAMRA-labelled INHPS, or (B) 3'-TAMRA-labelled LIGPS measured by fluorescence spectroscopy. The concentration of both ligands was 200 nM. Curves were fitted as described in the Materials and methods, giving K_d values of 10 ± 8 nM for mTLR9-cECD binding to HMGB1-INHPS-TAMRA and 23 ± 8 nM to HMGB1-LIGPS-TAMRA. Insets: titration of HMGB1 into 200 nM solutions of ligand. The K_d values for INHPS-TAMRA and LIGPS-TAMRA binding to HMGB1 are 60 ± 21 and 68 ± 37 nM, respectively. The arrows mark the concentration of HMGB1 (200 nM) that was selected for the experiments in the main panels. (C) Equilibrium binding of mTLR9-cECD to 3'-TAMRA-labelled 601H2 measured by fluorescence anisotropy polarization. The K_d for 601H2 binding to mTLR9-cECD is 140 ± 46 nM. (D) Equilibrium binding of mTLR9-cECD to a preformed complex of 601H2 and human histone H2A-H2B heterodimers measured by fluorescence anisotropy polarization. Curves were fitted as described in the Materials and methods, giving a K_d value of 5 ± 3 nM for mTLR9-cECD binding to H2A-H2B-601H2-TAMRA. Inset: titration of H2A-H2B into a 200 nM solution of 601H2-TAMRA. The K_d for H2A-H2B binding to 601H2-TAMRA is 31 ± 20 nM. The arrow marks the concentration of H2A-H2B (200 nM) that was selected for the experiment in the main panel.

TLR9. A contribution of the N-terminal region of the ectodomain to ligand binding would provide a logical explanation for this discrepancy. Various mutations within the N-terminal region of the ectodomain inactivate TLR9, implying that this region does indeed participate in ligand binding (Peter *et al*, 2009). Moreover, the C-terminal cleavage fragment of TLR9 produces significantly lower levels of NF- κ B activation than full-length TLR9, suggesting that residues 1–473 are required for maximal signalling activation (Figure 7). To reconcile a contribution of N-terminal ectodomain residues in DNA recognition with the requirement for proteolytic activation of the ectodomain (Ewald *et al*, 2008, 2011; Park *et al*, 2008), we propose that the N- and C-terminal ectodomain fragments remain associated after proteolytic cleavage in the endosome. Cleavage of the ectodomain may be necessary for the pre-formed TLR9 homodimer to undergo the ligand-induced conformational change that brings the TIR domains together and activates the receptor (Latz *et al*, 2007) (Figure 8). TLR3 and TLR7 are cleaved by the same endosomal proteases as TLR9 (Ewald *et al*, 2011), and the N-terminal region of TLR3 is known to contribute to RNA binding (Liu *et al*, 2008).

Thus, proteolytic activation may provide a general mechanism to regulate the activation of nucleic acid-sensing TLRs by structurally priming the receptors for the ligand-induced activating conformational change in the correct cellular compartment. A detailed structural analysis of TLR9 is needed to confirm this proposed mechanism of ligand binding and activation. We note that while our single-stranded PD-ODN binding data favour a mechanism in which TLR9 dimers transduce the signal across the membrane, the double-stranded 601H2 ligand does cause TLR9 aggregation. We therefore cannot rule out that TLR9 signalling depends on the formation of larger oligomers in the context of the full-length receptor *in vivo*. Multimerization has been reported for TLR5 (Mizel *et al*, 2003), and the TLR signalling adaptor MyD88 forms large oligomers during signalling (Lin *et al*, 2010).

Although the chemical specificity for CpG motifs or other patterns in microbial DNA remains controversial (Kindrachuk *et al*, 2007; Peter *et al*, 2009), there is mounting evidence that TLR9 can be activated in a sequence-independent manner (Barton *et al*, 2006; Haas *et al*, 2008; Ewald *et al*,

Figure 7 TLR9-dependent NF-κB signalling in cells stimulated with PS-ODN DNA ligands. HEK293 cells stably expressing an NF-κB-dependent luciferase reporter gene were transfected with the C-terminal TLR9 cleavage fragment (ΔTLR9, residues 474–1032), full-length TLR9 (FL-TLR9, residues 1–1032), water (Mock) or a noncoding plasmid (noncoding). The transfected cells were stimulated 48 h post-transfection with (A) LIGPS or (B) 14PS. Cellular signalling activation (vertical axis) is presented as mean luciferase units (MLU). Data are the mean ± s.d. of triplicate measurements representative of three experiments.

Figure 8 Proposed model for TLR9 activation. TLR9 (in light and dark blue; only the ectodomains are shown for clarity) forms dimers in the absence of ligand such that all of the dimer contacts are dependent on the N-terminal region of the ectodomain. TLR9 is cleaved near residue 477 in the endosome by asparagine endopeptidase (AEP) and cathepsins (cts). The two cleavage fragments remain associated. Cleavage of the ectodomain allows the preformed homodimer to undergo a ligand-induced conformational change that brings the TIR domains together and activates the receptor. Both the N- and C-terminal regions of the ectodomain participate in ligand binding (right panel). The DNA ligand (in black) binds in a bent or curved conformation. DNA curvature-inducing proteins including HMGB1 or core histones (in red) promote ligand binding by maintaining the ligand in a curved conformation. TLR9 may compete with the curvature-enhancing proteins and cause them to dissociate (not shown). Without proteolytic cleavage, the ectodomain is conformationally restricted and cannot fully engage with ligands or reach the activated conformation (bottom panel).

2011). Our work reinforces this notion, as PD-ODNs of at least 15 nucleotides induced dimerization of TLR9-cECD and bound with similar affinities, even when CpG motifs were methylated or absent. Thus, any specificity of TLR9 for CpG motifs, if it exists, would have to come from the N-terminal region of the ectodomain, or from a cofactor or co-receptor (Baumann *et al*, 2010).

Binding of PS-ODNs of any sequence to TLR9-cECD results in the formation of large aggregates and causes a large, specific and saturable loss in intensity in the far-UV circular dichroism spectrum. Both of these phenomena could be explained by a partial denaturation of the TLR9 ectodomain. The tendency of PS-ODNs to cause TLR9 aggregation even in the absence of G-tetrads or 3' poly-G extensions provides a novel explanation for why PS-ODNs induce greater TLR9-signalling responses than PD-ODNs *in vivo*. Together, our

data suggest that activation of TLR9 by PS-ODNs occurs by a distinct and more complex mechanism than TLR9 activation by natural DNA ligands.

HMGB1 enhances the recognition of certain ODNs by TLR9 by associating with the RAGE receptor and delivering HMGB1-DNA complexes into early endosomes (Tian *et al*, 2007). However, the molecular basis for the enhancement of TLR9 activation was unclear. We introduce here the hypothesis that TLR9 preferentially recognizes curved DNA backbones. The DNA curvature-inducing activity of HMGB1 (Paull *et al*, 1993; Pil *et al*, 1993; Stott *et al*, 2006) would then facilitate TLR9-ligand recognition. Indeed, HMGB1 enhances the binding affinity of TLR9-cECD for LIGPS and INHPS by 23- and 143-fold, respectively. The greater enhancement for INHPS may stem from the higher rigidity of INHPS, so that more work is required by HMGB1 to induce curvature in

INHPS than in LIGPS. The INHPS–HMGB1 complex may then bind to TLR9 without the entropic penalty shown by our ITC measurements to be associated with LIGPS binding. Also, we note that binding of TLR9–cECD to preformed HMGB1–ODN complexes resulted in a decrease in fluorescent intensity (Figure 6). This suggests that the fluorophore on the ODN may become more exposed to solvent on binding to TLR9, and hence that the ODN may be displaced from HMGB1 by TLR9.

Similar to HMGB1, histones induce bending of DNA (Luger *et al*, 1997), although the double helix is bent in different directions in the HMGB1–DNA and nucleosome complexes. The H2A–H2B core histone heterodimer also enhances the binding affinity of TLR9–cECD for a double-stranded PD-ODN ligand. Histones can act as autoantigens in SLE (Rekvis and Hannestad, 1980; Hardin and Thomas, 1983), and nucleosomes are thought to be one of the key endogenous nucleic acid–protein complexes that can induce TLR-dependent auto-immunogenic signalling during apoptosis, tissue injury or viral infections (Deane and Bolland, 2006). Our fluorescence binding assays are consistent with a competitive relationship between TLR9 and HMGB1 or H2A–H2B for DNA binding, suggesting that HMGB1 and H2A–H2B are displaced from the DNA upon TLR9 binding.

Apart from the enhancement of TLR9–DNA recognition by HMGB1 and H2A–H2B, three other lines of evidence indicate that TLR9 is selective for curved DNA ligands. First, TLR9–cECD shows a marked preference for supercoiled versus linearized plasmid. Second, binding of the relatively flexible LIGPS ligand to TLR9–cECD releases more enthalpic heat than binding of the more rigid INHPS ligand, suggesting that flexibility in the DNA backbone is required for optimal binding. Third, 601H2, a PD-ODN based on a strong histone octamer binding and positioning sequence bound to mTLR9–cECD with a higher affinity than LIGPD.

In conclusion, we have shown that ligand binding and ligand-induced dimerization of the cleaved TLR9 ectodomain is independent of the sequence and methylation state of DNA ligands. We have found that ligands with PS backbones cause TLR9 aggregation, which may be the primary reason for why PS-ODNs induce greater TLR9–signalling responses than biologically produced ODNs in many experimental systems. Moreover, we have presented evidence that protein-induced curvature of DNA ligands plays a previously unappreciated role in promoting ligand recognition by TLR9. Our results support the view that the ability of TLR9 to distinguish microbial DNA from self-DNA stems from the endosomal localization of the receptor combined with a structural requirement for proteolytic activation by endosomal proteases. The sequence and methylation state of DNA ligands, however, appear to be of lesser importance than previously assumed. Therapeutic strategies that specifically target DNA curvature-inducing proteins in endolysosomal compartments may offer a new approach to immune modulation for diseases associated with the nucleic acid–sensing TLRs.

Materials and methods

Oligonucleotides

All oligonucleotides were synthesized by Sigma-Aldrich, except the oligonucleotides with PS backbones and fluorescently labelled oligonucleotides, which were synthesized at the WM Keck

Biotechnology Resource Laboratory at Yale University or by Sigma-Aldrich. All fluorescent oligonucleotides were labelled with TAMRA at the 3' end. For 601H2, only one of the strands was fluorescently labelled. The oligonucleotide sequences are listed in Table I. We note that the sequences of LIGPD and LIGPS are identical to the stimulatory CpG DNA 2007 used in other studies (e.g., Latz *et al*, 2007), except that the context of the three CpG motifs was changed from GTCGTT to GACGTT because of the greater responsiveness of mouse TLR9 to the latter sequence (Bauer *et al*, 2001). INHPS is identical to inhibitory CpG ODN 2088.

Expression and purification of truncated mouse TLR9 ectodomain (mTLR9–cECD)

A gene encoding the proteolytically processed fragment (residues 474–824) of mouse TLR9 ectodomain (mTLR9–cECD) with an N-terminal eight-histidine purification tag, followed by the linker sequence Ser-Ser-Gly and a tobacco etch virus protease cleavage site (ENLYFQGP), was cloned into the pAcGP67-A vector (BD Biosciences) in frame with the baculovirus gp67 signal sequence. Sf9 insect cells (Invitrogen) were co-transfected with the mTLR9–cECD expression construct and Diamond Bac baculovirus genomic DNA (Sigma-Aldrich) to produce a recombinant baculovirus expressing mTLR9–cECD. Virus stocks were amplified with three sequential infections of Sf9 cells. For mTLR9–cECD expression, Tni insect cells (Expression Systems) were infected with 1% (vol/vol) of third-passage (P3) baculovirus stock. After culture in suspension for 72 h at 27°C, mTLR9–cECD was extracted from intracellular compartments with 50 mM Tris, pH 7.5, 500 mM NaCl, 10% glycerol, 5 mM β -mercaptoethanol, 1% Fos-Choline 12. mTLR9–cECD was purified by nickel affinity chromatography with a HisTrap HP column (GE Healthcare), followed by cation exchange chromatography with a MonoS column (GE Healthcare) and size-exclusion chromatography with a Superdex 200 10/300 GL column (GE Healthcare). The size-exclusion buffer was 20 mM MES, pH 5.6, 100 mM NaCl, 2 mM β -mercaptoethanol. This procedure typically yielded > 1 mg of pure mTLR9–cECD per litre of insect cell culture. Optionally, mTLR9–cECD was deglycosylated by treatment with PNGase F (New England Biolabs) at 25°C for 12 h, between the nickel affinity and cation exchange chromatography steps. Typical size-exclusion chromatograms of glycosylated and deglycosylated mTLR9–cECD with the corresponding SDS–PAGE gels are shown in Figure 1.

Expression and purification of mouse HMGB1

Mouse *hmgb1* was cloned into the pET28a vector (Novagen) using the *NcoI* and *BamHI* restriction sites. HMGB1 was expressed in *Escherichia coli* strain Rosetta (DE3) (Novagen) by induction at OD₆₀₀ = 0.6–0.8 with 1.0 mM isopropyl- β -D-thiogalactoside for 4 h at 37°C. The protein was purified by anion exchange chromatography on a Hitrap Q column (GE Healthcare), followed by size-exclusion chromatography on a Superdex 200 10/300 GL column in 50 mM Tris, pH 8.0, 50 mM NaCl, 3 mM β -mercaptoethanol.

Intrinsic fluorescence and fluorescence anisotropy polarization

To determine the binding affinities of various DNA oligonucleotides (ODNs) for mTLR9–cECD and/or HMGB1, mTLR9–cECD was titrated from a 0.2-mM stock solution into a 0.2 μ M solution of 3'-TAMRA-labelled LIGPS or INHPS ODN (see Table I for sequences) for final mTLR9–cECD concentrations between 20 nM and 10 μ M in 20 mM MES, pH 5.6, 100 mM NaCl, 2 mM β -mercaptoethanol. Binding studies were performed at 25°C. The intrinsic fluorescence or fluorescence anisotropy depolarization was recorded after 10 min equilibration with a PTI Quantamaster C-61 two-channel fluorescence spectrophotometer. The TAMRA fluorophore was excited at 555 nm and the emission was recorded at 575 nm, with a 6-nm slit width.

Determination of equilibrium binding constants

Fluorescence anisotropy was determined as described (Grove *et al*, 2010). Binding constants of HMGB1, human H2A–H2B (New England Biolabs) and mTLR9–cECD were determined with Origin 7 (OriginLab) by fitting the data to Equation 1 (Inglesse *et al*, 1989):

$$S_{\text{Tot}} = S_L + (S_{\text{EL}} - S_L) \times$$

$$\left[\frac{K_d + [L]_T + [E]_T - \sqrt{(K_d + [L]_T + [E]_T)^2 - 4[L]_T[E]_T}}{2[L]_T} \right] \quad (1)$$

where S_{Tot} is the signal (fluorescence intensity or fluorescence anisotropy) of ODN at protein concentration $[E]_T$; S_L is the signal of labelled ODN, L ; S_{EL} is the signal of ODN in the plateau region of the binding curve; $[L]_T$ is the total concentration of ODN added; K_d is the dissociation constant. The data were initially fit to three parameters $[L]_T$, S_{EL} and K_d using the measured values for S_L , S_{Tot} and $[E]_T$. The overall fit was selected such that the values for $[L]_T$ and S_{EL} coincided with the measured values for these parameters.

Determination of binding constants for mTLR9-cECD to HMGB1-ODN and H2A-H2B-ODN complex were determined by fitting the data to the following Equation (2) (Nolen and Pollard, 2008):

$$S_{\text{Tot}} = S_L + (S_{\text{EL}} - S_L) \times \left[\frac{-K_2K_1 - 2K_2[A]_T - K_1[X]_T + \sqrt{K_2^2K_1^2 + 2K_2K_1^2[X]_T + 2K_2K_1^2[A]_T + K_1^2[X]_T^2 - K_1^2[X]_T[A]_T + K_1^2[A]_T^2}}{2(-K_2K_1 - K_2[A]_T - K_1[X]_T + K_1^2 + K_1[A]_T)} \right] \quad (2)$$

where S_{Tot} is the signal (fluorescence intensity or fluorescence anisotropy) of TAMRA-ODN at mTLR9-trcECD concentration $[X]_T$; S_L is the signal of TAMRA-ODN in the plateau region of the binding curve, S_{EL} is the signal of TAMRA-ODN bound to HMGB1 or H2A-H2B, $[A]_T$ is the total concentration of HMGB1 or H2A-H2B, K_1 is the equilibrium dissociation constant of TAMRA-ODN for HMGB1 or H2A-H2B, K_2 is the equilibrium dissociation constant of TAMRA-ODN for mTLR9-cECD.

Isothermal titration calorimetry

The binding of oligonucleotides to mTLR9-cECD was analysed by ITC in 10 mM MES, pH 5.6, 100 mM NaCl, 2 mM β -mercaptoethanol at 25°C, with an iTC200 system (MicroCal). The sample cell contained 20 μ M mTLR9-cECD or buffer only and the syringe was loaded with 200 μ M oligonucleotide. An initial injection of 1.5 μ l of oligonucleotide was followed by 19 serial injections of 2.0- μ l oligonucleotide, each at 3-min intervals. The stirring speed was maintained at 1000 r.p.m. and the reference power was kept constant at 11 μ cal s⁻¹. The net heat absorption or release associated with each injection was calculated by subtracting the heat associated with the injection of oligonucleotide to buffer. Thermodynamic parameters were extracted from a curve fit to a single-site model with Origin 7.0 (MicroCal). The experiments were performed in triplicate.

Circular dichroism

In order to assess the secondary structure of ODN/mTLR9-cECD complexes, oligonucleotides were titrated into a solution of 3.75- μ M mTLR9-cECD in 20 mM MES, pH 5.6, 100 mM NaCl, 2 mM β -mercaptoethanol (or buffer only). Circular dichroism measurements were performed with a Chirascan CD spectrophotometer (Applied Photophysics). Each measurement was the average of five scans in steps of 0.5 nm at 20°C. For ODN/mTLR9-cECD complexes, the differential circular dichroism spectrum of the complex was obtained by subtracting the spectrum of the oligonucleotide alone (at the same concentration as in the complex) from the spectrum of the complex before the conversion to mean residue molar ellipticity values.

Analytical ultracentrifugation

To determine oligomeric states of mTLR9-cECD alone and in complex with various oligonucleotides, we measured the sedimentation velocities of the protein and its complexes with an Optima XL-I analytical ultracentrifuge equipped with both absorbance and interference optical detection systems (Beckman-Coulter). Purified mTLR9-cECD was dialysed overnight into the reference buffer, 20 mM MES, pH 5.6, 100 mM NaCl, 2 mM β -mercaptoethanol. The assembled cells with sample and reference buffer, and the rotor were temperature equilibrated at 20°C in the centrifuge for 1 h prior to each run. The protein solution was centrifuged at 20°C and 42 000 r.p.m. in an An-60 Ti rotor (Beckman-Coulter). Absorption at 260 nm and fringe displacement in interference optics were measured for 6–10 h. Data were fitted to a c(s) sedimentation coefficient distribution, which was calculated with Sedfit 12.1 (Brown *et al*, 2009) and regularized using the maximum entropy method with a value of 0.7. Values for protein vbar (0.739 ml g⁻¹), buffer density (1.00391 g ml⁻¹) and viscosity (0.01002 P) were calculated with SEDNTERP.

tation velocities of the protein and its complexes with an Optima XL-I analytical ultracentrifuge equipped with both absorbance and interference optical detection systems (Beckman-Coulter). Purified mTLR9-cECD was dialysed overnight into the reference buffer, 20 mM MES, pH 5.6, 100 mM NaCl, 2 mM β -mercaptoethanol. The assembled cells with sample and reference buffer, and the rotor were temperature equilibrated at 20°C in the centrifuge for 1 h prior to each run. The protein solution was centrifuged at 20°C and 42 000 r.p.m. in an An-60 Ti rotor (Beckman-Coulter). Absorption at 260 nm and fringe displacement in interference optics were measured for 6–10 h. Data were fitted to a c(s) sedimentation coefficient distribution, which was calculated with Sedfit 12.1 (Brown *et al*, 2009) and regularized using the maximum entropy method with a value of 0.7. Values for protein vbar (0.739 ml g⁻¹), buffer density (1.00391 g ml⁻¹) and viscosity (0.01002 P) were calculated with SEDNTERP.

Electrophoretic mobility shift assays

To assay for binding of mTLR9-cECD to oligonucleotide or plasmid DNA ligands, 67 μ M of protein, 300 ng of supercoiled pET21a plasmid or linearized pET21a plasmid, and various concentrations of INHPS were mixed in 10 mM MES, pH 5.6, 50 mM NaCl, 2 mM β -mercaptoethanol, and incubated at 25°C for 1 h. The mixtures were analysed by electrophoresis on 1% agarose gels in 1 $\times$ TBE buffer. Following electrophoresis, the gel was stained with SYBR Green (Molecular Probes) to visualize DNA.

NF- κ B-dependent luciferase reporter TLR9-signalling assay

HEK293 human embryonic kidney cells stably expressing the NF- κ B reporter gene pELAM-luciferase (a kind gift from Ruslan Medzhitov, Yale School of Medicine) were seeded at a density of 10⁶ cells per well in six-well plates. The cells were transfected the following day with pcDNA expression constructs containing full-length mouse TLR9 or the C-terminal cleavage fragment of TLR9 (residues 474–1032). After 48 h, the transfected cells were stimulated with increasing amount of LIGPS or 14PS (or no stimulus as a negative control). In the mock control, cells were transfected with the baculovirus transfer vector pAcgp67-A (BD Bioscience). Sixteen hours after stimulation, cell lysates were prepared in passive lysis buffer and luciferase activity was recorded with the luciferase assay system (Promega).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

This work was supported by a Novel Research Project grant from the Lupus Research Institute to YM, grant 233194878 from the Roche Organ Transplant Research Foundation to YM, contract W81XWH-07-1-0516 from the US Army Medical Research and Materiel Command to YM, and a Burroughs Wellcome Investigator in the Pathogenesis of Infectious Disease award to YM.

Author contributions: YL performed the cloning, expression and purification of TLR9 and HMGB1 proteins. YL performed all biochemical and ligand-binding experiments except analytical ultracentrifugation experiments, which were performed by ICB. YM devised the study. YM and YL wrote the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Ahmad-Nejad P, Hacker H, Rutz M, Bauer S, Vabulas RM, Wagner H (2002) Bacterial CpG-DNA and lipopolysaccharides activate Toll-like receptors at distinct cellular compartments. *Eur J Immunol* **32**: 1958–1968
- Akashi-Takamura S, Miyake K (2008) TLR accessory molecules. *Curr Opin Immunol* **20**: 420–425

- Alexopoulou L, Holt AC, Medzhitov R, Flavell RA (2001) Recognition of double-stranded RNA and activation of NF- κ B by Toll-like receptor 3. *Nature* **413**: 732–738
- Ashman RF, Goeken JA, Drahos J, Lenert P (2005) Sequence requirements for oligodeoxynucleotide inhibitory activity. *Int Immunol* **17**: 411–420

- Ashman RF, Goeken JA, Latz E, Lenert P (2011) Optimal oligonucleotide sequences for TLR9 inhibitory activity in human cells: lack of correlation with TLR9 binding. *Int Immunol* **23**: 203–214
- Barrat FJ, Meeker T, Gregorio J, Chan JH, Uematsu S, Akira S, Chang B, Duramad O, Coffman RL (2005) Nucleic acids of mammalian origin can act as endogenous ligands for Toll-like receptors and may promote systemic lupus erythematosus. *J Exp Med* **202**: 1131–1139
- Barton GM, Kagan JC, Medzhitov R (2006) Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nat Immunol* **7**: 49–56
- Bauer S, Kirschning CJ, Hacker H, Redecke V, Hausmann S, Akira S, Wagner H, Lipford GB (2001) Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition. *Proc Natl Acad Sci USA* **98**: 9237–9242
- Baumann CL, Aspalter IM, Sharif O, Pichlmair A, Bluml S, Gebrien F, Bruckner M, Pasierbek P, Aumayr K, Planyavsky M, Bennett KL, Colinge J, Knapp S, Superti-Furga G (2010) CD14 is a coreceptor of Toll-like receptors 7 and 9. *J Exp Med* **207**: 2689–2701
- Bishop JS, Guy-Caffey JK, Ojwang JO, Smith SR, Hogan ME, Cossum PA, Rando RF, Chaudhary N (1996) Intramolecular G-quartet motifs confer nuclease resistance to a potent anti-HIV oligonucleotide. *J Biol Chem* **271**: 5698–5703
- Botos I, Segal DM, Davies DR (2011) The structural biology of Toll-like receptors. *Structure* **19**: 447–459
- Brown PH, Balbo A, Schuck P (2009) On the analysis of sedimentation velocity in the study of protein complexes. *Eur Biophys J* **38**: 1079–1099
- Chen J, Nag S, Vidi PA, Irudayaraj J (2011) Single molecule *in vivo* analysis of toll-like receptor 9 and CpG DNA interaction. *PLoS One* **6**: e17991
- Christensen SR, Kashgarian M, Alexopoulou L, Flavell RA, Akira S, Shlomchik MJ (2005) Toll-like receptor 9 controls anti-DNA autoantibody production in murine lupus. *J Exp Med* **202**: 321–331
- Christensen SR, Shupe J, Nickerson K, Kashgarian M, Flavell RA, Shlomchik MJ (2006) Toll-like receptor 7 and TLR9 dictate autoantibody specificity and have opposing inflammatory and regulatory roles in a murine model of lupus. *Immunity* **25**: 417–428
- Cornelie S, Hoebeke J, Schacht AM, Bertin B, Vicogne J, Capron M, Riveau G (2004) Direct evidence that toll-like receptor 9 (TLR9) functionally binds plasmid DNA by specific cytosine-phosphate-guanine motif recognition. *J Biol Chem* **279**: 15124–15129
- Dalpe AH, Zimmermann S, Albrecht I, Heeg K (2002) Phosphodiester CpG oligonucleotides as adjuvants: polyguanosine runs enhance cellular uptake and improve immunostimulatory activity of phosphodiester CpG oligonucleotides *in vitro* and *in vivo*. *Immunology* **106**: 102–112
- de Bouteiller O, Merck E, Hasan UA, Hubac S, Benguigui B, Trinchieri G, Bates EE, Caux C (2005) Recognition of double stranded RNA by human toll like receptor 3 and downstream receptor signaling requires multimerisation and an acidic pH. *J Biol Chem* **280**: 38133–38145
- Deane JA, Bolland S (2006) Nucleic acid-sensing TLRs as modifiers of autoimmunity. *J Immunol* **177**: 6573–6578
- Diebold SS, Kaisho T, Hemmi H, Akira S, Reis e Sousa C (2004) Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **303**: 1529–1531
- Ding C, Wang L, Al-Ghawi H, Marroquin J, Mamula M, Yan J (2006) Toll-like receptor engagement stimulates anti-sRNP autoreactive B cells for activation. *Eur J Immunol* **36**: 2013–2024
- Duramad O, Fearon KL, Chang B, Chan JH, Gregorio J, Coffman RL, Barrat FJ (2005) Inhibitors of TLR-9 act on multiple cell subsets in mouse and man *in vitro* and prevent death *in vivo* from systemic inflammation. *J Immunol* **174**: 5193–5200
- Ehlers M, Fukuyama H, McGaha TL, Aderem A, Ravetch JV (2006) TLR9/MyD88 signaling is required for class switching to pathogenic IgG2a and 2b autoantibodies in SLE. *J Exp Med* **203**: 553–561
- Ewald SE, Engel A, Lee J, Wang M, Bogoy M, Barton GM (2011) Nucleic acid recognition by Toll-like receptors is coupled to stepwise processing by cathepsins and asparagine endopeptidase. *J Exp Med* **208**: 643–651
- Ewald SE, Lee BL, Lau L, Wickliffe KE, Shi GP, Chapman HA, Barton GM (2008) The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. *Nature* **456**: 658–662
- Favicchio R, Dragan AI, Kneale GG, Read CM (2009) Fluorescence spectroscopy and anisotropy in the analysis of DNA-protein interactions. *Methods Mol Biol* **543**: 589–611
- Grove TZ, Hands M, Regan L (2010) Creating novel proteins by combining design and selection. *Protein Eng Des Sel* **23**: 449–455
- Gursel I, Gursel M, Yamada H, Ishii KJ, Takeshita F, Klinman DM (2003) Repetitive elements in mammalian telomeres suppress bacterial DNA-induced immune activation. *J Immunol* **171**: 1393–1400
- Haas T, Metzger J, Schmitz F, Heit A, Muller T, Latz E, Wagner H (2008) The DNA sugar backbone 2' deoxyribose determines toll-like receptor 9 activation. *Immunity* **28**: 315–323
- Hardin JA, Thomas JO (1983) Antibodies to histones in systemic lupus erythematosus: localization of prominent autoantigens on histones H1 and H2B. *Proc Natl Acad Sci USA* **80**: 7410–7414
- Heil F, Hemmi H, Hochrein H, Ampenberger F, Kirschning C, Akira S, Lipford G, Wagner H, Bauer S (2004) Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* **303**: 1526–1529
- Inglese J, Blatchly RA, Benkovic SJ (1989) A multisubstrate adduct inhibitor of a purine biosynthetic enzyme with a picomolar dissociation constant. *J Med Chem* **32**: 937–940
- Janeway Jr CA, Medzhitov R (2002) Innate immune recognition. *Annu Rev Immunol* **20**: 197–216
- Jin MS, Kim SE, Heo JY, Lee ME, Kim HM, Paik SG, Lee H, Lee JO (2007) Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **130**: 1071–1082
- Kawai T, Akira S (2010) The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol* **11**: 373–384
- Kindrachuk J, Potter JE, Brownlie R, Ficzy AD, Griebel PJ, Mookherjee N, Mutwiri GK, Babiuk LA, Napper S (2007) Nucleic acids exert a sequence-independent cooperative effect on sequence-dependent activation of Toll-like receptor 9. *J Biol Chem* **282**: 13944–13953
- Krieg AM (2002) CpG motifs in bacterial DNA and their immune effects. *Annu Rev Immunol* **20**: 709–760
- Krieg AM, Yi AK, Matson S, Waldschmidt TJ, Bishop GA, Teasdale R, Koretzky GA, Klinman DM (1995) CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature* **374**: 546–549
- Latz E, Schoenemeyer A, Visintin A, Fitzgerald KA, Monks BG, Knetter CF, Lien E, Nilsen NJ, Espevik T, Golenbock DT (2004) TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nat Immunol* **5**: 190–198
- Latz E, Verma A, Visintin A, Gong M, Sirois CM, Klein DC, Monks BG, McKnight CJ, Lamphier MS, Duprex WP, Espevik T, Golenbock DT (2007) Ligand-induced conformational changes allosterically activate Toll-like receptor 9. *Nat Immunol* **8**: 772–779
- Leonard JN, Ghirlando R, Askins J, Bell JK, Margulies DH, Davies DR, Segal DM (2008) The TLR3 signaling complex forms by cooperative receptor dimerization. *Proc Natl Acad Sci USA* **105**: 258–263
- Lin SC, Lo YC, Wu H (2010) Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. *Nature* **465**: 885–890
- Liu L, Botos I, Wang Y, Leonard JN, Shiloach J, Segal DM, Davies DR (2008) Structural basis of toll-like receptor 3 signaling with double-stranded RNA. *Science* **320**: 379–381
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260
- Matsumoto M, Funami K, Tanabe M, Oshiumi H, Shingai M, Seto Y, Yamamoto A, Seta T (2003) Subcellular localization of Toll-like receptor 3 in human dendritic cells. *J Immunol* **171**: 3154–3162
- Mizel SB, West AP, Hantgan RR (2003) Identification of a sequence in human toll-like receptor 5 required for the binding of Gram-negative flagellin. *J Biol Chem* **278**: 23624–23629
- Nishiya T, Kajita E, Miwa S, Defranco AL (2005) TLR3 and TLR7 are targeted to the same intracellular compartments by distinct regulatory elements. *J Biol Chem* **280**: 37107–37117
- Nolen BJ, Pollard TD (2008) Structure and biochemical properties of fission yeast Arp2/3 complex lacking the Arp2 subunit. *J Biol Chem* **283**: 26490–26498
- Park B, Brinkmann MM, Spooner E, Lee CC, Kim YM, Ploegh HL (2008) Proteolytic cleavage in an endolysosomal compartment is required for activation of Toll-like receptor 9. *Nat Immunol* **9**: 1407–1414

- Park BS, Song DH, Kim HM, Choi BS, Lee H, Lee JO (2009) The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* **458**: 1191–1195
- Paull TT, Haykinson MJ, Johnson RC (1993) The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. *Genes Dev* **7**: 1521–1534
- Peter ME, Kubarenko AV, Weber AN, Dalpke AH (2009) Identification of an N-terminal recognition site in TLR9 that contributes to CpG-DNA-mediated receptor activation. *J Immunol* **182**: 7690–7697
- Pil PM, Chow CS, Lippard SJ (1993) High-mobility-group 1 protein mediates DNA bending as determined by ring closures. *Proc Natl Acad Sci USA* **90**: 9465–9469
- Rekvig OP, Hannestad K (1980) Human autoantibodies that react with both cell nuclei and plasma membranes display specificity for the octamer of histones H2A, H2B, H3, and H4 in high salt. *J Exp Med* **152**: 1720–1733
- Sen D, Gilbert W (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**: 364–366
- Sester DP, Naik S, Beasley SJ, Hume DA, Stacey KJ (2000) Phosphorothioate backbone modification modulates macrophage activation by CpG DNA. *J Immunol* **165**: 4165–4173
- Stott K, Tang GS, Lee KB, Thomas JO (2006) Structure of a complex of tandem HMG boxes and DNA. *J Mol Biol* **360**: 90–104
- Tian J, Avalos AM, Mao SY, Chen B, Senthil K, Wu H, Parroche P, Drabic S, Golenbock D, Sirois C, Hua J, An LL, Audoly L, La Rosa G, Bierhaus A, Naworth P, Marshak-Rothstein A, Crow MK, Fitzgerald KA, Latz E *et al* (2007) Toll-like receptor 9-dependent activation by DNA-containing immune complexes is mediated by HMGB1 and RAGE. *Nat Immunol* **8**: 487–496
- Vasudevan D, Chua EY, Davey CA (2010) Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J Mol Biol* **403**: 1–10
- Wagner H (2008) The sweetness of the DNA backbone drives Toll-like receptor 9. *Curr Opin Immunol* **20**: 396–400
- Yasuda K, Yu P, Kirschning CJ, Schlatter B, Schmitz F, Heit A, Bauer S, Hochrein H, Wagner H (2005) Endosomal translocation of vertebrate DNA activates dendritic cells via TLR9-dependent and -independent pathways. *J Immunol* **174**: 6129–6136