Single stranded DNA immune modulators with unmethylated CpG motifs: structure and molecular recognition by Toll-like Receptor 9

Krisztina Fehér[*a,b]

[a]*Department of Inorganic and Analytical Chemistry, Faculty of Natural Sciences and Technology, Debrecen University, Debrecen, Hungary;* [b]*Heidelberg Institute for Theoretical Studies, Heidelberg, Germany*

Abstract: Single stranded microbial DNA fragments with unmethylated deoxycytidylyl-deoxyguanosine dinucleotide (CpG) motifs are interpreted as danger signals by the innate immune system via recognition on Toll-like Receptor 9 (TLR9). Their synthetic analogues, Oligodeoxynucleotides (ODN) comprise a promising class of immune modulators with potential applications in the treatment of multiple diseases, such as cancer, autoimmune diseases or allergy. ODN molecules contain a core hexamer sequence, which is species specific consisting of GACGTT and AACGT for mouse and GTCGTT in humans. Assessment of structural features of different type of ODNs is highly challenging. NMR spectroscopic insights were gained for a short, single CpG motif containing ODN 1668. The structural basis of ODN recognition by TLR9 recently started to unravel as crystal structures of TLR9 orthologues in complex with ODN 1668 were solved. Systematic investigations of ODN sequences revealed that ODNs with a single CpG motif are capable of activating mouse TLR9, but two closely positioned CpG motifs are necessary for activation of human TLR9. Furthermore, longer ODNs with TCC and TCG sequences at the 5' end were shown to activate TLR9 with higher efficiency. It was revealed that 5'-xCx motif containing short ODNs (sODN) are able to augment the immune response of short, single CpG containing ODNs, which are incapable of activating of TLR9 alone. All these observations pointed to the existence of a second binding site on TLR9, which was characterized in crystal structures that delivered further insights of the nucleic acid recognition of the innate immune system by TLR9.

1. INTRODUCTION

The modulation of the immune system activity is a required pharmacological intervention for many medical conditions bearing particular importance in the treatment of inflammatory diseases, autoimmune diseases, in transplant procedures, infectious diseases and in developing new vaccines against infections and cancer [1]. Mechanisms leading to immunomodulation may involve either the innate or the adaptive immune system, the underlying molecular players are, however, distinctly different. The innate immune system recognizes microbes [2] by means of Pattern Recognition Receptors (PRR), which are conserved in evolution to bind unique molecular signatures of microbial metabolism not produced by the host organism. These highly conserved structural motifs are known as Pathogen-Associated Molecular Patterns (PAMP) and are used as templates for the development of novel therapeutic immune modulators.

*Address correspondence to this author at the Department of Inorganic and Analytical Chemistry, Debrecen University, 4032 Egyetem tér 1. Debrecen, Hungary; Tel: +36-52-512-900; E-mails: feher.krisztina@science.unideb.hu, krisztina.feher@h-its.org

Among PAMPs of microbial origin, CpG motif containing bacterial or viral single stranded DNA (ssDNA) fragments are recognized as unique microbial danger signals by TLR9 [3]], which belongs to an important family of PRRs of the innate immunity. TLRs [4] are critical for effective host defenses against numerous pathogens being able to link innate immunity to adaptive immunity. TLR9 [5] belongs to the subgroup of TLRs, which sense nucleic acids and is localized intracellularly in the endosomal compartment. It is expressed in B cells, plasmacytoid dendritic cells (pDC) and monocyte-macrophage-lineage cells in humans [6] and in macrophages and bone marrow-derived dendritic cells in mouse [7]. Activation of TLR9 triggers the MyD88-dependent signaling cascade that induces transcription factors (NF-κβ, AP1, IRF7) in the nucleus initiating the production of inflammatory cytokines and chemokines. TLR9 activation leads to Th1 dominated immune responses, inducing B-cell-proliferation, maturation of pDCs and the secretion of cytokines [8].

Since ssDNA fragments in the endosomes are a results of degradation of large double stranded DNA (dsDNA) of bacterial or viral origin, activation of immune responses also requires processing [9] of larger microbial dsDNA as well as cellular and endosomal uptake [10] [11].

Synthetic analogues of bacterial and viral DNA, CpG ODNs are important immune modulators and represent the most advanced potential adjuvants [12]. CpG ODNs are able to activate the innate immunity by binding to TLR9 and trigger Th1-dominated immune responses. Chemically, CpG ODNs are synthetic ssDNA sequences with unmethylated CpG-motifs. When used as vaccine adjuvants, they improve the function of professional antigen-presenting cells and boost the generation of humoral and cellular vaccine-specific

immune responses [5]. Due to their Th1 response inducing ability, CpG ODNs are in clinical testing as adjuvants in cancer immunotherapy [13] [14].

Conformational properties and determinants of receptor recognition are important in understanding the structural requirements for receptor activation and inhibition, for cellular and endosomal uptake and for formulation of delivery systems for CpG ODN adjuvants in vaccines, thus, in general for improving the therapeutic use of CpG ODNs. In this mini-review, an overview of studies investigating the sequences, structural features and the molecular determinants of receptor recognition of CpG ODNs will be given.

1.1. CpG ODN sequences and structure activity relationships

The immune stimulatory properties of CpG ODNs is dependent on the length of the sequence, the number of CpG motifs and the residues flanking the CpG motif, the distance between the motifs and the resulting structural properties [15]. In addition sequence selectivity of TLR9 recognition is also species specific.

The specific sequence motif recognized by TLR9 consists of an unmethylated CpG dinucleotide flanked by two 5' purines and two 3' pyrimidines with the general formula of RRCGYY, where R is a purine containing nucleotide and Y is pyrimidine based. In mice, the core hexamer is GACGTT and AACGT, while in humans GTCGTT is the optimal minimal immune stimulatory sequence [16].

There are three main classes and one subclass of CpG ODNs described in primates as shown in Figure 1 [17].

A type CpG ODNs (former D type) contain a single hexameric purine / pyrimidine / CG / purine / pyrimidine motif flanked by self-complementary bases capped at the 3' end by a poly G tail [18] as shown in Figure 1. A type CpG ODNs are poor stimulators of B cells and only modestly induce pDC maturation inducing secretion of type I IFNs (IFN-α/β), TNF-α, IL-12, and IFN-γ–IP10 [19-23].

B type CpG ODNs (former K type) contain multiple isolated CpG motifs (depicted in Figure 1) with phosphorothiolate bond in place of the phosphodiester linkage. The phosphorothiolate bond decreases the susceptibility to DNase digestion and improves uptake resulting in longer *in vivo* half-life/bioavailability than phosphodiester bond containing ODNs, however, it was reported recently that the phosphodiester bond is strongly favored by TLR9 [24]. B type CpG ODNs are the most stimulatory among the classes with being poor inducers of IFN-α with the ability to induce robust maturation of pDC and strong B-cell proliferation [3, 25].

C type ODNs contain a TCGTCG sequence at the 5' end and include a CG-rich palindrome (shown in Figure 1). C type CpG ODNs combine the immune stimulatory properties of A and B type ODNs [26-28] by being able to directly stimulating B cells to secrete IL-6 and pDCs to produce IFN-α.

P-class CpG ODNs were developed from C type CpG ODNs [29] and are a variation of thereof. They contain two palindromes as depicted in Figure 1 and induce efficient INF-α responses.
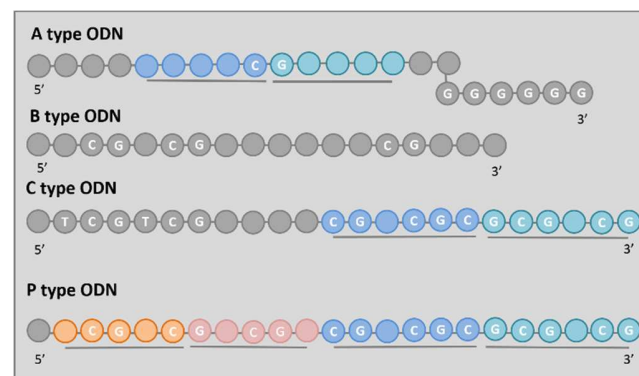

Figure 1 The sequences of different types of ODNs. Self-complementary palindromic sequences are depicted different shades of the same color and underlining, while the nucleotides of the CpG motifs are denoted in white letters.

Recent systematic analysis of B and C type CpG ODNs for human TLR9 have shown that the minimal sequence motif required for the activation of human [15] and mouse TLR9 are different [15]: one CpG motif localized close (4-6 nucleotide) to the 5' end is sufficient for the activation of mouse TLR9, while two, closely positioned (6-10 nucleotide distance) CpG motifs are necessary for human TLR9.

Furthermore, it has been determined that, in addition to the CpG motif, a TCG or TCC sequence at the 5' end of the molecule and length of 20-30 nucleotides more efficiently activate TLR9 [15, 30]. This indicated that nucleotides at the 5' end are also important and also suggested an additional binding site apart from the site for the CpG motif (see more details in 1.3.2). Specifically, mouse TLR9 activation [30] is augmented by a 5'TCC sequence one to three nucleotides from the CG present. Based on this result, they proposed a 23-mer CpG ODN, called minM80, with the sequence 5'-TCCT^{3}CGT^{15}. For efficient human TLR9 activation [15] the first CpG motif is preceded by the 5'-thymidine and the elongated poly-thymidine tail at the 3' end. Based on these findings they proposed a human specific agonist, minH75, a 24-mer with the sequence of sequence of 5'-TCGT^{7}CGT^{12}.

Furthermore, short ssDNA sequences with a 5' TCG motif (sODN), while unable of activate TLR9 themselves, were found to augment the TLR9 activation both in mouse and human cells lines [31-32]. Also, while single CpG motif containing ODNs are incapable of human TLR9 activation alone [15], together with sODNs they are able to bypass the requirement for two CpG motifs to potentiate human TLR9 [31-32].

1.2. Structural features of CpG ODNs in solution

While much of attention has been focused on the solution structure of dsDNA, ssDNA is less well characterized. The physical basis of structure formation of ssDNA is driven by base pairing between complementary nucleotides, while stacking interaction between adjacent bases and electrostatic interactions also play a crucial role [33] [34]. As a result, ssDNA structure is varied involving local secondary structure elements interspaced with random coil conformations, and it is also affected by solution state conditions (salt, pH, temperature, GC content, sequence). Self-complementary palindrome sequences enable intra- or intermolecular interactions via base pairing. The base pairing could involve canonical Watson-Crick or alternative Hoogsteen bonds. Oligomer formation could be a result of intermolecular base pairing interactions.
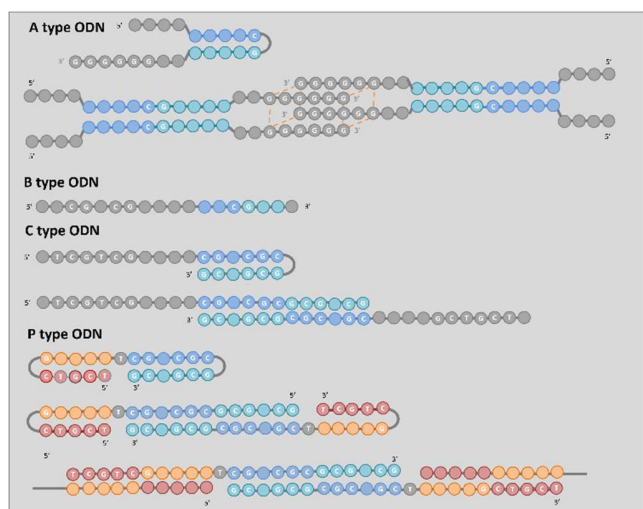

Figure 2 Schematic depiction of possible structural arrangements of different type of ODNs based on base pairing between complementary palindrome sequences.

Structural aspects, such as oligomerization of CpG ODNs has been shown to modulate their activities [29]. Structural propensities of CpG ODNs can be inferred from the presence of self-complementary palindrome sequences. Possible secondary structures could be predicted based on favorable canonical base pairing in the sequences as implemented in the *mfold* [35] program, Structures observed experimentally for different type of CpG ODNs are summarized in Figure 2.

A type CpG ODNs contain palindrome capable of forming a stem-loop structure (shown in Figure 2) capped at the 3' end by a poly G tail. The latter is suspected to form G tetrads/quadruplexes via Hoogsteen base pairs and form nanoparticles [36] in endolysosomes or to convert to short linear B type segments upon degradation by DNase II [9]. This latter hypothesis, however, makes it difficult to explain differing activities of A and B type CpG ODNs, unless uptake is enhanced by nanoparticle formation.

B type CpG ODNs were assumed to occur mostly in extended, random coil conformation [29], however, later a dimerization interface of ODN 1668 as depicted in Figure 2 was shown by NMR [37] [38] (see below). Formation of dimer structures were observed at low temperature (283K), in

solutions with 10 mM phosphate buffer at pH 7. It is, however, debated if they are also formed under physiological conditions (at 37ºC) and it is widely accepted that B type ODNs are extended short linear ssDNAs.

C type CpG ODN also contain a palindrome [26, 39-40], which may form either an intramolecular stem-loop structure in monomeric conformation or a serves as a dimerization interface as show in Figure 2.

P type ODNs contain a double palindrome, which can resolve into monomeric structures with a double hairpin-like conformation, a dimer containing a single hairpin or multimeric structures [29] as depicted in Figure 2.

Structural features have only been investigated in the case of B type CpG ODN 1668 and its fragments (shown in Figure 3) using NMR spectroscopy[37] [38]. The high resolution structure of the core hexamer, **1**, of B type CpG ODN 1668 (Figure 3) was determined by NMR [38] and found to form a dimer double helix. A decamer with 2-2 residues flanking the core hexamer, **2**, the tetradecamer, **3**, the eicosamer full ODN 1668, **4,** and the pentaeicosamer, **5** also showed signs of partial duplex formation in 10mM phosphate buffered, pH 7, 150 mM NaCl solution [38]. The concentrations of the samples were between 2-5 mM and spectra were measured 283K. These conditions do not represent physiological conditions, where ODN concentrations are lower and temperature is higher, however, these conditions could be relevant for *in vitro* immobilization of ODN immune modulators in drug delivery formations. The duplex formation was attributed to intermolecular base pairing via hydrogen bonds at the core hexamer, while the terminals were likely folded back onto the core region forming non-canonical structures.

1.   **GACGTT**
2.   AT**GACGTT**CC
3.   CCAT**GACGTT**CCTG
4. TCCAT**GACGTT**CCTGATGTC   ODN-1668
5. TCCAT**GACGTT**CCTGATGCTG**GGGG**

Figure 3 Fragments of CpG ODN 1668 studied by NMR.

1.3. Receptor bound CpG ODN conformations

The structural determinants of CpG ODN binding by TLR9 were revealed by crystal structures [41] that were obtained for several orthologues of TLR9 in a liganded and unliganded states together with a short B type CpG ODN 1668. The picture was further refined after a second binding site has been suggested for ssDNA in human TLR9 [15] [31], structures with occupying the second site of human TLR9 by sODN or by a longer immune stimulatory CpG ODN with two CpG motifs has been solved [32]. Later structural information on mouse TLR9 activation was also provided [42].

1.3.1. TLR9

Structurally TLR9, like all TLRs, is a single transmembrane protein, which consists of an N terminal ligand-binding comprised of Leucine Rich Repeats (LRR), a single membrane spanning domain and a C terminal signaling Toll/1L-1R Homology (TIR) domain. The LRR domain that recognizes PAMPs, which leads to oligomerization and activation of signaling mediated by the TIR domain.

1.3.2. Structures of TLR9 with agonistic CpG ODN and inhibitory DNA

The crystal structure of the unliganded C terminal domain of mouse TLR9 has been solved [43], however, revealed little regarding the molecular determinants of CpG motif containing ssDNA binding.

Figure 4 Shape and arrangement of LLR domain of TLR9 a)
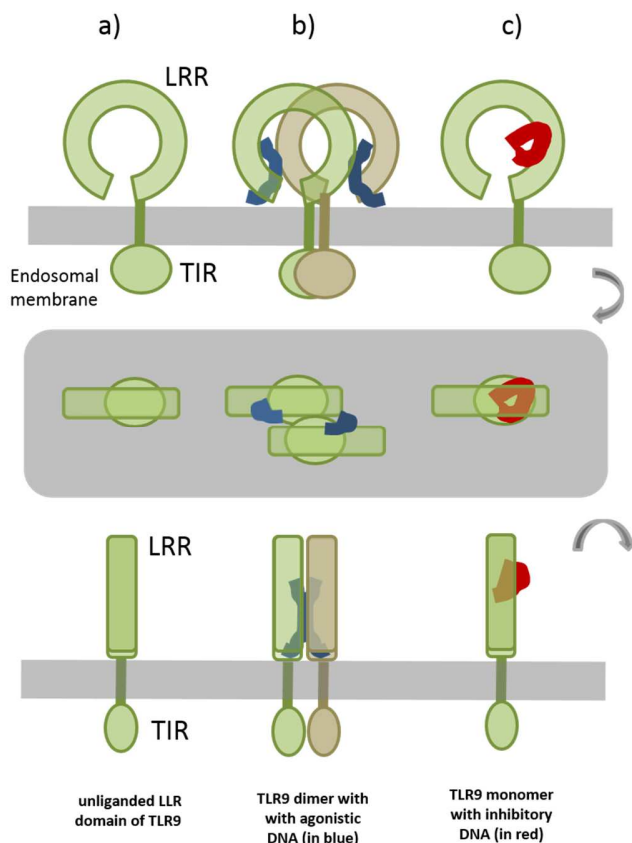
unliganded, b) with agonistic CpG ODN and c) inhibitory DNA.

The structural basis of CpG ODN recognition by TLR9 has first been revealed by crystallographic structures on complexes between TLR9 orthologues and a short, B type CpG ODN with a single CG motif [41]. The crystal structures solved also show structural determinants for molecular recognition of inhibitory ssDNA. Crystal structures of orthologues from equine (*Equus caballus*), from bovine (*Bos Taurus*), and murine (*Mus musculus*) were investigated. Unliganded equine (3WPB) and murine (3WPF) forms of

TLR9 were found to be monomeric as well as equine (3WPD) and murine (3WPG, 3WPH, 3WPI) TLR9 complexes with inhibitory ssDNA as opposed to agonistic ssDNA bound equine (3WPC) TLR9, which was shown to be dimeric. The arrangements are summarized in Figure 4.

Agonistic ssDNA, B-type ODN 1668 depicted in Figure 4b, acts as a molecular glue between the LLR domains of two TLR9 receptors leading to the formation of a TLR9 dimer that in turn initiates the signaling cascade leading to the activation of proinflammatory agents. The agonistic DNA bound TLR9 have an "m"-shaped 2:2 complex, whereby the C termini of the TLR9 protomers are positioned in the center. The agonistic ssDNA binds with its 3' end to the N terminal LLR0-11 of one TLR9 and the 5' part binds to the C terminal LLR20-22 of the other TLR9.

It has been shown that only disordered CpG containing ssDNA can act as an agonist, CpG motif containing ssDNA with secondary structure binds with much lower affinity to TLR9 LRRs. This highlights the importance of interactions formed by the bases of the ssDNA, which are buried in the interior of DNA forming duplexes. Indeed, in the crystal structure of agonistic ssDNA, interface-1 (Figure 5a) recognizes the bases of the CpG motifs, interactions at interface-2 primarily involves interactions with the backbone of the CpG ssDNA.

Figure 5 Schematic representations of crystallographic
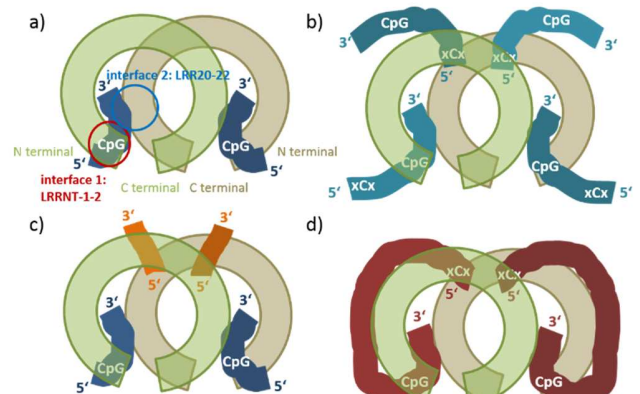
structures of TLR9 complexes with a) two agonistic CpG ODN (1668) with two TLR9 molecules (first TLR9 in green, second TLR9 in braun) at interfaces-1 and -2, b) short, 10-mer, CpG ODN and a short (6-mer) ssDNA with 5' xCx motif, c) short CpG ODN with a 5' xCx motif and c) longer CpG ODN with a 5' xCx motif.

The bases of the nucleotides are involved in direct and water mediated hydrogen bonds to amino acids of interface-1 consisting of LRRNT-1-2 of the N terminal part of the first TLR9 molecule. Stacking interactions were found between the nucleotide bases and aromatic amino acids such as Phe or Trp. Backbone phosphate groups of ssDNA interact with positively charged Lys, Arg and His sidechains. Interactions of the N2 atom in guanine in the CpG motif appears to be defining specificity of the CpG ODN towards the LRR domain of TLR9. At interface-2 containing LLR-20-22 from the C-terminal fragment of the second TLR9, there are

hydrogen bonds towards phosphate groups and van der Waals interactions to the sugar moieties of the ssDNA.

Complexes of TLR9 with inhibitory DNA were obtained and found to be monomeric as depicted in Figure 4c. The inhibitory ssDNA formed stem-loop structures by intramolecular base pairing that fit snugly into the interior ring of the LLR domain of TLR9. The stem-loop structure is formed by intramolecular base pairing between complementary 3-4 bases, which may also contain a mismatched pair and a short (3 bases) or a long (8 bases) loop. The recognition is driven by interactions involving the backbone of the inhibitory ssDNA, but also required contacts to the base of a non-base paired guanidine at the 3' end. Antagonistic effect of inhibitory ssDNA to disordered CpG ssDNA arises due to the fact that their binding sites overlap.

1.3.2. Structural determinants of the second ssDNA binding site of TLR9

It was shown that human TLR9 requires at least two CG motifs separated by 6-10 nucleotides for activation [15]. These double CpG ODNs are called immune stimulatory ODNs [31-32]. However, sODNs, which specifically augment TLR9 response with double CpG ODNs, can bypass the requirement of two CpG motifs in human TLR9 [31-32]. In same study it was shown that longer CpG ODN sequences with a TCG or a TCC motif at the 5' end were shown to activate TLR9 more efficiently than shorter CpG ODN without such motif at the 5' end [15]. All these findings suggested the existence of second ssDNA binding site on LLR domain of TLR9.

Model building of human TLR9 based on the crystal structures of orthologous TLR9 combined with a mutagenesis study [31] have been used to localize the residues important for the putative second binding site. Various biochemical biophysical and cell biological investigations [42] have shown that binding to the second site cooperatively promote receptor dimerization and activation. Investigations of the sequence requirements for the second binding site have shown that an ssDNA containing cytosine at the second position from the 5'-end is needed and thus was called a 5'-xCx motif. Finally, crystallographic structures of equine (*Equus caballus*) 5Y3J, 5Y3K, 5Y3L and bovine (*Bos Taurus*) 5Y3M TLR9 orthologues in complex with various CpG ssDNA pinpointed the exact place and binding mode of the second binding site as summarized in Figure 5b, c and d.

Using two ssDNA sequence, one with a short (10-mer) CpG ODN and an even shorter 6-mer with 5'-xCx motif, the former assumed the same binding mode as has been found previously [41], while the latter bound to the newly identified binding site in the middle of the LLR domain at LRR17-19 on one TLR9 molecule and LRR11 on the other TLR9. Indeed, only the last three 3' residues made contact with the ligand binding domains of the two TLR9, T1 and C2 connecting to both while G3 bound to only one TLR9 as shown in Figure 5b. These short 5'-xCx motif containing ssDNA sequences cannot activate TLR9 alone, and bind to the second site only when the first one is occupied.

CpG ODNs containing both CpG and 5'-xCx motifs are strong activators of TLR9. These so called immune stimulatory CpG ODNs [44] and showed two alternative types of binding modes. The short sequence (Figure 5c) was unable span both binding sites, thus bound to both individually, while the long sequence (Figure 5d) was able to stretch out to cover both binding sites.

Thus it has been shown that human TLR9 contains two binding site, a primary CpG motif site and a secondary site for the 5' xCx motif, which acts synergistically to induce TLR9 activity.

Finally, structural features of mouse TLR9 activation were clarified [42]. A single CpG motif containing ODN is sufficient for efficient activation, however, dimerization of mouse TLR9 can also be further enhanced by 5'-xCx ssDNA as with human TLR9. The latter effect was also observed for equine TLR9. No crystal structures could be obtained for mouse TLR9 and CpG ODN alone, only in addition of a sODN. Crystal structure of mouse TLR9 in complex with 10-mer CpG DNA (AGGCGTTTTT) and 6-mer 5'-xCx DNA (TCGCCA) was solved (5ZLN), in which both ligands bound to their corresponding binding sites with similar conformations as observed with equine complex (5J3J). Most residues contributing to CpG ODN binding were conserved in mouse and equine TLR9 complexes with the exception of 4 amino acids. Furthermore, some differences in the residues contributing to the protein-protein interactions at the dimerization interaction were found. It was found with Size Exclusion Chromatography analysis that the mouse TLR9 dimer complex with CpG ODN alone and with CpG ODN and sODN combination are more similar to each other than analogous equine structures. It was concluded that in spite of overall similarity, the murine TLR9 is more easily dimerizes than the equine orthologue due to species specific differences in their sequences.

CONCLUSION

Selectivity between pathogenic microbial and endogenous DNA sequences of the host is based on multiple factors. CpG motifs are more frequent in prokaryotic than in eukaryotic genome, furthermore, in the latter cysteine residues are often methylated [45]. Compartmentalization of TLR9 in the endosome and its acidic environment with nuclease activity also contributes to the selectivity between pathogenic and endogenous DNA fragments[46]. Endogenous DNA from apoptotic cells is also protected by histones from DNases. The requirement for binding to the second binding site in human TLR9 was proposed to contribute to further selectivity [44].

Crystal structures of TLR9 in complex with a short CpG ODN have shown that recognition of DNA by TLR9 receptors requires unstructured regions of ssDNA for binding, while inhibitory ssDNA with structured, duplex regions prevents oligomerization of TLR9 and thus inhibits dimerization and subsequent activation of immune stimulatory signaling.

The structural features of unbound CpG ODNs free in solution are less clear than that of the ODNs in the bound form. While an NMR study has found intermolecular dimer formation for B type CpG ODN 1668 *in vitro*, studies under physiological conditions indicate disordered monomers. Conformational studies of A and C type CpG ODNs, however, have shown formation of secondary structure elements either by intra- or intermolecular interactions.

If such structured states of CpG ODNs exist in solution, they have to unfold or shift equilibrium between folded and unfolded states towards disordered binding-competent conformation to bind the receptor as depicted in Figure 6. The unfolding event is associated with an entropic penalty, which needs to be compensated by enthalpic contributions to the binding coming from favorable interactions with the receptor. If the binding-competent disordered conformation is dominant in the conformational ensemble, the binding is entropically favored in accordance with the principle of conformational selection. Similar reasoning is valid for inhibitory ssDNA, which binds TLR9 in a duplex containing a stem-loop form.

Although, it is worth noting, that the equilibrium could also be shifted upon digestion of duplex containing states by DNases [9]. It has been shown that immune response to A type CpG ODNs is impaired without the presence of DNase II enzymes in endolysosomes, which was observed to digest 20-mer A type CpG ODNs molecules into 11-12-mers containing the 3' end including the CG motif. These fragments resemble linear fragments of B type CpG DNA, which do not show the same dependence on DNase digestion. As a possible explanation the authors reason that duplex regions of A type CpG ODNs prevent recognition of the CG motif.
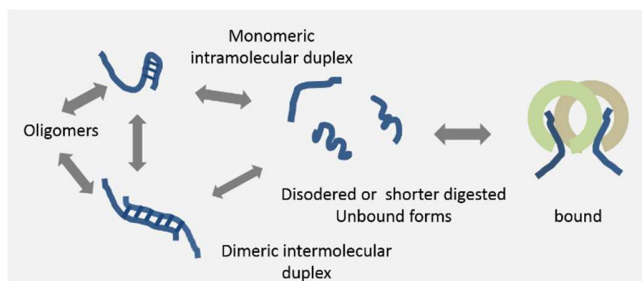

Figure 6 Conformational selection of CpG ODNs for binding TLR9

By conformational selection, the thermodynamic character influencing binding strength is dependent on the conformational ensemble of CpG ODNs. Further insights in the structural properties of CpG ODNs affecting the energetics of the binding could be exploited for the design of CpG ODN sequences that are better activators of TLR9, or more suited for surface adsorption on carriers of delivery systems or more efficiently taken up. Conformational studies of ssDNA are, however, highly challenging due their intrinsic flexibility. As described above, some attempts were made to investigate the structural propensities of CpG ODNs in the years following discovery of their immune stimulatory properties, but comprehensive understanding of their conformations is still missing.

Crystal structures provided crucial insight into the interactions between ssDNA and TLR9, however, they represent static snapshots without information on dynamics of the complex. Indeed, it has been proposed that the stability of equine and murine TLR9 complexes with the same ligands are distinctly different. Biomolecular simulations applied to the complex structure could provide detailed insights into the dynamic processes underlying these observations.

From a structural point of view, the case of unstructured ssDNA versus dsDNA is similar to the case of Intrinsically Disordered Proteins (IDPs) versus structured proteins: description of conformational properties is complex and few methods are suitable under physiologically relevant conditions. NMR spectroscopy would be a suitable methods to characterize conformational ensembles, however, NMR studies are hindered by the size of the ssDNA fragments in samples containing natural abundance NMR active isotopes and by the complexity of methods for stable isotope labelling. Again, biomolecular simulations may provide alternative means to gain insights into the conformational properties immune modulatory ssDNA [47]. With improvement of computational capacities and the development of enhanced sampling algorithms and multiscale models, biomolecular simulations may provide a more comprehensive description of structural and dynamic features of CpG ODNs.

CONFLICT OF INTEREST

No conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1] Bartneck, M. Immunomodulatory Nanomedicine. *Macromol. Biosci.* **2017,** *17* (10), 13.
[2] Medzhitov, R.; Janeway, C. A. An ancient system of host defense. *Current Opinion in Immunology* **1998,** *10* (1), 12-15.
[3] Krieg, A. M.; Yi, A. K.; Matson, S.; Waldschmidt, T. J.; Bishop, G. A.; Teasdale, R.; Koretzky, G. A.; Klinman, D. M. CPG MOTIFS IN BACTERIAL-DNA TRIGGER DIRECT B-CELL ACTIVATION. *Nature* **1995,** *374* (6522), 546-549.
[4] Medzhitov, R. Toll-like receptors and innate immunity. *Nature Reviews Immunology* **2001,** *1* (2), 135-145.

[5] Kumagai, Y.; Takeuchi, O.; Akira, S. TLR9 as a key receptor for the recognition of DNA. *Adv. Drug Deliv. Rev.* **2008,** *60* (7), 795-804.

[6] Hornung, V.; Rothenfusser, S.; Britsch, S.; Krug, A.; Jahrsdorfer, B.; Giese, T.; Endres, S.; Hartmann, G. Quantitative expression of Toll-like receptor 1-10 mRNA in cellular subsets of human peripheral blood mononuclear cells and sensitivity to CpG oligodeoxynucleotides. *J. Immunol.* **2002,** *168* (9), 4531-4537.

[7] Hemmi, H.; Takeuchi, O.; Kawai, T.; Kaisho, T.; Sato, S.; Sanjo, H.; Matsumoto, M.; Hoshino, K.; Wagner, H.; Takeda, K.; Akira, S. A Toll-like receptor recognizes bacterial DNA. *Nature* **2000,** *408* (6813), 740-745.

[8] Klinman, D. M. Immunotherapeutic uses of CpG oligodeoxynucleotides. *Nat. Rev. Immunol.* **2004,** *4* (4), 248-257.

[9] Chan, M. P.; Onji, M.; Fukui, R.; Kawane, K.; Shibata, T.; Saitoh, S.; Ohto, U.; Shimizu, T.; Barber, G. N.; Miyake, K. DNase II-dependent DNA digestion is required for DNA sensing by TLR9. *Nat. Commun.* **2015,** *6*, 10.

[10] Hara, T.; Tanegashima, K.; Takahashi, R.; Nuriya, H.; Naruse, N.; Tsuji, K.; Shigenaga, A.; Otaka, A. A Novel Function of a CXC-Type Chemokine CXCL14 As a Specific Carrier of CpG DNA into Dendritic Cells for Activating Toll-like Receptor 9-Mediated Adaptive Immunity. *Blood* **2016,** *128* (22), 5.

[11] Tanegashima, K.; Takahashi, R.; Nuriya, H.; Iwase, R.; Naruse, N.; Tsuji, K.; Shigenaga, A.; Otaka, A.; Hara, T. CXCL14 Acts as a Specific Carrier of CpG DNA into Dendritic Cells and Activates Toll-like Receptor 9-mediated Adaptive Immunity. *EBioMedicine* **2017,** *24*, 247-256.

[12] Montomoli, E.; Piccirella, S.; Khadang, B.; Mennitto, E.; Camerini, R.; De Rosa, A. Current adjuvants and new perspectives in vaccine formulation. *Expert Rev. Vaccines* **2011,** *10* (7), 1053-1061.

[13] Shirota, H.; Tross, D.; Klinman, D. M. CpG Oligonucleotides as Cancer Vaccine Adjuvants. *Vaccines* **2015,** *3* (2), 390-407.

[14] Scheiermann, J.; Klinman, D. M. Clinical evaluation of CpG oligonucleotides as adjuvants for vaccines targeting infectious diseases and cancer. *Vaccine* **2014,** *32* (48), 6377-6389.

[15] Pohar, J.; Krajnik, A. K.; Jerala, R.; Bencina, M. Minimal Sequence Requirements for Oligodeoxyribonucleotides Activating Human TLR9. *J. Immunol.* **2015,** *194* (8), 3901-3908.

[16] Heeg, K.; Dalpke, A.; Peter, M.; Zimmermann, S. Structural requirements for uptake and recognition of CpG oligonucleotides. *Int. J. Med. Microbiol.* **2008,** *298* (1-2), 33-38.

[17] Klinman, D. M. Use of CpG oligodeoxynucleotides as immunoprotective agents. *Expert Opin. Biol. Ther.* **2004,** *4* (6), 937-946.

[18] Krug, A.; Rothenfusser, S.; Hornung, V.; Jahrsdorfer, B.; Blackwell, S.; Ballas, Z. K.; Endres, S.; Krieg, A. M.; Hartmann, G. Identification of CpG oligonucleotide sequences with high induction of IFN-alpha/beta in plasmacytoid dendritic cells. *European Journal of Immunology* **2001,** *31* (7), 2154-2163.

[19] Hartmann, G.; Krieg, A. M. Mechanism and function of a newly identified CpG DNA moth in human primary B cells. *J. Immunol.* **2000,** *164* (2), 944-952.

[20] Vollmer, J.; Jurk, M.; Samulowitz, U.; Lipford, G.; Forsbach, A.; Wullner, M.; Tluk, S.; Hartmann, H.; Kritzler, A.; Muller, C.; Schetter, C.; Krieg, A. M. CpG oligodeoxynucleotides stimulate IFN-gamma-inducible protein-10 production in human B cells. *Journal of Endotoxin Research* **2004,** *10* (6), 431-438.

[21] Guiducci, C.; Ott, G.; Chan, J. H.; Damon, E.; Calacsan, C.; Matray, T.; Lee, K. D.; Man, R. L. C.; Barrat, F. J. Properties regulating the nature of the plasmacytoid dendritic cell response to Toll-like receptor 9 activation. *Journal of Experimental Medicine* **2006,** *203* (8), 1999-2008.

[22] Verthelyi, D.; Ishii, K. J.; Gursel, M.; Takeshita, F.; Klinman, D. M. Human peripheral blood cells differentially recognize and respond to two distinct CpG motifs. *J. Immunol.* **2001,** *166* (4), 2372-2377.

[23] Gursel, M.; Verthelyi, D.; Gursel, I.; Ishii, K. J.; Klinman, D. M. Differential and competitive activation of human immune cells by distinct classes of CpG oligodeoxynucleotide. *Journal of Leukocyte Biology* **2002,** *71* (5), 813-820.

[24] Pohar, J.; Lainscek, D.; Kunsek, A.; Cajnko, M. M.; Jerala, R.; Bencina, M. Phosphodiester backbone of the CpG motif within immunostimulatory oligodeoxynucleotides augments activation of Toll-like receptor 9. *Sci Rep* **2017,** *7*, 11.

[25] Klinman, D. M.; Yi, A. K.; Beaucage, S. L.; Conover, J.; Krieg, A. M. CpG motifs present in bacterial DNA rapidly induce lymphocytes to secrete interleukin 6, interleukin 12, and interferon gamma. *Proceedings of the National Academy of Sciences of the United States of America* **1996,** *93* (7), 2879-2883.

[26] Vollmer, J.; Weeratna, R.; Payette, P.; Jurk, M.; Schetter, C.; Laucht, M.; Wader, T.; Tluk, S.; Liu, M.; Davis, H. L.; Krieg, A. M. Characterization of three CpG oligodeoxynucleotide classes with distinct immunostimulatory activities. *European Journal of Immunology* **2004,** *34* (1), 251-262.

[27] Abel, K.; Wang, Y. C.; Fritts, L.; Sanchez, E.; Chung, E.; Fitzgerald-Bocarsly, P.; Krieg, A. M.; Miller, C. J. Deoxycytidyl-deoxyguanosine oligonucleotide classes A, B, and C induce distinct ctokine gene expression patterns in rhesus monkey peripheral blood mononuclear cells and distinct alpha interferon responses in TLR9-expressing rhesus monkey plasmacytoid dendritic cells. *Clinical and Diagnostic Laboratory Immunology* **2005,** *12* (5), 606-621.

[28] Marshall, J. D.; Fearon, K. L.; Higgins, D.; Hessel, E. M.; Kanzler, H.; Abbate, C.; Yee, P.; Gregorio, J.; Dela Cruz, T.; Lizcano, J. O.; Zolotorev, A.; McClure, H. M.; Brasky, K. M.; Murthy, K. K.; Coffman, R. L.; Van Nest, G. Superior activity of the type C class of ISS in vitro and in vivo across multiple species. *DNA and Cell Biology* **2005,** *24* (2), 63-72.

[29] Samulowitz, U.; Weber, M.; Weeratna, R.; Uhlmann, E.; Noll, B.; Krieg, A. M.; Vollmer, J. A Novel Class of Immune-

Stimulatory CpG Oligodeoxynucleotides Unifies High Potency in Type I Interferon Induction with Preferred Structural Properties. *Oligonucleotides* **2010,** *20* (2), 93-101.

[30] Pohar, J.; Lainscek, D.; Fukui, R.; Yamamoto, C.; Miyake, K.; Jerala, R.; Bencina, M. Species-Specific Minimal Sequence Motif for Oligodeoxyribonucleotides Activating Mouse TLR9. *J. Immunol.* **2015,** *195* (9), 4396-4405.

[31] Pohar, J.; Lainscek, D.; Ivicak-Kocjan, K.; Cajnko, M. M.; Jerala, R.; Bencina, M. Short single-stranded DNA degradation products augment the activation of Toll-like receptor 9. *Nat. Commun.* **2017,** *8*, 13.

[32] Ohto, U.; Ishida, H.; Shibata, T.; Sato, R.; Miyake, K.; Shimizu, T. Toll-like Receptor 9 Contains Two DNA Binding Sites that Function Cooperatively to Promote Receptor Dimerization and Activation. *Immunity* **2018**.

[33] Martinez, J. M.; Elmroth, S. K. C.; Kloo, L. Influence of sodium ions on the dynamics and structure of single-stranded DNA oligomers: A molecular dynamics study. *J. Am. Chem. Soc.* **2001,** *123* (49), 12279-12289.

[34] Zhang, Y.; Zhou, H. J.; Ou-Yang, Z. C. Stretching single-stranded DNA: Interplay of electrostatic, base-pairing, and base-pair stacking interactions. *Biophys. J.* **2001,** *81* (2), 1133-1143.

[35] Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003,** *31* (13), 3406-3415.

[36] Kerkmann, M.; Costa, L. T.; Richter, C.; Rothenfusser, S.; Battiany, J.; Hornung, V.; Johnson, J.; Englert, S.; Ketterer, T.; Heckl, W.; Thalhammer, S.; Endres, S.; Hartmann, G. Spontaneous formation of nucleic acid-based nanoparticles is responsible for high interferon-alpha induction by CpG-A in plasmacytoid dendritic cells. *Journal of Biological Chemistry* **2005,** *280* (9), 8086-8093.

[37] Narayanan, S.; Dalpke, A. H.; Siegmund, K.; Heeg, K.; Richert, C. CpG Oligonucleotides with modified termini and nicked dumbbell structure show enhanced immunostimulatory activity. *J. Med. Chem.* **2003,** *46* (23), 5031-5044.

[38] He, G. Y.; Patra, A.; Siegmund, K.; Peter, M.; Heeg, K.; Dalpke, A.; Richert, C. Immunostimulatory CpG oligonucleotides form defined three-dimensional structures: Results from an NMR study. *ChemMedChem* **2007,** *2* (4), 549-560.

[39] Hartmann, G.; Battiany, J.; Poeck, H.; Wagner, M.; Kerkmann, M.; Lubenow, N.; Rothenfusser, S.; Endres, S. Rational design of new CpG oligonucleotides that combine B cell activation with high IFN-alpha induction in plasmacytoid dendritic cells. *European Journal of Immunology* **2003,** *33* (6), 1633-1641.

[40] Marshall, J. D.; Fearon, K.; Abbate, C.; Subramanian, S.; Yee, P.; Gregorio, J.; Coffman, R. L.; Van Nest, G. Identification of a novel CpG DNA class and motif that optimally stimulate B cell and plasmacytoid dendritic cell functions. *Journal of Leukocyte Biology* **2003,** *73* (6), 781-792.

[41] Ohto, U.; Shibata, T.; Tanji, H.; Ishida, H.; Krayukhina, E.; Uchiyama, S.; Miyake, K.; Shimizu, T. Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9. *Nature* **2015,** *520* (7549), 702-U303.

[42] Ishida, H.; Ohto, U.; Shibata, T.; Miyake, K.; Shimizu, T. Structural basis for species-specific activation of mouse Toll-like receptor 9. *FEBS letters* **2018**.

[43] Collins, B.; Wilson, I. A. Crystal structure of the C-terminal domain of mouse TLR9. *Proteins* **82** (10), 2874-2878.

[44] Pohar, J.; Yamamoto, C.; Fukui, R.; Cajnko, M. M.; Miyake, K.; Jerala, R.; Bencina, M. Selectivity of Human TLR9 for Double CpG Motifs and Implications for the Recognition of Genomic DNA. *J. Immunol.* **2017,** *198* (5), 2093-2104.

[45] Stacey, K. J.; Young, G. R.; Clark, F.; Sester, D. P.; Roberts, T. L.; Naik, S.; Sweet, M. J.; Hume, D. A. The molecular basis for the lack of immunostimulatory activity of vertebrate DNA. *J. Immunol.* **2003,** *170* (7), 3614-3620.

[46] Barton, G. M.; Kagan, J. C. A cell biological view of Toll-like receptor function: regulation through compartmentalization. *Nature Reviews Immunology* **2009,** *9* (8), 535-542.

[47] Feher, K. Atomistic simulations of immune stimulatory single stranded bacterial DNA. In *ISQBP Presindet's Meeting 2018*, Barcelona, Spain, **2018**.