Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Cytosolic DNA sensing via cGAS: long, U-shaped and other DNA ligands and cellular co-factors

von

Liudmila Andreeva

aus

Moskau, Russland

2018

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordung vom 28. November 2011 von Herrn Prof. Dr. Karl-Peter Hopfner betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 06.03.2018

………………………………………….

Liudmila Andreeva

Dissertation eingereicht am	06.03.2018
1. Gutachter	Prof. Dr. Karl-Peter Hopfner
2. Gutachter	Prof. Dr. Veit Hornung
Mündliche Prüfung am	16.04.2018

This thesis has been prepared from August 2013 to October 2017 in the laboratory of Professor Dr. Karl-Peter Hopfner at the Gene Center of the Ludwig-Maximilians-University Munich.

This is a cumulative thesis based on following publications:

Andreeva, L., Hiller, B., Kostrewa, D., Lässig, C., de Oliveira Mann, C. C., Jan Drexler, D., Maiser, A., Gaidt, M., Leonhardt, H., Hornung, V., & Hopfner, K.-P. (2017) cGAS senses long and HMGB/TFAM-bound U-turn DNA by forming protein–DNA ladders. *Nature*, 549(7672): 394-398.

Herzner, A.-M., Hagmann, C. A., Goldeck, M., Wolter, S., Kubler, K., Wittmann, S., Gramberg, T., Andreeva, L., Hopfner, K.-P., Mertens, C., Zillinger, T., Jin, T., Xiao, T. S., Bartok, E., Coch, C., Ackermann, D., Hornung, V., Ludwig, J., Barchet, W., Hartmann, G., & Schlee, M. (2015) Sequence-specific activation of the DNA sensor cGAS by Y-form DNA structures as found in primary HIV-1 cDNA. *Nat Immunol*, 16(10): 1025-1033.

Mankan, A. K., Schmidt, T., Chauhan, D., Goldeck, M., Höning, K., Gaidt, M., Kubarenko, A. V., Andreeva, L., Hopfner, K. P., & Hornung, V. 2014. Cytosolic RNA:DNA hybrids activate the cGAS–STING axis. *The EMBO Journal*, 33(24): 2937-2946.

Hopfner K.-P., Andreeva L., Drexler D.J. (Filed 24 July 2017) "A fluorescent cyclic dinucleotide and its use in methods of identifying substances having an ability to modulate the cGAS/STING pathway." European Patent Application EP 17182689.4.

Table of contents

Summary

Innate immune sensing of cytosolic nuclear acids is executed by pattern recognition receptors and is a powerful tool to counteract viral and bacterial infection. Nucleic acids with their essential function as genetic information carriers serve as a very general pathogen-derived pattern and therefore trigger a powerful immune response. The innate immune system evolved to distinguish between pathogen-specific patterns such as secondary structures or modifications of RNA and host nucleic acids. In case of DNA, however, such strategies as DNA compartmentalization are used to prevent self-DNA recognition. Indeed, cellular DNA normally present in nuclei and mitochondria is immunosilent, whereas accumulation of cytosolic DNA triggers an inflammatory response. Such DNA can emerge due to viral infection, mitochondrial and nuclear stress or due to dysfunction of key proteins responsible for elimination of excessive DNA amounts in the cytosol and lysosomes.

cGAS is a central sensor of cytosolic DNA expressed in almost all cell types. cGAS recognizes cytosolic dsDNA in a broad sequence-indiscriminatory manner and synthesizes the second messenger cyclic GMP-AMP (pG(2'-5')pA(3'-5'), 2'3'-cGAMP) from ATP and GTP. Unlike bacterial cyclic dinucleotides (CDNs) with canonical 3'-5' linkages, 2'3'-cGAMP is a unique metazoan CDN that comprises both 3'-5' and a non-canonical 2'-5' phosphodiester linkages connecting adenosine with guanosine and guanosine with adenosine, respectively. 2'3'-cGAMP produced upon infection binds and activates the downstream adaptor stimulator of interferon genes (STING). Residing on the endoplasmic reticulum, STING binds cGAS-generated cGAMP or bacterial CDNs and undergoes trafficking to the Golgi complex in perinuclear space where it recruits TANK-binding kinase 1 (TBK1) and transcription factor interferon regulatory factor 3 (IRF3). As a result of the pathway activation IRF3 gets phosphorylated, dimerizes and translocates into the nucleus resulting in type I interferons (IFNs) production.

A range of cGAS structures and biochemical studies revealed activation and catalytic mechanisms of cGAS. cGAS was discovered to dimerize upon DNA binding in a way that two DNA molecules are sandwiched between two cGAS protomers and such dimerization was shown to be necessary for cGAS activation. The nature and physiological function of cGAS dimerization, however, remains elusive, since such conformation was not found for functionally and structurally similar 2'-5'-oligoadenylate synthetases (OAS) or other proteins suggesting it to be a unique feature of cGAS. Moreover, the proposed dimerization does not explain why short DNA constructs of 14-20 base pairs (bp) used for crystallization fail to fully activate cGAS *in vivo*, though they are capable of inducing all conformational changes known for cGAS activation in the crystal. Furthermore, the composition of such cGAS dimers would lead to steric clashes between two bound DNA molecules, if the length of DNA strands exceeds 18 bp present in the structure. The instability of such dimeric cGAS conformation is another enigma to be clarified, since $cGAS_2:DNA_2$ species could only be observed in non-physiological high concentrations of cGAS and DNA.

In this work a mechanism of cGAS activation by biologically relevant ligands was studied. Furthermore, a model for cooperative sensing of long DNA by cGAS was established. Mab-21 domain of cGAS was found to have an intrinsic capability to measure DNA length, since its activity dramatically increased with DNA length by the same number of cGAS binding sites *in vitro* and in cell-based experiments. In order to investigate cGAS activity, a novel high-throughput fluorescence-based assay was developed. The first

crystal structure of cGAS in complex with stimulatory DNA of 39 bp presented in this thesis provided an insight into cGAS activation by fibril formation. Such cGAS oligomers were found to make protein-DNA ladders with two nearly parallel DNA strands as "ladder sides" and cGAS dimers as "rungs" holding them together. cGAS-DNA oligomers, first observed in crystal packing, were confirmed by isothermal titration calorimetry (ITC) and by size exclusion chromatography coupled to right-angle light scattering (SEC-RALS). These methods revealed the stoichiometry and molecular weights of cGAS complexes with different DNA species that together determined an exact complex composition and confirmed the presence of $(cGAS_2)_n$:DNA_2 complexes in solution. According to our model, the formation of the first dimer is highly unfavorable resulting in an unstable complex, however, it parallelizes two DNA strands and enables an effective binding of the subsequent cGAS dimer. This results in cooperative binding of cGAS on long stimulatory DNA with cooperativity emerging already from cGAS-DNA interactions. cGAS dimers are mutually stabilized within a fibril resulting in higher cGAS activity. Furthermore, potential cGAS co-factors were proposed. Based on similarities in DNA conformations and a recent research revealing mitochondrial DNA to be a cGAS activator, mitochondrial transcription factor A (TFAM), as well as high-mobility group box 1 (HMGB1) and bacterial nucleoid HU proteins were discovered to enhance cGAS activity. According to our hypothesis, TFAM and other DNA-bending proteins introduce a U-turn in long DNA, prearrange DNA in a manner favorable for cGAS dimerization and thus serve as nucleation points for cGAS-DNA ladder formation.

Another part of this thesis was dedicated to discover other physiological nucleic acid ligands of cGAS. Cell line experiments with different synthetic constructs revealed RNA:DNA hybrids to be potent cGAS activators. Furthermore, ssDNA hairpins formed by HIV-1 reverse-transcribed ssDNA were found to stimulate cGAS during HIV-1 infection. Substitution of guanines flanking the double-stranded hairpin regions abolished cGAS activity and specific short Y-shaped DNA with G-overhangs (YSD) were found to be potent cGAS ligands. Intriguingly, such YSDs were capable of cGAS activation despite their short length (< 20 bp) leaving the question open, whether the cGAS-DNA ladder model is applicable in this case.

Taken together, this work presents a range of cGAS activators that include RNA:DNA hybrids, HIV-1 ssDNA hairpins and partly unfolded mitochondrial and bacterial nucleoids rather than naked dsDNA as specific cGAS ligands. The oligomerization mechanism of cGAS described in this thesis provides a link between a peculiar cGAS dimerization, on the one hand, and high stimulatory activity of long dsDNA *in vivo*, on the other hand. Proposed cGAS co-factors further clarify the complexity of DNA recognition by cGAS in the living cell suggesting that analogically to RNA sensing by RIG-I, cGAS recognizes specific structures of dsDNA rather than just mislocalized DNA fragments. However, further studies are needed to evaluate the relevance of the oligomerization mechanism for cGAS sensing of other ligands and to discover other cellular co-factors facilitating cGAS activity *in vivo*.

Introduction

1. Immunity and innate immune system

Recognizing and fighting invading pathogens and foreign molecules are the key mechanisms of organism survival. Through the whole tree of life organisms developed mechanisms aimed to response to invading pathogens that are called immune system [2]. For example, bacteria possess a set of inducible nucleases to eliminate phage DNA [3]. The innate immune system is an evolutionary conserved part of the immune system, developed early in multicellular organisms. It recognizes the invading pathogens through pathogen-associated molecular patterns (PAMPs), the presence of altered self through damage-associated molecular patterns (DAMPs) or the absence of self, like in case of MHC class I recognition by natural killer cells [4, 5]. The innate immune system is represented by a limited number of genes and can be characterized by its precision in pattern recognition, immediate activation and recognition of evolutionary conserved molecular patterns [6]. Vertebrates are the only phylum that together with innate immunity developed an adaptive immune system that allows a much higher variability of receptors and sensed patterns and immunological memory of invading pathogens. Though the adaptive immune system is more flexible and pathogen-specific in comparison to the innate immune system, it is also slower and more prone to allergic and autoimmune reactions[6]. Both the innate and adaptive immune systems create a strong defense against pathogen invasion: innate immunity is the first line of defense, initiating inflammatory response upon binding to pathogen-specific proteins, nucleic acids and lipids, whereas adaptive immune response fine-tunes the pathogen recognition by adding adjustable antigen-specific receptors [7]. However, the presence of innate immune responses are crucial for adaptive immunity: activation of the innate immune system, represented by macrophages, dendritic cells and NK-cells, results in production of proinflammatory cytokines, chemokines and type I and II interferons that recruit and activate T and B lymphocytes of the adaptive immune system [8-10]. Moreover, activation of innate immune system is a crucial step for antigen-presentation to T-lymphocytes by dendritic cells [11].

Dysregulation of the innate immune system is associated with various autoimmune diseases, such as rheumatoid arthritis, Crohn's disease (CD), systemic lupus erythematosus (SLE), Acardi Goutières syndrome (AGS), and Sjogren's syndrome [12]. Dysfunction of the innate immune system, on the other hand, can result in higher susceptibility to viral and bacterial infections. Thus, deep understanding of precise mechanisms of innate immune responses, their regulation and inhibition by invading pathogens is absolutely necessary for developing of new vaccines and therapies against the huge range of diseases of bacterial, viral or autoimmune origin.

2. Pattern recognition receptors

The main targets of innate immune recognition are pathogen- and damage-associated molecular patterns (PAMPs and DAMPs), such as lypopolysaccharides, lipopeptides, flagellin, cyclic dinucleotides (CDNs) or various nucleic acids emerging from bacterial and viral replication or from the damaged self. Such molecular patterns have several features in common: they are produced only by pathogen or in damaged host and not by healthy host organism; they are essential for pathogen viability, so that pathogens can not lose them through evolution; the same type of PAMPs or DAMPs are usually introduced by several groups of pathogens and stress signals, so that the innate immune system can use the same tools for

multiple kinds of invasions; and PAMPs are chosen to carry information about the type of infection in order to trigger the best suitable immune response [4]. These PAMPs and DAMPs are sensed by the innate immune system using a range of evolutionarily conserved pattern recognition receptors (PRRs).

Major PRR families include Toll-like receptors, recognizing a variety of bacterial and fungous structural elements, as well as nucleic acids; C-type lectin receptors (CLRs) sensing mannose, fucose and glucan carbohydrate structures present in bacterial cell walls; absent in melanoma 2 (AIM2)-like and nucleotide-binding and oligomerization domain (NOD)-like receptors involved in inflammasome formation and caspase-1 activation; retinoic acid inducible gene I (RIG-I)-like receptors (RLRs) sensing various cytosolic RNA species; and cGAS-STING pathway responsible for cytosolic DNA sensing. Activation of these receptors leads to signal transduction through different adaptor proteins resulting in activation of various transcription factors, such as NF-κB, interferon regulatory factor 3 (IRF3) and 7 (IRF7) [13-15]. These cascades of interactions result in transcription activation and production of IFN-β, cytokines and chemokines that in turn activate the neighboring cells. IFN-β is produced in early stages of infection, binds to IFN-α/β receptor (IFNAR) and induces transcription of IFN-stimulated genes (ISGs) through the JAK-STAT pathway, thus increasing the anti-viral defense of the neighboring cells, activating natural killer (NK) cells and macrophages and increasing antigen presentation [16-19].

Though the main function of the majority of PRRs is inducing inflammation, their activation can be associated with various types of cell death in order to eliminate the pathogens and to trigger a potent immune response. Activation of ALRs or NLRs resulting in inflammasome formation and activation of caspase-1 leads to IL-1β and IL-18 maturation on the one side, but also to highly inflammatory pyroptotic cell death [20-22]. TLR2, 3 and 4 were shown to be involved in apoptosis by association of their adaptor Toll/IL-1R domain-containing adapter inducing IFN-β (TRIF) with Fas-associated death domain (FADD) and caspase-8-dependent apoptotic pathway [23-27]. The cytosolic DNA sensor DNA-dependent activator of interferon regulatory factors (DAI) is able to directly interact with receptor interacting protein kinase 3 (RIP3) and thus to mediate necroptosis in response to viral infection [28]. Additionally, activation of stimulator of interferon genes (STING) was recently shown to induce lysosomal cell death and to additionally activate inflammasome pathway resulting in high inflammatory response [29].

PRRs are present in different cell compartments. Toll-like and C-type lectin receptors are membrane-associated PRRs and are localized in the cell or endosome membranes. RLRs, ALRs, NLRs and cGAS are cytosolic PRRs. Some PRRs like mannan-binding protein or C-reactive protein are secreted into the blood stream where they participate in microorganism opsonisation and activation of the complement system, inflammation and phagocytosis [30].

The cooperative action and regulation of PRRs on the cell an organism levels provide a diverse and specific platform for pathogen and damage recognition serving as the first line of defense against the changes in organism homeostasis.

3. Nucleic acids sensors

Though recognition of bacterial structural elements by surface PRRs is extremely important for the rapid response to pathogens, sensing of intracellular pathogens is far more challenging for the innate immune system, since the exact discrimination between self and non-self is required. The most common group of

DAMPs occurring in all intracellular pathogens is represented by nucleic acids arising from bacterial and viral genomes and their replication products. Invertebrates recognize such nucleic acids mostly using RNA interference (RNAi). For example, in *Drosophila* the mutations in Dicer-2 or Argonaute-2 proteins result in higher susceptibility to viral infections [31, 32]. In contrast, vertebrates mostly use sequence-unspecific type I IFNs-dependent immune response. Surprisingly, the RNAi is even down-regulated in mammals upon viral infection, while it inhibits the excessive expression of interferon-stimulated genes (ISGs) in the absence of immunogenic stimuli [33]. The innate immune system of vertebrates recognizes nucleic acids as foreign according to their local concentration, specific localization, such as in cytosol or lysosomes and specific structural elements and chemical groups usually absent in the host, such as cytosolic double-stranded RNA or RNA with 5'-triphosphates at the RNA end [34].

Nucleic acid receptors can be divided into two groups: receptors with a direct antiviral activity and PRRs that activate the downstream signaling cascade and transcription factors resulting in the synthesis of effector proteins and cytokines. The expression level of the nucleic acid sensors depends of the cell type. For example, Toll-like receptors are largely expressed in specialized immune cells like dendritic cells, monocytes, macrophages and T- and B-lymphocytes. Epithelial cells of skin and lungs are equipped with high levels of cytosolic DNA and RNA sensors. Some PRRs, like cGAS, are expressed in almost all cell types. Recent studies showed that the same PRRs can be activated in presence of self-DNA, if it appears in unusual compartment. For example, mitochondrial and nuclear stresses can result in a leakage of cellular DNA into the cytosol where it is recognized by the same PRRs that usually detect pathogenic genomes. Specific groups of PRRs involved in nucleic acid sensing will be discussed in detail below.

3.1. Toll-like receptors

Toll-like receptors are an evolutionary conserved and one of the most studied groups of PRRs. The first Toll-like receptor was discovered in *Drosophila* and was found to be crucial for dorsal-ventral polarity establishment in embryogenesis [35]. Except for its function during development, this receptor was shown to be involved in *Drosophila*'s defense against fungal infection [36]. Homologous proteins were soon found in vertebrata that now correspond to one of the largest families of PRRs – toll-like and IL-1 receptor (TLR/IL-1R) family.

TLRs are expressed in epithelial cells and fibroblasts, as well as in immune cells like macrophages, dendritic cells and B- and T- cells [37]. In humans the TLR family comprises 10 members. Extracellular TLRs 1, 2, 4, 5 and 6 recognize structural components of bacteria, such as lipopeptides, peptidoglicans, lipopolysaccharides and flagellin. Other TLRs monitor the lumen of lysosomes for the presence of foreign nucleic acids. TLR3 is activated upon binding of synthetic and viral dsRNA, TLR 7 and 8 are responsible for GU-rich ssRNA emerging form viruses and bacteria [38-40]. TLR9 recognizes unmethylated CpG motifs of bacterial and viral DNA [41]. IL-1R family receptors are direct binders of proinflammatory cytokines like IL-1β, IL-18, IL-33 that regulate innate and adaptive immune responses [42, 43]. A couple of TLRs absent in humans were recently discovered in mice. TLR11 and 12 were shown to recognize profilin and flagellin emerging from apicomplexan parasites and bacteria [44-46]. Mouse TLR13, also absent in humans, specifically senses bacterial 23S rRNA of Gram-positive and Gram-negative bacteria [47].

TLRs are membrane-bound receptors that consist of a common cytoplasmic signaling domain – Toll-IL-1R (TIR) homology domain, transmembrane part and different ectodomains containing leucine-rich repeats that are responsible for PAMP sensing [48]. Upon ligand binding the TLRs form homo- or heterodimers that lead to TIR conformational changes and adaptor recruitment. For example, TLR2 forms a heterodimer with TLR6 or TLR1 for effective sensing of Gram-positive bacteria by recognizing peptidoglycan and di- or triacylopeptides, respectively [49]. TLR3, 4, 5 and 9 are known to form homodimers [50-53]. Such cooperative interactions between TLRs not only enhances the recognition of pathogens by combining several PAMPs recognition, but also plays important role in regulation of TLRs. TLR8 dimerization with TLR7 or TLR9 was shown to antagonize activation of these receptors, whereas TLR9 inhibits TLR7, indicating the complexity of TLRs activation and cross-talk [54].

Following the ligand binding, TLRs recruit adaptor proteins myeloid differentiation primary-response protein 88 (MYD88) or TIR domain-containing adaptor protein inducing IFNβ (TRIF) that lead to nuclear factor kappa light-chain enhancer of activated B cells (NF-κB) activation and proinflammatory cytokine production or type I IFN response, respectively (Figure 1). Most of TLRs use MYD88 which forms a complex with IL-1R-associated kinases (IRAKs) via its death domain forming a complex called Myddosome [55]. Myddosome initiates a kinase cascade leading to phosphorylation of NF-κB inhibitory protein IκBa and NF-κB translocation into the nucleus [56]. In addition to NF-κB pathway, TLR7 and TLR9 induce type I IFN production in MYD88-dependent manner by phosphorylating transcription factors IRF1 or IRF7 that lead to IFN-β activation [57]. In contrast to other TLRs, TLR3 (and internalized TLR4), a sensor of endosomal dsRNA, binds adaptor TRIF that associates with kinases TANK-binding kinase 1 (TBK1) and IKK-i which in turn phosphorylate

Figure 1 Toll-like receptor signaling pathways.
Upon ligand binding TLRs form homo- or heterodimers and recruit adaptors MyD88 (myeloid differentiation primary-response protein 88) or TRIF (TIR domain-containing adaptor protein inducing IFNβ) for downstream signaling resulting in production of proimflammatory cytokines and type I IFNs. Adopted from *Du B., et al., 2016* [1].

and activate transcription factor IRF3 and IRF7 resulting in IFN-β induction and expression of IFN-inducible genes [58-60]. TLR signaling pathways are summarized in Figure 1.
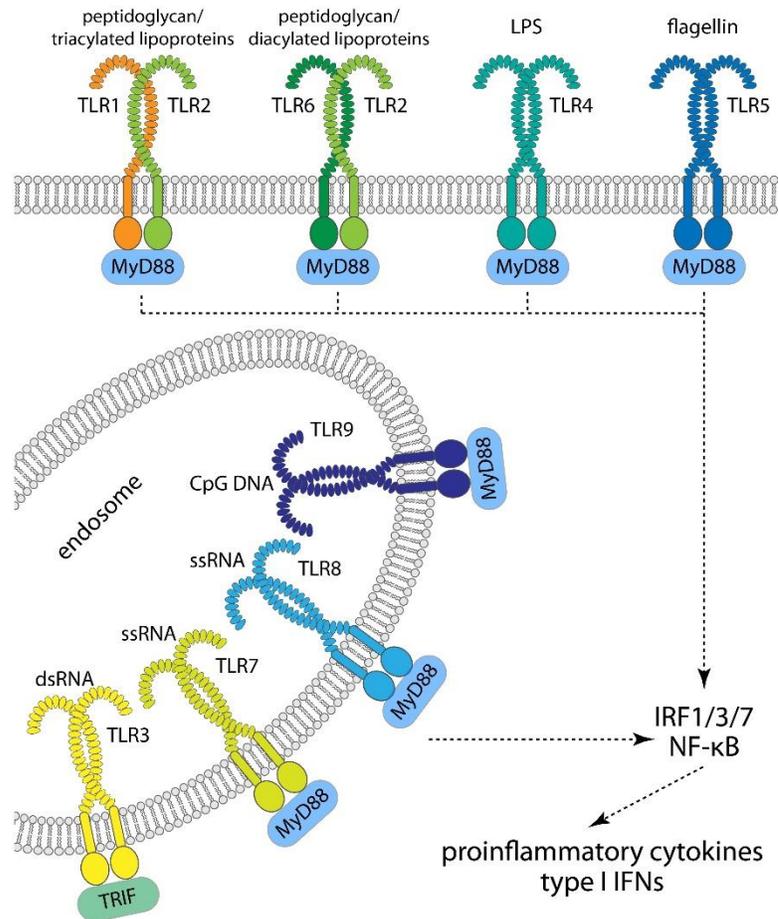
3.2. AIM2- and NOD-like receptors

Another group of PRRs sensing a broad range of pathogen associated molecules consists of absence in melanoma 2 (AIM2)-like receptors (ALRs) and nucleotide oligomerization domain (NOD)-like receptors (NLRs) [61]. ALRs and NLRs are known to form multi-protein signaling complexes called inflammasomes and to activate inflammatory caspases, leading to IL-1β and IL-18 secretion and cell death. Unlike TLRs, ALRs and NLRs are cytosolic receptors and are ubiquitously expressed in all cell types, however, different NLRs and ALRs are not equally distributed and can be preferentially expressed in particular tissue [62]. Invertebrates and plants have proteins similar to NLRs in domain architecture, like *Arabadopsis* resistance (R) proteins or NLR-proteins in corals, however the NLR homologues are missing in nematodes and *Drosophila* [61, 63, 64].

In humans ALRs and NLRs family counts 24 members that all possess a ligand binding domain and at least one signaling domain of death domain superfamily (Figure 2). AIM-2 family members consist of N-terminal effector pyrin domain (PYD) and one or two C-terminal hemopoietic expression, interferon-inducibility, nuclear localization (HIN) domains responsible for dsDNA binding [65-68]. NLRs comprise N-terminal effector domain followed by central nucleotide oligomerization domain (NOD) and C-terminal leucine-rich repeats (LRR) responsible for PAMP recognition. N-terminal effector domain is represented by PYD domain in NLRP subfamily, CARD in NLRCs or baculoviral inhibitory repeat (BIR)-like domains in NLR family-apoptosis inhibitory proteins (NAIPs). NLRs respond to a variety of PAMPs of bacterial and viral origins, as well as to cellular damage signals. NLRP1 senses a protease from *Bacillus anthracis*, NLRC4 in complex with NAIP is responsible for flagellin and bacterial type III secretion system (T3SS) recognition [69]. NLRP3 ligands include pathogen derived toxins, such as Syk kinase of *Candida albicans*, pore-forming toxins of *S. aureus* or *L. monocytogenes* and alpha-toxin of *S. aureus*, or host-derived molecules indicating cellular stress like extracellular ATP, ion influx, amyloid-β peptide, uric acid crystals or elevated level of extracellular glucose [70-75].

Some ALRs and NLRs are reported to retain their autoinhibited state in the absence of signal by interaction between ligand-binding and signaling or NOD domains [76-78]. Ligand binding results in oligomerization of the receptors. AIM2 and IFI16 create fibrils along dsDNA stretches. Whereas IFI16 oligomerization is driven by clustering of pyrin domains, DNA-binding HIN-domains of AIM2 are sufficient for filament formation [79, 80]. Though no high-resolution structure is available so far, based on homology with apoptosome and electron microscopy studies of NLRP1 NLRs are believed to assemble into a ring-like structure via CARD-domain association [81-83]. ALRs and NLRs oligomerization results in PYD or CARD-domain aggregation that recruits adaptor protein apoptosis-associated speck-like protein containing a CARD (ASC) via PYD-PYD or CARD-CARD domains interactions, respectively (Figure 2) [83-85]. Interaction of ASC with PYD-containing ALRs or NLRs nucleates ASC^{PYD} filament formation that brings its CARD domains in close proximity and thus creates a platform for caspase-1 recruitment and activation [86-89]. NLRs containing a CARD domain can directly bind and activate caspase-1, however, the presence of ASC as mediator of inflammasome formation results in more efficient signaling (Figure 2) [78, 85, 90]. Several oligomerization steps of inflammasome formation result in a star-shaped protein assembly with ALR/NLR-ASC complex attached to long caspase-1 filaments [87]. Activated caspase-1 processes IL-1β and IL-18 and cleaves gasdermin D, which induces pyroptosis by forming membrane pores (Figure 2) [84, 91-93].
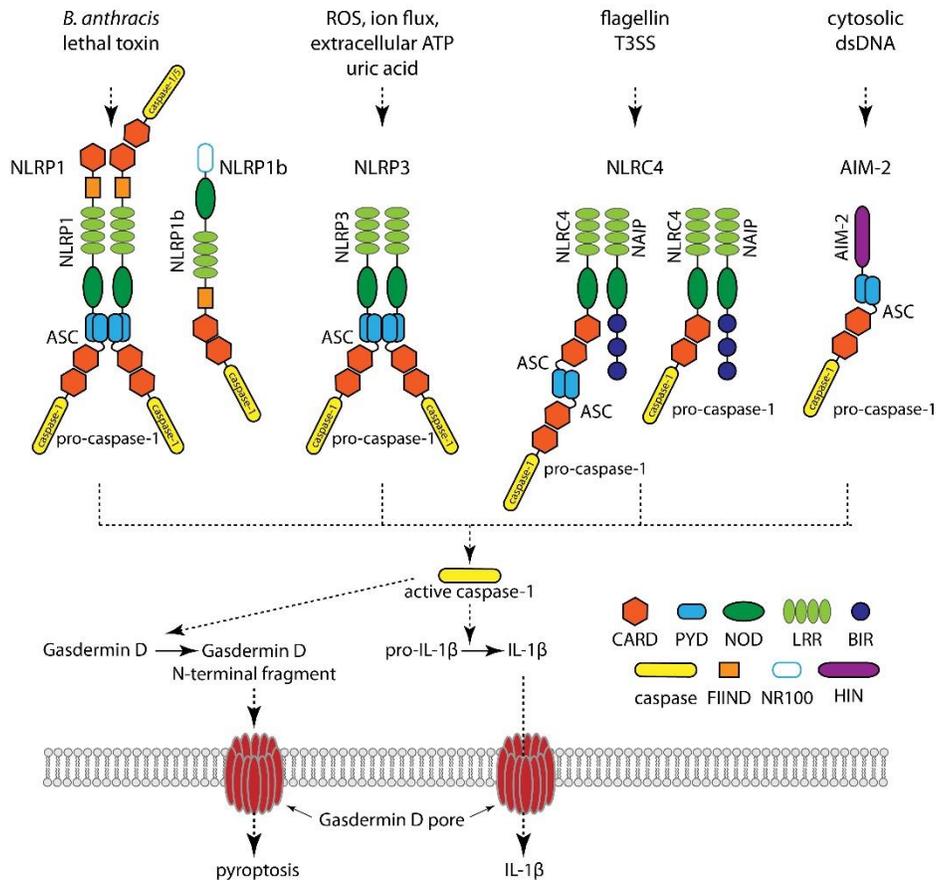
Figure 2 Inflammasome signaling pathways.
NLRP1 (NLRP1b in mice), NLRP3, NLRC4/NAIP and AIM-2 recruit adaptor ASC via PYD-PYD interactions. CARD domains of ASC assemble pro-caspase-1 filaments via CARD-CARD interaction leading to caspase-1 activation and cleavage of pro-IL-1β and gasdermin D followed by gasdermin D-mediated pore formation, pyroptotic cell death and IL-1β release. Depicted domains include CARD (caspase activation and recruitment domain), PYD (pyrin domain), NOD (nucleotide oligomerization domain), LRR (leucin-rich repeats), BIR (baculoviral inhibitory repeat), caspase, FIIND (function-to-find domain), NR100 (amino-terminal domain of rodent NLRP1 of about 100 amino acids), HIN (hemopoietic expression, interferon-inducibility, nuclear localization). A schematic representation of domains is shown in figure. Adopted from *Man S.M. and Kanneganti T.-D., 2016* [94].

Except for PAMP sensing and pyroptosis, inflammasomes are described to be involved in physiological processes unrelated to inflammation. For example, NLRP1 or NAIP/NLRC4 inflammasomes activation is associated with caspase-1 mediated activation of phospholipase A2 and eicosanoid synthesis [95]. NLRP3 inflammasome supports phagosome maturation and antigen cross-presentation in macrophages [96]. NLRC4 induced activation of caspase-1 was reported to impair autophagy through TRIF cleavage, as well as to cleave glycolysis-related enzymes [97, 98]. ALRs and NLRs can also function inflammasome-independently. For instance, AIM2 is involved in colorectal cancer suppression, NLPR3 influences transcription program in T helper cells by directly engaging with nuclear transcription factors, NLRP6 negatively regulates NF- κB activation by Toll-like receptors, but stimulates type I IFN response upon EMCV infection [99-102].

3.3. RIG-I-like receptors

Unlike Toll-like receptors and AIM-2/NOD-like receptors RIG-I-like receptors (RLRs) are specialized in foreign RNA recognition. RLRs trigger innate immune response counteracting broad range of RNA viruses like influenza A virus, Newcastle disease virus, Sendai virus, vesicular stomatitis virus, hepatitis C virus and picornaviruses by detecting their genomes or replication products [103]. DNA viruses such as herpes simplex virus-1, adenovirus, Epstein-Barr virus, vaccinia virus and hepatitis B virus were also reported to trigger an immune response through RLRs [104, 105]. RLRs are expressed in most tissues, enabling immune sensing in all cell types.

RLR family consists of three receptors: retinoic acid-inducible gene I (RIG-I), melanoma differentiation-associated gene 5 (MDA5) and laboratory of genetics and physiology 2 (LGP2) [106]. Though cellular RNAs like mRNA, tRNA, rRNA and miRNA are also present in the cytosol, their length, structure and modifications restrict their recognition by RLRs. RIG-I specifically recognizes dsRNA >19 bp with 5'-tri or di-phosphates, though short fragments of poly(I:C) and very long dsRNA molecules were reported to be bound by RIG-I independently of 5'-phosphates [107-110]. Interestingly, even nucleoprotein-encapsidated dsRNAs with 5'-triphosphate are recognized by RIG-I [111]. RIG-I is also indirectly involved in poly(dA:dT) DNA sensing through its RNA intermediate transcribed by RNA polymerase III [112]. MDA5 is known for recognition of long > 1kbp dsRNA, poly(I:C) and AU-rich viral mRNAs originating from measles virus [113-115]. RIG-I and MDA5 are also suggested to have direct antiviral activity and to interfere with viral polymerases by binding to the replicating viral genomic RNA [105, 116]. LGP2, the only RLR lacking downstream signaling domains, preferentially binds dsRNA ends and was first described as negative regulator of RIG-I and MDA5 signaling [106, 117, 118]. However, more recent studies revealed LGP2 to assist in MDA5-RNA interaction and enhance MDA5 RNA recognition of EMCV virus [119, 120].

RLRs comprise of an SF2 ATPase domain and a C-terminal domain (CTD) that are responsible for RNA binding and ATP binding and hydrolysis (Figure 3). RIG-I and MDA5 additionally have an N-terminal tandem caspase recruitment domain (2CARD) for signal transduction. The CTD is responsible for substrate specificity: in RIG-I it specifically binds 5'-triphosphate, in LGP2 – dsRNA termini, and in MDA5 – dsRNA stem [121-124]. ATPase activity of RLRs is supposed to weaken dsRNA binding and thus support discrimination between self and non-self dsRNAs. ATP hydrolysis is reported to disassemble MDA5 fibrils formed on short dsRNAs thereby driving MDA5 filament formation only to long virus-derived RNAs [125]. Mutations in the RIG-I ATPase domain trapping the protein in its ATP-bound state were described to result in a constitutive activation of RIG-I due to its association with self-RNA and elevated IFN response in the absence of infection [126, 127].

Inactive RIG-I is present in an autoinhibited state with CARDs bound to the ATPase domain. Upon RNA and ATP binding the ATPase domain together with CTD forms a ring around dsRNA, and CARD domains of RIG-I are released to enable downstream signaling [128, 129]. CARD domains of RIG-I then assemble into a helical tetramer that is additionally stabilized by ubiquitin chains thus forming a scaffold for interaction with the adaptor MAVS (Figure 3) [130-132]. Unlike RIG-I, CARD domains of MDA5 are not inhibited by protein conformation [133]. MDA5 forms a ring-like structure on long dsRNA similar to RIG-I and undergoes further oligomerization through interactions between ATPase domains thus bringing
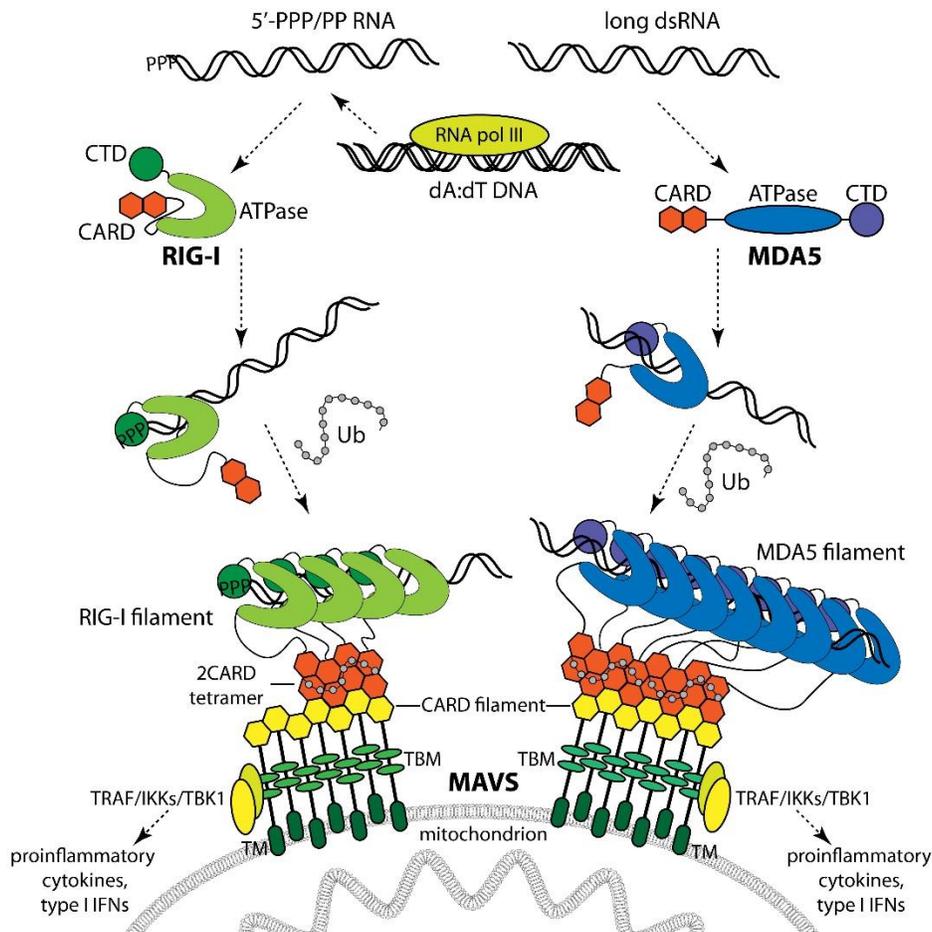
Figure 3 RIG-I and MDA5 RNA-sensing pathways.
RIG-I and MDA5 recognize 5'-PP/PPP and long dsRNA, respectively, that emerge from viruses or reverse transcribed by RNA polymerase II (RNA pol III) AT-rich dsDNA and form oligomers. 2CARD signaling domains assemble into tetramers stabilized by ubiquitin (Ub) and initiate oligomerization of CARD domains of MAVS. Further recruitment of TRAFs results in activation of IKK and TBK1 kinases resulting in phosphorylation of transcription factors and synthesis of type I IFNs and proinflammatory cytokines. Adopted from *Yoneyama M., et al, 2015* [134].

CARD domains in close proximity and facilitating their interaction (Figure 3) [125, 135, 136]. According to structural studies, LGP2 also assembles into filaments similar to MDA5 proposing filament nucleation to be a mechanism for its MDA5 signaling enhancement [137]. In presence of ATP RIG-I also can form active filaments along dsRNA, similar to those formed by MDA5 [138]. Taken together, dsRNA-driven oligomerization of RIG-I and MDA5 represents a general mechanism of cooperative dsRNA sensing by RLRs.

Activated RIG-I and MDA5 recruit adaptor mitochondrial antiviral signaling (MAVS) through CARD domains interaction (Figure 3) [139]. CARD-filaments originating from RIG-I and MDA5 are believed to serve as nucleation point for MAVS prion-like fibrils formation on mitochondrial surface [140, 141]. Tumor necrosis factor Receptor-Associated Factor (TRAF) binding motifs (TBMs) of MAVS oligomers bind TRAF proteins that activate IKK and TBK-1 kinases resulting in NF-κB or IRF3 activation and IFN response [142-147]. RLRs key features and pathways are summarized in Figure 3.

3.4. OAS proteins

Another class of RNA-sensing PRRs is represented by 2'-5' oligoadenylate synthase (OAS) proteins. OAS/RNase L pathway is one of first discovered innate immune pathways leading to IFN production and counteracting viral infection. OASs are template-independent nucleotidyltransferases that recognize intracellular dsRNA and produce 2′-5′-linked oligoadenylate (2-5A) second messenger [148, 149]. This second messenger binds and activates a downstream effector – endoribonuclease L (RNase L) that degrades viral and host RNA preventing protein synthesis and causes apoptosis of the infected cell (Figure 4) [150-152]. OAS proteins are widely expressed in different cell types and their expression is enhanced in the presence of type I IFN [148, 149, 153]. OAS/RNaseL axis detects dsRNA emerging from viral genomes or during the viral replication and is involved in antiviral defense against picornaviruses, such as encephalomyocarditis virus (EMCV), vaccinia virus, hepatitis C virus (HCV) and HIV [154-158].

OAS/RNaseL homologs can be found in most tetrapods, however, OAS-like proteins exist throughout metazoans and can be found even in sponges, indicating an ancient origin of 2'-5'-specific catalysis [159-161]. The human OAS family consists of 4 members – OAS1, OAS2, OAS3 containing one, two or three basal OAS units, respectively, and OAS-like protein (OASL) comprised of one OAS unit and a C-terminal ubiquitin-like domain [162-166]. OAS1 and OAS2 both exist as two splice variants with different C-terminal parts, whereas only one isoform of OAS3 was detected [167-169]. Different OAS proteins and their isoforms are characterized by different expression level, localization and biological functions [170]. For example, OAS3 was demonstrated to be localized with ribosomal RNA, whereas OAS2 was reported to be associated with nuclear envelope and rough endoplasmic reticulum and OAS1 localizes to the nucleus and cytoplasm or can have antiviral functions in the extracellular space [162, 171-174]. OAS proteins are expressed during different stages of viral infection and combine their functions for an optimal antiviral activity [175-177].

OAS proteins are nucleotidyltransferases and one OAS unit consists of a two lobes characteristic for this family of enzymes (Figure 4) [178, 179]. OAS1-3 bind dsRNA with a positively charged "spine" between two lobes that results in active site rearrangement and show preferential binding to long dsRNAs [178-180]. Other OAS ligands include ssRNA with defined secondary structure, such as a stem-loop of HIV mRNA in 5'-untranslated region [181]. Activated OAS binds two Mg^{2+} ions and two ATP and catalyzes a nucleophilic attack by the 2'OH of the acceptor ATP or polyadenylate to triphosphate of donor ATP – the reaction mechanism proposed for the whole class of nucleotidyltransferases [179, 182, 183]. 2'-5' bond specificity emerges from the positioning of an acceptor ATP in the active site: in contrast to 3'-5' specific nucleotidyltransferases such as poly(A)-polymerase, OAS1 structure revealed the absence of bulky residues in proximity to the acceptor binding site, thus enabling the rotation of the acceptor ATP and positioning of 2'-OH at the reactive center [179, 184]. The base of the donor ATP is coordinated with unspecific hydrophobic interactions that explain the possibility of other OAS substrates. OAS can catalyze $pppA(pA)_n(2'-5')pG$ and $pppG(2'-5')pG$ formation, as well as elongate an RpA molecule like NAD^+, Ap_4A or tRNA [185-188]. Unlike OAS1-3, OASL has mutations in catalytic residues responsible for triphosphate recognition and therefore lacks enzymatic activity [164, 165].

The activity of OAS 1 and OAS2 depends on an intact tripeptide motif (CFK) that is necessary for their tetramerization and dimerization, respectively [170, 189]. OAS3 and OASL lack the conservation in tripeptide motif resulting in monomeric form of these proteins [170]. The oligomerization of OAS

correlates with the enzyme processivity: tetrameric OAS1 and dimeric OAS2 produce trimeric and tetrameric 2'-5'-linked polyadenylates that are capable of strong RNaseL activation and the mutations disrupting oligomerization of the enzymes at least for OAS2 abolish the enzyme activity [190-193]. In contrast to OAS2, OAS1 was also reported to be active in monomeric state [173]. Monomeric OAS3 synthesizes only dimeric pppA(2'-5')pA that are not efficient RNaseL activators [193, 194].

Polyadenylates produced by OAS proteins are bound by autoinhibited monomeric RNaseL that leads to its homodimerization and activation (Figure 4) [195, 196]. Active RNaseL degrades viral and cellular ssRNAs including rRNA and preferentially cuts after UA and UU sequences [197-200]. The resulting fragments of cellular RNA and dsRNA fragment arising from cleavage of single-stranded parts of viral replication products are recognized by the cell as non-self RNAs and are reported to stimulate IFN response through RIG-I and MDA5 pathways [201, 202]. RNaseL antiviral functions also include induction of apoptosis of the infected cell further eliminating viral infection (Figure 4) [203, 204].

OAS proteins were reported to have RNaseL-independent functions. Enzymatically inactive family member OASL was shown to increase resistance to single-stranded RNA and DNA viruses via its C-terminal ubiquitin-like domain and to suppress viral replication by enhancing RIG-I signaling by poly-ubiquitin mimicking [205, 206]. Mouse OASL1, on the other side, inhibits translation of IFN-regulating transcription factor 7 (IRF7) and thus reduces type I IFN production resulting in increased viral persistence [207, 208]. OAS3 was shown to have RNaseL-independent antiviral function in response to Chikungunya virus [209].

The importance of OAS/RNaseL axis in antiviral defenses is supported by the evolvement of viral mechanisms inhibiting the pathway. For example, L protein from Theiler's virus directly inhibits RNaseL and phosphodiesterases from murine coronavirus and group A rotavirus were reported to degrade polyadenylates to reduce RNaseL activation [210-212].

3.5. cGAS-STING pathway

DNA emerging from pathogens or damaged cells is also a potent DAMP recognized by the innate immune system [4]. Unlike pathogen-associated RNAs that can be distinguished by the various structural and chemical features from host RNAs, sensing of foreign DNA relies mostly on DNA compartmentalization [4]. The presence of DNA in cytosol is a danger-associated signal and misplaced self-DNA and DNAs emerging from viral or bacterial genomes are detected in cytosol by a range of PRRs. cGAS/STING axis displays a central pathway of cytosolic DNA sensing and is responsible for DNA recognition in the majority of cell types [213]. Other DNA-sensing PRRs like TLR9 and AIM2 can not fully rescue IFN induction in the absence of cGAS/STING [214, 215]. The cGAS/STING axis is responsible for detection of various DNA viruses, such as herpes simplex virus 1 (HSV-1), murine gamma-herpesvirus 68 (MHV68), Kaposi's sarcoma-associated herpesvirus (KSHV), vaccinia virus (VACV), hepatitis B virus (HBV), retrovirus HIV-1 through its replication intermediates, as well as bacterial and synthetic dsDNAs [213, 215-223]. Cellular DNA emerging in the cytosol is also an activator of the cGAS/STING pathway. For example, DNA exposure to UV irradiation or reactive oxygen species results in DNA oxidation and its recognition by cGAS/STING [224]. Mitochondrial nucleoids were recently reported to leak to the cytosol under mitochondrial stress conditions and induce cGAS/STING mediated type I IFN synthesis [225-227]. Chromatin fragments arising from missegregation of DNA
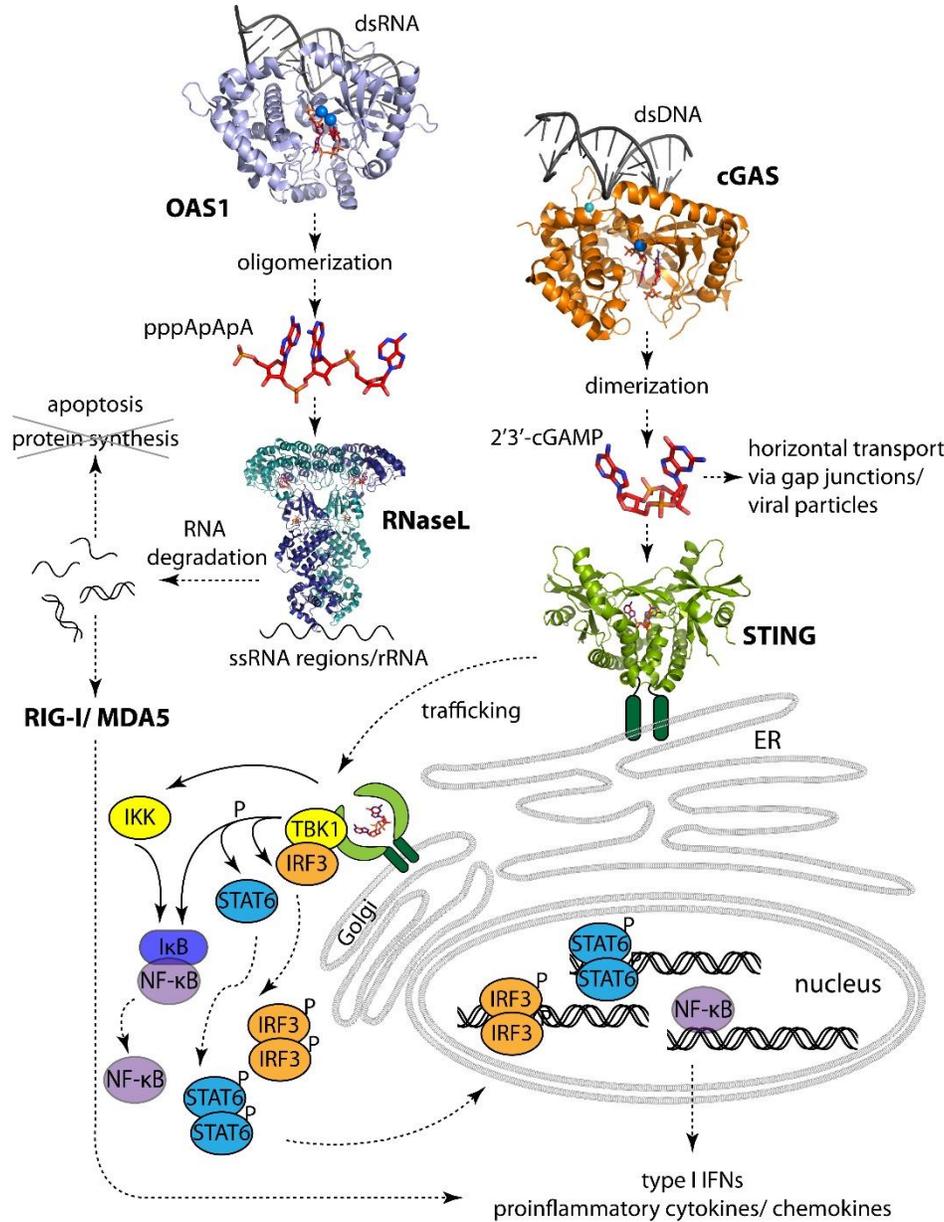
Figure 4 cGAS/OAS nucleic acid sensing pathways.
Upon dsRNA binding OAS (PDB:4RWN) produces polyadenylates (pppApA$_n$) that activate RNaseL (PDB:4OAU). RNaseL digests cellular and viral ssRNA abolishing translation and produces short dsRNA fragments recognized by RIG-I and MDA5.

cGAS (PDB: 4KB6) is activated by cytosolic dsDNA and synthesizes 2'3'-cGAMP that activates the adaptor STING (PDB: 4KSY) residing on endoplasmic reticulum (ER). STING undergoes trafficking to Golgi complex in perinuclear region and recruits kinase TBK1 for IRF3, STAT6 and NF-κB transcription factors activation. Additionally, STING mediates NF-κB activation through IKK kinases. Active IRF3, NF-κB and STAT6 are translocated into the nucleus and facilitate expression of type I IFNs and proinflammatory cytokines and chemokines.

during the mitosis are also detected by cGAS/STING [228, 229]. Interestingly, cGAS was also found to recognize chromatin fragments in senescent and cancer cells and to promote proinflammatory response

and paracrine senescence through production of various cytokines, chemokines, extracellular matrix proteins and growth factors, collectively referred to as the senescence-associated secretory phenotype (SASP) [230, 231].

Cyclic GMP-AMP synthase (cGAS) is a dsDNA sensor binding various DNA species sequence-unspecifically [232]. It belongs to the family of nucleotidyltransferases and shares a similar fold with OAS proteins (Figure 4) [233-235]. The dsDNA ligand is bound with a positively charged helix on the opposite side of the active center similar to OAS and introduces structural changes in the active site of cGAS [233, 234, 236]. Though long dsDNA is reported to be the most potent cGAS activator, ssDNA and dsRNA can also be bound by cGAS, however, binding of these ligands does not lead to cGAS activation unless ssDNA has a secondary structure with double-stranded stem regions [232, 233, 235, 237]. Active cGAS generates a second messenger cyclic GMP-AMP (cGAMP) from ATP and GTP (Figure 4). Unlike bacterial cyclic dinucleotides, composed of 3'-5'-phosphodiester linkages, cGAS catalyzes formation of an uncanonical 2'-5' linkage between 2'OH of GMP and 5'-phosphate of AMP [234, 238, 239]. In contrast to OAS that produces linear 2'-5' polyadenylates in a one-step mechanism, cGAS synthesizes two phosphodiester linkages via two-step mechanism catalyzed by a single active site. During the first step GTP and ATP bind to acceptor and donor binding sites, respectively, and pppG(2'-5')pA linear intermediate is formed. In order to perform the cyclization, the linear product has to rebind to cGAS in reverse order, so adenine part in positioned in acceptor and guanine part in donor binding sites, so the second 3'-5' linkage can be synthesized [234, 240]. Resulting 2'3'-cGAMP can be transferred between the adjacent cells via gap junctions or by packaging into viral particles and induce IFN signaling not only in infected cells, but also in the neighboring cells [241-243].

Second messenger cGAMP binds and activates downstream adaptor stimulator of interferon genes (STING) residing on endoplasmic reticulum (Figure 4) [215, 244, 245]. STING is a dimeric membrane protein with large C-terminal cytosolic domain and a flexible C-terminal tail responsible for signal transduction [244, 246, 247]. cGAMP binds between monomeric parts of STING and introduces an active closed STING conformation [248-250]. Upon cGAMP binding STING is translocated to the perinuclear region and recruits TANK-binding kinase 1 (TBK1) responsible for phosphorylation of the C-terminal domain of STING [143, 215, 251]. Phosphorylated STING builds a complex with TBK1 and interferon regulatory factor 3 (IRF3) resulting in IRF3 phosphorylation, dimerization and translocation into the nucleus, where is initiates expression of type I IFNs (Figure 4) [245, 247]. STING was also reported to recruit transcription factor STAT6 in a TBK-1-dependent manner leading to chemokine production [252]. Moreover, STING was proposed to induce NF-κB transcription factor activation through kinases TBK-1 and IKK (Figure 4) [253]. Interestingly, cGAS and STING are both IFN-stimulated genes and their expression is strongly enhanced by type I IFNs creating a positive feedback loop [254, 255].

STING serves not only as downstream adaptor of cGAS, but also as direct sensor of bacterial CDNs. Though 2'3'-cGAMP was shown to be the most affine STING ligand, STING triggers type I IFN signaling in response to c-di-GMP and c-di-AMP produced by such bacteria, as *L. monocytogenes*, *M. tuberculosis* and *C. trachomatis* [256-260]. Surprisingly, the cGAS/STING pathway was also reported to be involved in antiviral defense against some RNA viruses, such as West Nile virus or vesicular stomatitis virus, presumably through cGAS activation with intrinsic DNAs or STING trafficking or interaction with RIG-I [215, 216, 244].

Structural basis of cGAS/STING sensing and its regulation will be described in more detail in the next chapters.

3.6. Other DNA sensors

The cells are equipped with DNases, such as TREX1, DNaseI and DNaseII residing in cytosol, extracellular space and lysosomes, respectively, which eliminate unwanted DNA accumulation. However, the malfunction or inactivation of these DNases by pathogens triggers a potent type I IFN response and is associated with autoimmune diseases [261-264]. Apart from cGAS, several other proteins were reported to be involved in type I IFN signaling upstream of STING, including DAI, DNA-dependent serine/threonine protein kinase complex (DNA-PK), MRE11, DEAD box protein 41 (DDX41), polyglutamine binding protein 1 (PQBP1), Sox2 and IFI16. Though cGAS seems to be the central cytosolic dsDNA sensor, these proteins may play an important role in IFN signaling in specific cases and cell types.

One of the first proposed DNA sensors to be involved in IRF3-dependent type I IFN production was DNA-dependent activator of IFN-regulatory factors (DAI, also known as Z-DNA binding protein 1, ZBP1) [265]. Overexpression of DAI resulted in enhanced production of type I IFNs and its knock-down in the murine fibroblast L929 cell line strongly inhibited IFN response to HSV-1. Co-immunoprecipitation assays showed association of DAI with TBK1 and IRF3, suggesting the mechanism of DAI-mediated IRF3 activation [265]. Moreover, DAI was shown to be necessary for type I IFN synthesis and inhibited human cytomegalovirus (HCMV) replication in human fibroblasts [266]. The later studies, however, did not confirm the role of DAI in type I IFN induction in mouse bone marrow–derived macrophages, mouse embryonic fibroblasts or *in vivo* [68, 267]. DAI was also shown to stimulate necrosis in fibroblasts during murine cytomegalovirus infection and thus to activate antiviral response [268].

DNA damage response proteins were also reported to participate in cytosolic DNA sensing. DNA-PK complex is a complex of catalytic subunit DNA-PKcs kinase with Ku70 and Ku80 and is known to be responsible for non-homologous end joining in DNA repair [269]. Remarkably, DNA-PKcs was shown to co-localize with cytosolic viral DNA and its knock-out impaired IRF3-dependent IFN response to dsDNA and modified virus Ankara (MVA) in mouse embryonic fibroblasts [270]. Another complex member Ku70 responsible for DNA-ends recognition was recently discovered to recognize cytosolic DNA and interacts with STING for IRF1 and IRF7-dependent synthesis of IFN-λ1 (a member of type III IFN) [271, 272]. Other proteins normally involved in DNA-damage response - meiotic recombination 11 (MRE11) and RAD50 - were also described to participate in cytosolic DNA sensing. MRE11 and RAD50 together with Nijmegen breakage syndrome 1 (NBS1) form the MRN complex that is crucial for double-strand DNA breaks recognition and repair [273]. MRE11 knock-down or mutations associated with ataxia-telangiectasia–like disorder dramatically decreased type I IFN production in response to transfected DNA but not *Listeria monocytogenes* or HSV-1 and its involvement in IFN expression is STING-dependent [274]. The importance of MRE11 for cytosolic DNA sensing is supported by the finding that LANA protein of Kaposi Sarcoma herpesvirus (KSHV) inhibits NF-κB signaling through MRE11 thus promoting lytic replication of KSHV [275]. RAD50 was reported to activate NF-κB and to facilitate IL-1β production in response to cytosolic DNA via its direct interaction with caspase recruitment domain-

containing protein 9 (CARD9) and deficiency of this pathway resulted in reduced production of IL-1β in response to vaccinia virus *in vivo* [276].

DExD/H box helicase DDX41 was also revealed to induce type I IFN response to cytosolic DNA with involvement of STING, TBK1 and IRF3 [277, 278]. DDX41 was shown to bind to both DNA and STING and its knock-down abolished type I IFN production in response to transfected DNA, HSV-1 or *Listeria monocytogenes* infection in myeloid dendritic cells [278]. Another study shows the ability of DDX41 to directly bind bacterial second messengers cyclic di-GMP and cyclic di-AMP and induce type I IFN response [279]. E3 ubiquitin ligase TRIM21 was shown to target DDX41 for degradation resulting in decreased levels of IFN-β in response to intracellular DNA, further supporting the involvement of DDX41 in cytosolic DNA sensing [280]. However, several studies did not confirm the importance of DDX41 for DNA-induced type I IFN production [223, 281]. Other DExD/H box helicases were described to be involved in DNA signaling. DHX9 and DHX36 were reported to bind CpG DNA and participate in MyD88-dependent TNF-α and IFN-α signaling, as well as take part in NF-kB and IRF3 activation in response to dsRNA in MAVS-dependent manner in dendritic cells [282-284].

Another protein - polyglutamine binding protein 1 (PQBP1) – was shown to act as cGAS co-factor in HIV-1 sensing [219]. According to the study, PQBP1 binds HIV transcript, directly interacts with cGAS and is necessary for cGAS activation by HIV-1 transcript upstream of STING in human monocyte-derived dendritic cells and monocyte-like THP-1 cell line. Interestingly, such effects were not observed for DNA or RNA viruses, suggesting PQBP1 as a specific binder of retroviral DNA products. Moreover, mutations in PQBP1 were found to be associated with reduced immune response to HIV-1 infection in patients. [219].

High-mobility-group box (HMGB)-containing proteins Sox2 and HMGB1 may also participate in cytosolic DNA recognition. Sox2 is a transcription factor playing an important role during development and tissue regeneration [285, 286]. In neutrophils Sox2 was discovered to be responsible for sequence-specific recognition of bacterial genomes leading to expression of proinflammatory cytokines, but not IFNs [287]. Upon DNA binding Sox2 dimerizes and forms a complex with kinase TAK1 and its binding partner TAB2 leading to activation of transcription factors NF-κB and AP-1. Other hematopoietic cells were shown to activate cGAS-STING pathway upon DNA stimulation and do not express Sox2 [287]. HMGB1 is a nuclear architectural chromatin protein participating in chromatin maintenance, nucleosome sliding and transcription regulation [288-290]. Secreted HMGB1 was reported to be involved in inflammation acting as a cytokine interacting with TLRs and chemokine CXCL12 and to trigger cytokine production in monocytes [291-297]. However, according to recent studies HMGB1 and its homologs HMGB2 and HMGB3 play a crucial role in recognition of intracellular nucleic acids [298]. Depletion of HMGB1 resulted in reduced IFN production in response to intracellular DNA and poly(I:C), the lack of HMGB2 impaired response to intracellular DNA and the loss of all HMGBs lead to defects in AIM2-dependent IL-1β production and reduced IFN response to both DNA and RNA viruses. HMGBs depletion also reduced TLR3, TLR7 and TLR9 signaling [298]. However, the exact nature of these effects remains to be discovered.

IFI16 was also suggested to be a DNA sensor. IFI16 is a member of PYHIN family of proteins and shares similar domain organization with AIM2 [299]. IFI16 recognizes foreign DNA in the nucleus by scanning the long stretches of unprotected pathogenic DNA and assembling into oligomeric structures [300].

Though IFI16 is a nuclear protein, it was discovered to co-precipitate with cytosolic DNA and co-localize with dsDNA during HSV-1 infection [68]. Moreover, IFI16 forms filaments upon DNA binding similar to those of AIM2 and was proposed to induce inflammasome formation upon Kaposi's sarcoma herpesvirus or Epstein-Barr virus infection [67, 76, 80, 301]. The depletion of IFI16 or its mouse homolog p204 resulted in impaired IFN response to HSV-1 and HIV-1 in STING-dependent manner [68, 237, 302]. IFI16 was also shown to directly interfere with viral replication [303]. IFI16 was linked to cGAS sensing and the direct interaction of the proteins and their co-localization on cytosolic DNA were shown [304]. Importance of IFI16 in DNA sensing is further supported by the presence of viral proteins targeting IFI16 for degradation [305]. However, the latest studies revealed the role of IFI16 as cGAS/STING co-factor, rather than an independent DNA sensor. Two recent studies showed that both cGAS and IFI16 are essential for IFN response, since IFI16 directly interacts with STING and promotes its association with TBK1, whereas the direct influence of IFI16 on cGAS activity remains controversial [306, 307].

4. Structural basis of cytosolic DNA recognition by cGAS-STING pathway

4.1. cGAS structure and activation mechanism

Since its discovery in 2013 many structures of cGAS from different species, in different activation states and with various substrates were published. cGAS belongs to nucleotidyltransferase family of proteins, which share a similar protein fold and catalyze phosphodiester bond formation between donor nucleoside monophosphate and acceptor hydroxyl group residing on protein, nucleic acid or small molecules [182, 308-310]. Nucleotidyltransferases are involved in such processes, as replication and repair (DNA polymerases, DNA-ligases), transcription and RNA processing (RNA-polymerases, mRNA capping enzymes) and signal transduction (OAS, cGAS). Though the members of different nucleotidyltransferase classes share little similarity on the sequence level, they all share a common α/β-fold with conserved catalytic aspartates/glutamates responsible for divalent ions coordination and polarization of the acceptor hydroxyl group and hG[GS] pattern (h – hydrophobic amino acid) that helps by substrate binding [308, 309]. The NTase core is composed of three-stranded β-sheet flanked by 4 α-helices with the catalytic triade of aspartates/glutamates positioned on the β-strands and can be accompanied by additional domains for different families of nucleotidyltransferases [308].

cGAS is a 60 kDa protein with an unconserved unstructured N-terminus of 150 amino acids and highly conserved C-terminal catalytic male abnormal 21 (Mab21) domain that belongs to the NTase superfamily and shares sequence homology with catalytic domain of OAS1 [223, 310]. The Mab21 domain was first described for *Caenorhabditis elegans* protein Mab21 and is sufficient for both DNA binding and enzymatic activity in cGAS [223, 311]. The Mab21 domain of cGAS was a subject of extensive structural studies [233-236, 312-314]. In contrast, the exact function and structure of the N-terminal part of cGAS remains unclear. Several studies revealed that it participates in DNA binding and thus increases cGAS affinity towards DNA and gains its fold first upon DNA binding [223, 315].

The catalytic domain of cGAS has an α/β-fold and harbors a bilobal structure with a catalytic cleft between two lobes connected by a long "spine" helix (Figure 5a). The N-terminal Lobe 1 represents an NTase fold with 8 β-sheets flanked by two helices. Lobe 1 contains the catalytic residues E225, D227 and D319 (numbering is given for human cGAS, correspond to E211, D213, D307 in mouse and E200, D202,

D294 in pig) on the side of the cleft that coordinate two catalytic Mg^{2+} ions and the nucleotides (Figure 5b). GS-motif characteristic for NTAses is represented by G212/S213 (G198/S199 in mouse and G187/S188 in pig) and is located between the first and second β-strands in an "activation loop" (Figure 5b). C-terminal Lobe 2 consists of α-helices and a Zn-thumb with a highly conserved histidine and cysteines H390, C396, C397 and C404 (H378, C384, C385 and C392 in mouse and H365, C371, C372 and C379 in pig) coordinating the Zn^{2+} cation (Figure 5c). The molecular surface opposite to the active site between the "spine" and Zn- thumb is flat and positively charged and forms a "platform" for DNA binding (Figure 5c). Inactive cGAS is a monomeric protein with a rather unstructured nucleotide-binding activation loop. Apo-cGAS was co-crystallized with UTP and a variety of CDNs including reaction
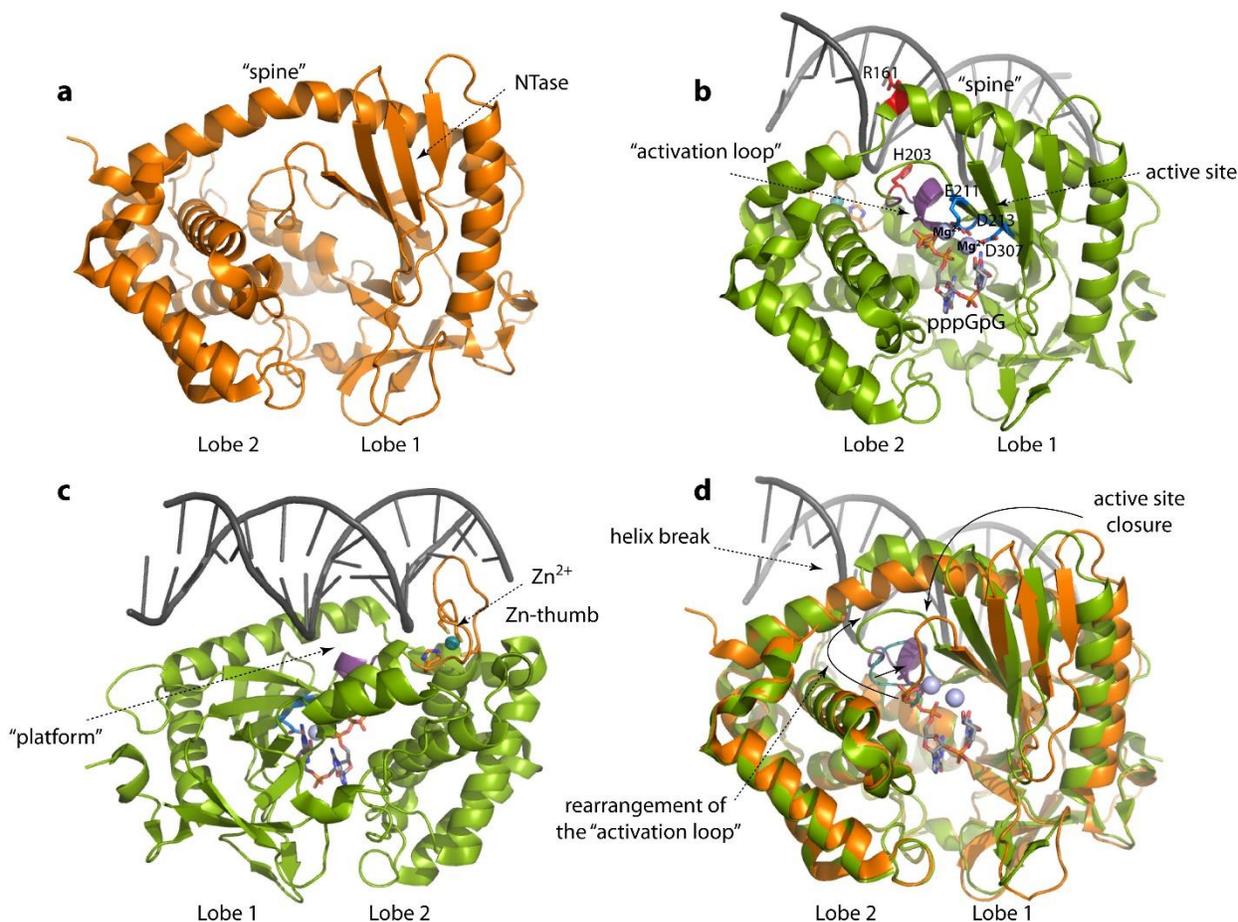
Figure 5 cGAS structure.
   (a) Apo-form of mouse cGAS Mab21 domain (PDB: 4K8V, orange) represents a bilobal structure (N-terminal Lobe 1 and C-terminal Lobe 2) connected with a "spine" helix.
   (b) DNA-bound mouse cGAS (PDB: 4K98, green). Activation loop (purple) rearrangement opens cGAS active site. Catalytic residues (blue) coordinate two Mg^{2+} cations (spheres) and linear reaction product (pppGpG). DNA binding is stabilized by R/H fingers (red).
   (c) DNA-binding by cGAS. DNA interacts with a positively charged "platform" and is stabilized by Zn-thumb (orange) with chelated Zn^{2+} cation (blue sphere).
   (d) Superposition of apo- (orange) and DNA-bound (green) cGAS forms. Conformational changes upon cGAS activation are indicated with arrows and include "spine" helix break, rearrangement of the activation loop (teal in apo-cGAS and purple in DNA-bound cGAS) and closure of the active site. Active cGAS coordinates Mg^{2+} (spheres) and linear intermediate (pppGpG) for the second reaction step.

product 2'3'-cGAMP, however, though some coordination of catalytic residues took place, the activation loop was not fully ordered and the substrate and product were not correctly positioned [233, 312, 313]. Surprisingly, cyclic di-UMP introduced changes in the activation loop of apo-cGAS corresponding to cGAS's active state, whereas the catalytic pocket remained unchanged, suggesting that activation loop and active site conformations are two distinct events needed for cGAS activation [312]. Moreover, the B-factors of the Lobe 1 containing catalytic residues were two times higher than those of C-terminal Lobe 2, indicating a high level of active site flexibility in apo-form of cGAS [314].

Dramatic changes in cGAS active site are induced by B-form DNA binding (Figure 5d). DNA is bound by the cGAS "platform" between "spine" α-helix and a "Zn-thumb". Consistent with sequence-independent DNA binding, cGAS contacts mostly the sugar-phosphate backbone of DNA. Positively charged residues of the "spine" helix and the "platform" interact with the minor groove of DNA via side- or main-chain contacts, whereas the "Zn-thumb" grips the DNA at the major groove resulting in cGAS specificity towards B-form DNA and not A-form RNA ligands. DNA binding is additionally stabilized by two arginine/histidine fingers R176/H217 (R161/H203 in mouse, R150/R192 in pig), positioned on the "spine" helix and "platform", respectively, that are inserted into the minor groove of DNA (Figure 5d). Upon DNA binding the "spine" helix breaks in two parts inducing Lobe 1 closure and the opening of the entrance to catalytic pocket of cGAS. The catalytic residues undergo realignment and the "activation loop" with conserved GS motif is restructured allowing Mg^{2+} coordination and effective nucleotide binding (Figure 5d) [233, 234, 314].

Though apo-cGAS is a monomeric protein, a closer look at the structure of apo- and DNA-bound cGAS reveals the packing of cGAS into dimers (Figure 6) [233, 234, 313, 314]. Small-angle X-ray scattering and analytical ultracentrifugation studies confirmed that DNA binding introduces $cGAS_2:DNA_2$ complex formation with 16-20 bp DNA [313, 314]. Two DNA molecules in such complex are "sandwiched" between two cGAS protomers and point to

Figure 6 cGAS dimer.
cGAS protomers (green and orange, PDB: 4LEZ) bound to two dsDNA molecules (grey). A single cGAS protomer binds two DNA molecules with DNA-binding sites A (includes "spine" helix) and B (shown with arrows). Annotated DNA-binding residues of site A (cyan) and site B (red), as well as protein-protein contacts in dimer interface (blue) are represented as sticks. Zn^{2+} ions are depicted as blue spheres.

each other, so that the same binding mode will result in clashes between long DNA strands, if cGAS is bound to the middle of such long DNA. cGAS:DNA 2:2 complex is stabilized by protein-protein interactions between cGAS molecules and the secondary DNA-binding site that contributes to ~35-40% of the total surface area between cGAS and DNA. This secondary DNA-binding site (site "B") interacts with the same DNA molecule that is bound by the main DNA-binding site (site "A" formed by "spine"-helix, "platform" and Zn-thumb) of partner cGAS protomer (Figure 6). Site "B" forms two salt bridges to
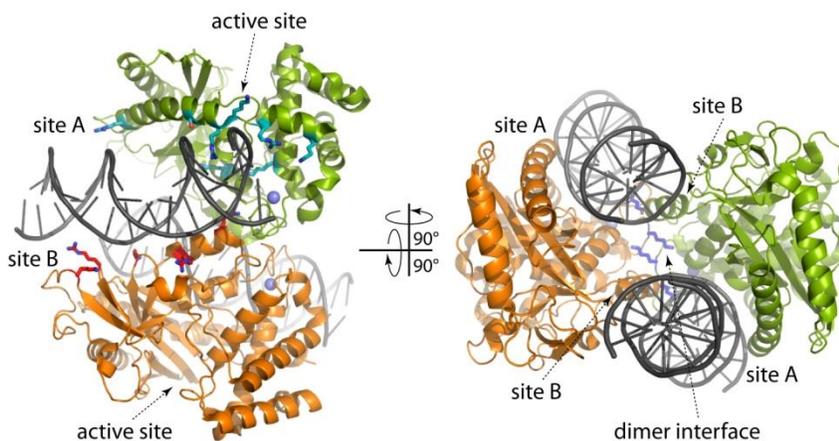
sugar-phosphate backbone of DNA (residues R236/K347 in human, R222/K335 in mouse and R211/K322 in pig) and an arginine finger (L354 in human, R342 in mouse and R329 in pig) is inserted into the minor groove but does not interact with the DNA backbone. Protein-protein contacts are mediated by electrostatic interactions between K394 and E398 (K382/E386 in mouse and K369/E373 in pig) residues of two cGAS protomers and are situated in the Zn-binding loop (Figure 6). Mutational studies *in vitro* and in the cells confirmed the importance of these contacts for cGAS dimer formation and its activity. Amino acid substitutions in protein-protein interaction loops or DNA-binding site "B" decreased DNA binding, impaired 2:2 complex formation and significantly reduced cGAS activity for both full length and Mab21-domain containing constructs indicating that cooperative binding of DNA by both DNA-binding sites and cGAS dimerization are necessary for cGAS activity [313, 314].

The DNA-bound cGAS fold is very similar to the RNA-bound OAS1 state, likely representing a common activation mechanism for these evolutionary connected enzymes (Figure 4) [178, 179, 316]. RNA-binding induces the same break in the "spine" helix and active site rearrangement necessary for subsequent nucleotide and Mg^{2+} binding by OAS1. Unlike cGAS, OAS1 does not have a Zn-thumb and binds two minor grooves segments by its interactions with the sugar-phosphate backbone of RNA. Moreover, cGAS possesses a second DNA-binding site and forms 2:2 active complexes, whereas no high-order structures of OAS1:RNA were observed [178, 313, 314]. Taken together, cGAS and OAS1 use differences in nucleic acid topology and nucleic acid binding mode to distinguish between DNA and RNA ligands. Intriguingly, according to the published structures, both OAS1 and cGAS nucleic acid binding sites are much shorter than the reported RNA or DNA length needed for enzymes full activation [180, 281]. Thus, additional mechanisms might be needed to achieve full activation of cGAS and OAS1.

4.2. Mechanism of the second messenger cGAMP synthesis by cGAS

Bacteria are well known for using CDNs as signal molecules involved in such processes, as biofilm formation, antibiotic signaling, virulence and bacterial stress response [317]. However, bacterial diguanylate and diadenylate cyclases differ significantly from eukaryotic OAS and cGAS proteins. The most of bacterial CDN-synthases are obligate oligomers in which each of the monomers catalyzes formation of one phosphodiester linkage between ATPs or GTPs [318-320]. In contrast, cGAS generates a mixed-based cyclic dinucleotide G[2′–5′]pA[3′–5′]p (2'3'-cGAMP) – a novel second messenger consisting of one 3'-5' and one unusual 2'-5' linkage - with one active site [234, 239, 240]. The same 2'-5'-linkage seems to be a hallmark of eukaryotic signaling shared by both cGAS and OAS proteins [149]. Interestingly, prokaryotic enzyme DncV from *Vibrio cholerae* was discovered to generate a mixed-based G[3′–5′]pA[3′–5′]p (3'3'-cGAMP) [321]. Despite a very low sequence homology with cGAS, DncV is a structural and functional homolog of cGAS revealing evolutionary connection between prokaryotic and eukaryotic signaling systems [322].

Upon DNA binding cGAS coordinates ATP and GTP in the catalytic site (Figure 7). During the first step of reaction donor- and acceptor-binding pockets are occupied by ATP and GTP, respectively, so that pppG2'-5'pA linear intermediate can be produced. Two nucleotides are sandwiched between T321 and Y436 (I309/Y421 in mouse and I296/Y411 in pig) hydrophobic residues and stacked in 90° rotated orientation [233]. Triphosphate from ATP is coordinated by two Mg^{2+} ions bound to catalytic triade E225, D227 and D319 (E211, D213, D307 in mouse and E200, D202, D294 in pig) and additionally stabilized

by hydrogen-bonds formed by residues S213 and S435 (S199/S420 in mouse and S188/S410 in pig). The adenine base is stacked with Y436 (Y421 in mouse and Y411 in pig) and the ribose of ATP makes hydrogen-bonds with E383 (E371 in mouse and E358 in pig) providing additional coordination [233, 234]. GTP binding to acceptor pocket is stabilized by hydrogen bonding with S378, S380, T211 and R376 (S366/S368/T197/R364 in mouse and S353/S355/T186/R351 in pig) providing specificity towards guanine base (Figure 7). Nucleotide coordination results in close proximity of α-phosphate of donor ATP to the attacking 2'-OH of GTP, 2'-OH polarization by D319 (D307 and D294 in pig) and linear product pppGpA formation (Figure 8a) [233]. Consistent with specific hydrogen-bonding to guanine base and only stacking interaction with adenine, cGAS can tolerate GTP as donor and produce cyclic di-GMP [234, 240].

Figure 7 Coordination of ATP and GTP in cGAS catalytic pocket.
OH groups of ribose in ATP and GTP base are coordinated with hydrogen bonds (dashed lines) with E371 (green) and T197/S368/S366/R364 (cyan) residues, respectively. Numbering is given for mouse cGAS.

The cyclization step requires a 180° flip of linear intermediate in cGAS catalytic pocket, so that guanine and adenine bases occupy donor- and acceptor-binding sites, respectively (Figure 8b). It was recently proposed to dissociate and rebind to cGAS competing for binding with ATP and GTP [312]. According to structural studies of cGAS in complex with pppG(2'-5')pG or pG(2'-5')pA, triphosphate of the linear product is coordinated by Mg^{2+} cations analogically to ATP binding mode during the first step of reaction (Figure 8c). Guanine and adenine bases reside in *syn* and *anti* glycosidic torsion orientations in pppG(2'-5')pA, respectively, and guanosine ribose is flipped in donor site in comparison to ATP binding mode during the first reaction step [234]. Such orientation leads

Figure 8 A mechanism of 2'3'-cGAMP synthesis by cGAS.
(a) The first reaction step: after DNA binding ATP and GTP are sandwiched between residues Y421 and I309 (teal) and triphosphate of ATP is positioned by Mg^{2+} ions coordinated by catalytic triad (blue) and hydrogen bonds (dashed lines) with S420 and S199 (purple). D307 polarizes 2'OH of GTP for a nucleophylic attack (black arrows) of α-phosphate of ATP.
(b) The resulting pppG(2'-5')pA linear intermediate has to flip over (black arrow) in cGAS catalytic pocket to enable the second cyclization step.
(c) The second reaction step: 3'OH of adenosine attacks α-phosphate of pppG(2'-5')pA (black arrows) resulting in cyclic product with mixed phosphordiester linkages 2'3'-cGAMP.
(d) 2'3'-cGAMP leaves cGAS catalytic pocket.
Residue numbering corresponds to mouse cGAS.

to polarization of 3'-OH from adenosine and 2'3'-cGAMP formation that is indeed observed in crystal structure (Figure 8c) [314]. Though the riboses and bases of cGAMP interact with cGAS, the close proximity of phosphodiester bond 3'-5' pApG to catalytic negatively charged residue D227 (D213 in mouse and D202 in pig) destabilizes the binding and presumably promotes product dissociation (Figure 8d) [314].

The catalytic mechanism of 2'-5' polyadenylates synthesis by OAS1 is very similar to that of cGAS. Interestingly, OAS1 coordinates donor ATP with a number of hydrogen-bonds required for adenine specificity. Like in cGAS, OAS1 binds the donor nucleotide by unspecific hydrophobic interactions [179].

Such specificity towards acceptor adenine results in OAS adding any nucleotide to 2'-OH group of RpA [166]. However, unlike cGAS, OAS1 catalyzes only a one-step reaction producing linear polyadenylates probably by suppressing product binding in reverse order and thus preventing full cyclization.

Striking similarity in structure and catalytic mechanism can be observed between cGAS and DncV from *Vibrio cholerae* producing 3'3'-cGAMP (Figure 9) [322-324]. DncV has the same organization of the catalytic core as cGAS and OAS proteins and unlike these proteins maintains an autoactivated state by bracing two lobes with α-helices surrounding the catalytic core (Figure 9a). Interestingly, despite the constitutive activation DncV possesses a positively charged "platform" with conserved residues opposite to the active site at the position of DNA binding by cGAS. DncV also lacks regulatory loops closing the active site and maintains an open entrance into the catalytic pocket. In contrast to cGAS, DncV binds GTP in the donor pocket and produces pppA(3'-5')pG linear product during the first reaction step, however, *syn* and *anti* configurations within pppA(3'-5')pG correspond to same orientations of bases within pppG(2'-5')pA cGAS intermediate (Figure 9b, c). Strikingly, single nucleotides as well as linear intermediate in DncV is rotated ~40° away from the catalytic triad as compared to cGAS and this positioning is defined by the catalytic pocket residues of DncV and cGAS that would clash with alternative substrate or linear product binding mode (Figure 9b, c). In cGAS the long side chain of R376
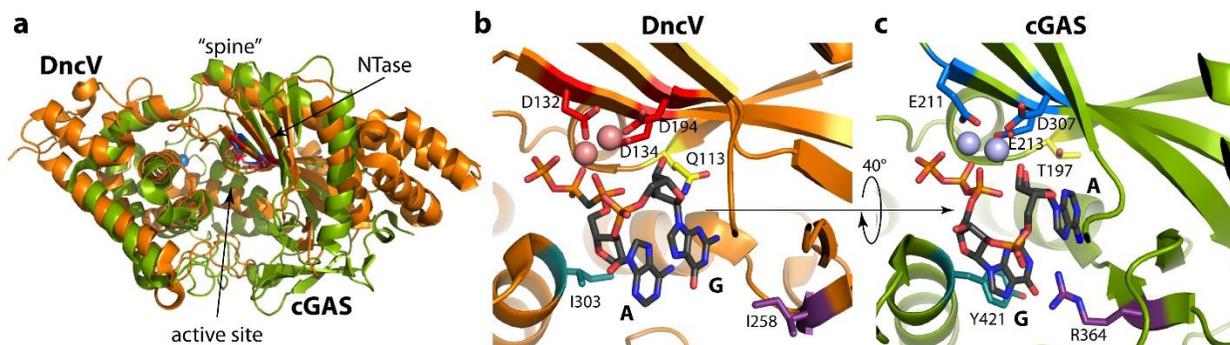
Figure 9 Comparison of cGAS and DncV structures.
(a) Superposition of cGAS (PDB: 4K98) and DncV (PDB: 4TY0). Active site residues are depicted in blue and red for cGAS and DncV, respectively.
(b-c) DncV (b) and cGAS (c) active sites with pppA(3'-5')pG and pppG(2'-5')pA linear intermediates, respectively. Active site residues (red and blue, respectively) coordinate Mg^{2+} ions (spheres) and a triphosphate. Corresponding residues responsible for substrate orientation and linkage specificity are depicted in teal, yellow and purple. Relative rotation of linear intermediates is indicated with black arrow.

residing on the opposite site of catalytic triad and short T211 (R364/T197 in mouse and R351/T186 in pig) push pppG(2'-5')pA closer to catalytic residues. DncV has a short I258 and long Q113 side chains in corresponding positions pushing the linear product away from catalytic triad. Additional coordination of intermediate in enzyme-specific way emerges from base stacking with Y436 (Y421 in mouse and Y411 in pig) in cGAS, whereas DncV and other bacterial enzymes favor non-aromatic non-polar residues in this position (I303 in DncV) (Figure 9b, c). Exchanging of cGAS R376/T211 with DncV I238/Q113 resulted in cGAS synthesizing 3'3'-cGAMP, indicating the mechanism of phosphodiester bond specificity [322].

4.3. STING structure and mechanism of activation by cyclic dinucleotides

Stimulator of interferon genes (STING, also called MITA, ERIS, MPYS or TMEM173) is a central adaptor protein for cytosolic DNA-sensing pathway and a direct sensor of cyclic dinucleotides [215, 256, 257]. Though it was also reported to interact with major histocompatibility complex class II molecules and participate in interferon response to cytosolic RNA, these findings remain controversial [244, 245, 252, 257, 325]. STING is widely expressed in endothelial and epithelial cells, as well as in T cells, macrophages and dendritic cells and its overexpression leads to strong activation of IRF3 and NF-κB transcription factors and type I IFN production [244, 245].

STING is a 42 kDa protein and consists of 4 transmembrane helices anchored to the endoplasmic reticulum (ER) membrane (aa 1-154), a large globular C-terminal domain pointing towards the cytosol that binds CDNs (155-342) and a flexible C-terminal tail (CTT, aa 343-379) responsible for signal transduction and interaction with TBK1 kinase and transcription factor IRF3 (Figure 10a) [244, 247, 250, 326]. STING recognizes a variety of cyclic dinucleotides and was shown to trigger type I IFN production in response to metazoan second messenger 2'3'-cGAMP, bacterial c-di-AMP, c-di-GMP and 3'3'-cGAMP and synthetic 3'2'- and 2'2'-cGAMP [250]. Upon activation it re-localizes to the Golgi complex and assembles into punctate structures containing TBK1 in perinuclear region [215]. STING re-localization and regulation is a function of its N-terminal transmembrane domain. Though no structural data about this region are available, deletions or mutations of the N-terminal domain result in impaired STING activity and trafficking [244, 327].

Though some structural similarities between STING and leucine-rich repeat kinase 2 can be observed, the overall STING structure is unique [248]. Some sequence similarities of STING could be found with *L. monocytogenes* diadenylate cyclase, however, there seems to be no homology between STING and bacterial CDN receptors [326]. Cytosolic CDN-binding domain structures and solution studies show that STING operates as a butterfly-shaped dimer (Figure 10b) [248-250, 328-333]. Each protomer has an α+β-fold in which a curved five-strand β-sheet is surrounded by α-helices. The dimer is held together by hydrophobic interactions between α-helices at the bottom of the V-shaped structure. Some studies proposed STING dimerization upon CDN binding to be an activating mechanism, however, the nature of STING intermolecular interaction and a strong conservation of the interacting amino acids suggest that STING maintains dimeric conformation in both inactive and activated states. The C-terminal tail of STING is flexible and was not observed in crystal structures. A region surrounding the CDN-binding cleft – "lid" region - is also unstructured in apo-STING (Figure 10b). The *goldenticket* mutation I200N (I199N in mouse) that was reported to abolish STING activity is located in the β-sheet region and impairs STING folding and expression [257, 331].

The apo-form of human STING represents an open conformation in which the distance between the "wings" of butterfly-structure is ~60Å [329, 330]. CDN binding induces a closure of the V-shape and decrease in protomer distance to ~35Å. Upon 2'3'-cGAMP binding the "lid" region of STING (aa 224-245) gains its fold and arranges into a two-strand β-sheet closing the CDN binding pocket increasing the interaction interface within the dimer (Figure 10b) [249, 250]. cGAMP interacts with STING via purine bases stacking with aromatic residues Y167 and Y240 and α-phosphates contacting R238 and R232 (Figure 10c). The interaction is further stabilized by polar contacts between free 3'-OH of 2'3'-cGAMP with S162 and between guanine base and E260 and T263 side chains resulting in specific 2'3'-cGAMP recognition (Figure 10c) [250]. 2'3'-cGAMP specificity is also defined by its intrinsic free-ligand conformation. Unlike 3'2'- and 3'3'-cGAMP, 2'3'-cGAMP conformation has a "closed" geometry with a 90° rotated guanine group stabilized by a hydrogen-bond between guanine base and adenosine phosphate
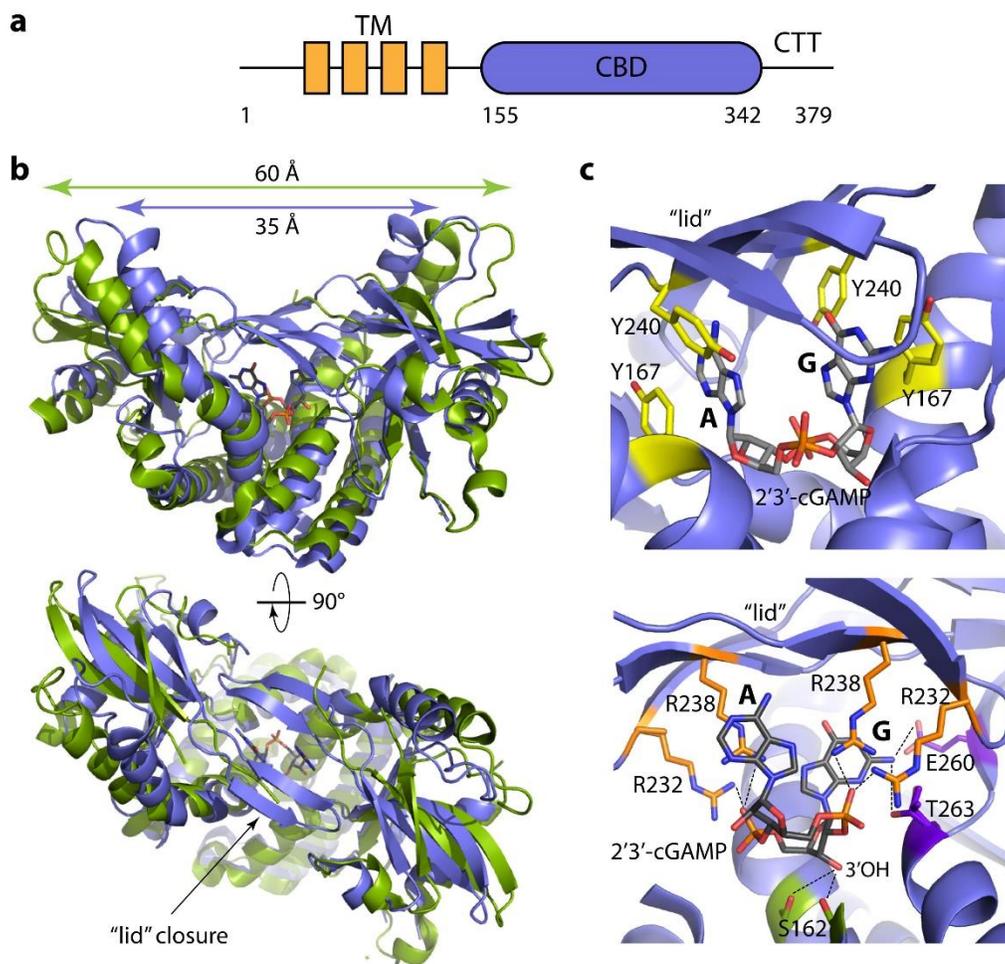
Figure 10 STING structure.
(a) Domain structure of human STING. STING consists of transmembrane domain (TM, orange), CDN-binding domain (CBD, blue) and C-terminal tail (CTT). Numbers indicate position within the sequence.
(b) Superposition of apo- (PDB: 4EMU, green) and 2'3'-cGAMP-bound human STING (PDB: 4KSY, blue).
(c) Close-up view on CDN-binding pocket (PDB: 4KSY). 2'3'-cGAMP is coordinated by stacking with aromatic residues (yellow) and hydrogen bonds (dashed lines) between arginines of the "lid" (orange) and phosphates. Additionally, 3'-OH forms hydrogen bonds (dashed lines) with S162 (green), and guanine base – with E260 and T263 residues (purple).

group. Such conformation matches well with the STING-bound 2'3'-cGAMP mode and thus requires less entropy cost for binding to STING in comparison to other combinations of phosphodiester linkages [332].

Interestingly, STING variants were discovered to have different affinities to CDNs [334]. R232H variant of human STING and R231A variant of mouse STING were shown to be activated by 2'3'-cGAMP, but not c-di-GMP [239]. By contrast, human STING[R232] and mouse STING[R231] respond to all CDNs with their affinities decreasing in the following order: 2'3'-cGAMP/3'3'-cGAMP/c-di-GMP/c-di-AMP. Though in general the CDN affinity towards STING correlates with STING stimulation potential, human STING[H232] responding only to 2'3'-cGAMP shows similar dissociation constants for 2'3'-cGAMP and 3'3'-cGAMP revealing that differential activation might occur due to additional conformational changes [249, 250, 314]. Consistent to type I IFN production rates and *in vitro* binding assays, the structures of human STING[H232] with c-di-GMP represent an open conformation [248, 329, 331], whereas one of two STING[R232] structures with c-di-GMP shows a closed active conformation with ordered lid region [328].

Unlike human STING, mouse protein resembles a closed conformation independently of CDN binding. However, the presence of c-di-AMP, c-di-GMP, 2'3'-cGAMP or small compounds 10-carboxymethyl-9-acridanone (CMA) and 5,6-dimethylxanthenone-4-acetic acid (DMXAA) induce the rearrangement of the "lid" region [249, 333, 335, 336]. Functional analysis revealed that mouse STING shows less preference for cGAS product 2'3'-cGAMP and is strongly activated by all CDNs, CMA and DMXAA [333, 337]. These results suggest that not the STING closure itself, but the proper folding of the "lid" is the major conformational change driving STING into its active state. The importance of the "lid" region of STING for CDN specificity and signaling is further supported by single mutation that increases stability of β-sheets of the "lid" driving human STING towards DMXAA sensing [249].

According to the current model CDN binding induces reorganization of the C-terminal tail of STING that serves as a scaffold for TBK1 and IRF3 recruitment and activation [247, 331]. CTT directly interacts with TBK1 and is phosphorylated at residues S355, S358, S366 and L374 that facilitates IRF3 transcription factor binding [247]. A conserved phosphorylation pLxIS (p- the hydrophilic residue, x - any residue, S - the phosphorylation site) motif common for all adaptor proteins activating IRF3 – STING, MAVS and TRIF – was shown to bind C-terminal domain of IRF3 in extremely similar mode revealing a general mechanism of IRF3 activation [338]. Moreover, the same motif is responsible for phosphorylation and dimerization of IRF3 itself and is released from an autoinhibited state by adaptor protein binding [338, 339]. Rotavirus protein NSP1 also possesses pLxIS motive on its C-terminus and is able to mimic the host adaptor proteins binding [338]. NSP1 targets IRF3 for ubiquitination and degradation thus abolishing IRF3 activation and IFN-β synthesis [340, 341].

Molecular dynamics simulations and mutational analysis of STING recently shed light on CTT rearrangement upon CDN binding. In *in silico* simulations the lid of human STING[R232] remained in stable closed conformation only in 2'3'-cGAMP-bound model, whereas c-di-GMP ligand failed to maintain the structured lid region. CTT end was making a contact with the lid regions on both STING subunits in ligand-bound STING, but only in 2'3'-cGAMP bound STING CTT arranged into a short β-sheet [342]. Notably, this local secondary structure formation appeared in the region of CTT sensible for TBK1-mediated phosphorylation and responsible for IRF3 binding [338]. Mutations impairing this β-sheet formation decreased IFN-β production. The importance of CTT interaction with the lid region of STING was also shown by exchanging the non-conservative lid residue K224 in human STING to methionine,

present in mouse protein in the same position. The resulting increase of lid hydrophobicity leads to more stable interaction of the CTT end with the lid region and higher IFN-β induction by human STING [342]. This observation also clarifies a surprisingly reduced activity of human STING in comparison to its mouse homolog and lower CDN specificity of mouse STING [239, 249, 333, 342].

4.4. Evolutionary origins of cGAS and STING

Though IFN signaling is a feature of the innate immune system in vertebrata, homologs of cGAS and STING can be found in most metazoans and in unicellular choanoflagellate *M.brevicollis* – a member of a phylum evolutionarily close to animals [343, 344]. Other kingdoms – fungi, plants and protists – do not have cGAS/STING homologs leaving these proteins to be a unique feature of multicellular animals. cGAS and STING ancestors can be found as early in metazoan evolution, as in cnidarians like *Nematostella vectensis*, but some phyla like nematodes, some insects and flat worm *Schistosoma mansoni* lost them during evolution (Figure 11). Interestingly, most analyzed animals either have both cGAS and STING homologs or have lost both proteins. Except for some species having two or three homologs of cGAS or STING, most of the metazoans possess only one copy of each gene.

Bioinformatic and structural analysis of STING homologs revealed that CDN recognition is one of the most ancient characteristics of STING [343-345]. All STING homologs share a conserved dimerization and CDN-binding domains. The amino acid sequence analysis revealed a strong conservation of the residues responsible for CDN binding (ex. Y167 and R238) and lid closure (ex. R232) [343, 344]. STING proteins from different phyla are capable of CDN binding, except for STING homologs from the *Drosophila* family that do not interact with CDN (Figure 11). Surprisingly, STING from most vertebrates including non-mammalian species *Danio rerio* and *Xenopus tropicalis* shows preferential binding of 2'3'-cGAMP compared to other CDNs [345]. This indicates that specific 2'3'-cGAMP recognition is not the most recent evolutionary achievement and evolved in the beginning of evolution of vertebrates. Despite the poor similarity on the amino acid sequence level, STING proteins seem to share a common dimeric structure and CDN binding mechanism. The structure and CDN binding mode of the most ancient STING from *Nematostella vectensis* is nearly identical to that of the human protein. Crystal structure of *N. vectensis* (nv) STING shows a possibility for both "open" and "closed" conformation corresponding to previously reported human and mouse STING, respectively. Likely, the apo-STING is a flexible dimer and upon CDN binding adopts an active conformation similar from see anemone to humans. Indeed, though nvSTING apparently recognizes 3'3'-cGAMP in see anemone in its "open" conformation, its binding mode to 2'3'-cGAMP is nearly indistinguishable from that of human STING indicating that the nature of the second messenger and not the changes of the protein itself resulted in different STING conformations upon CDN binding. Though the major conformation and CDN binding residues are conserved in all metazoan species, subtle changes in the lid region moved the mode of CDN recognition from a simple base staking in *N. vectensis* to specific recognition of phosphodiester linkages in mammals [345].

Though the CDN-binding domain seems to be highly conserved in structure and function within all metazoans, STING homologs share similarities in other domains as well. Most STING homologs were discovered to have three to four N-terminal transmembrane helices, except for three arthropod species [343]. In contrast to CDN-binding domain and N-terminal transmembrane part, C-terminal tail responsible for signal transduction in mammals is a rather recent evolutionary feature and was found only
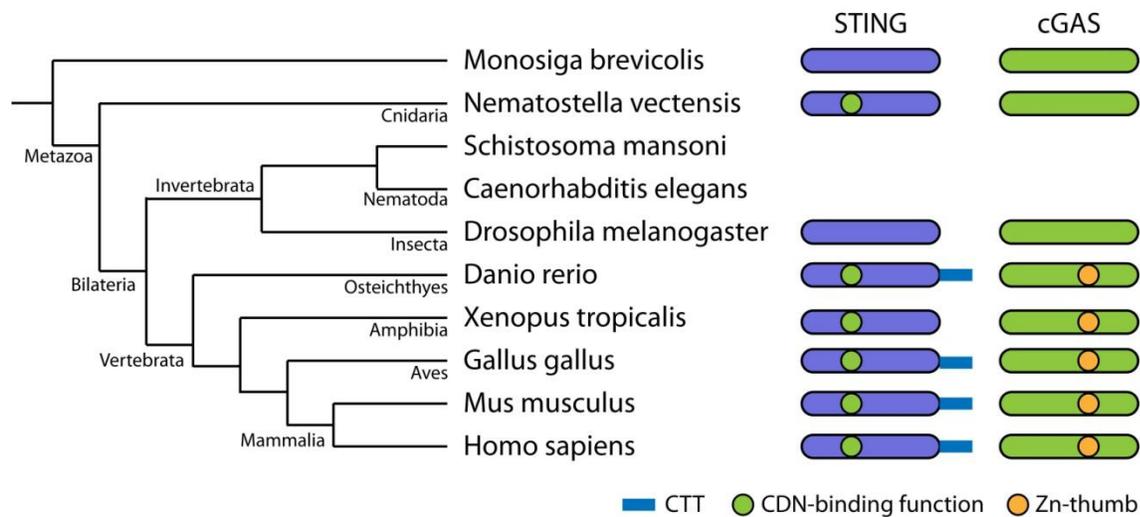
Figure 11 Distribution of cGAS and STING homologs in metazoans.
Branch lengths in species tree are not scaled. STING (blue) and cGAS (green) homologs are depicted as ovals corresponding to the species and were defined by *Wu X., et al, 2014* [343]. C-terminal tail (CTT) in STING homologs is depicted as blue line. CDN-binding function of STING is indicated with green circle and was defined by *Kranzusch P.J. et al, 2015* [345]. cGAS homologs possessing Zn-thumb are indicated with orange circle. Adopted from *Margolis S.R., et al, 2017* [346].

in vertebrates, except for amphibians, that correlates with the evolvement of the IFN system (Figure 11) [247, 343]. The CTT is critical for TBK1 recruitment and activation in humans and zebrafish and even can trigger IFN production when attached to *N. vectensis* STING homolog [247, 338, 345, 347]. This illustrates how IFN signaling might have evolved from STING ancestors with only CDN binding function. However, the functions of the cGAS-STING system in invertebrates seem to be independent of the CTT and remain to be discovered.

Like STING, cGAS homologs can be found throughout the most of the metazoan species and *M. brevicollis* (Figure 11) [343, 344]. However, the functionality of these proteins is difficult to predict, since many other NTases and Mab21-containing proteins, like Mab21L1 or Mab21L2 proteins involved in development, exist that despite a tremendous similarity to cGAS in sequence and structure do not bind DNA and produce CDN [348-350]. Moreover, cGAS shows a remarkable structural and functional similarity to OAS1 and bacterial DncV protein producing linear polyadenylates and 3'3'-cGAMP, respectively, making it even more difficult to search for functional analogs of cGAS [178, 322].

Sequence analysis revealed a strong conservation of the key cGAS residues throughout the evolution [343]. Except for the catalytic triad present in all NTases including OAS1 and DncV, cGAS homologs in vertebrates and cephalochordates share conserved substrate/product binding residues (K362, R376 and S378 of human cGAS). The DNA-binding residue K407 on the main cGAS DNA-binding site and Y436 participating in base stacking are highly conserved in all metazoans. Higher degree of conservation can be observed in vertebrates in the main DNA-binding helix, the dimer interface and the second DNA-binding site of cGAS that differs from other phyla.

According to bioinformatic analysis of the domain organization cGAS homologs gained a Zn-thumb DNA-interacting insertion and therefore their DNA-binding function in vertebrata (Figure 11). Moreover,

vertebrate cGAS homologs share a long (~167 aa) N-terminal domain, except for chicken and turkey, whereas the N-terminal cGAS domain in other phyla is much shorter (7-70 aa). According to recent studies, the N-terminal domain of cGAS plays an important role in DNA binding along with a Zn-thumb. The gain of DNA-binding domains correlates with the evolvement of STING signaling C-terminal tail during evolution with an exception in amphibian (Figure 11). The Zn-thumb was observed in the cGAS homolog in *Xenopus tropicalis* that does not have a CTT on its STING homolog. Interestingly, evolution of the downstream components of cGAS-STING pathway matches the evolvement of CDN binding properties of STING rather than cGAS DNA-binding properties. Though IFN signaling and transcription factors IRF3 and 7 evolved only in vertebrates together with cGAS DNA-binding and STING C-terminal tail, kinases TBK1 and IKKε, as well as NF-κB transcription factor can be found in cnidarians, cephalochordates, mollusks and insects that may indicate a mechanism of CDN-induced NF-κB activation different from a well-studied cGAS-STING-IRF3 signaling [346].

Though the most functional features of cGAS seem to have evolved in vertebrates, other metazoans including *N. vectensis,* as well as bacterium *Vibrio cholerae* seem to have structural and functional analogs of cGAS that may have other functions instead of DNA recognition [322, 345]. However, only in primates cGAS and OAS seem to have undergone similar rapid evolution in nucleic acid binding sites driven by positive selection [351, 352]. Though cGAS dominates DNA-sensing in vertebrates, other species likely utilize alternative pathways.

4.5. Other Mab21-domain containing proteins

Apart from the best studied male-abnormal 21 (Mab-21) family member cGAS (Mb21D1) 11 other Mab-21 proteins were discovered in humans [310]. They include Mb21D2, three Mab21-like proteins Mb21L1, Mb21L2 and Mb21L3, mitochondrial dynamics proteins 49 and 51 (MiD49 and MiD51), transmembrane protein 102 (TMEM102, also common β chain associated protein, CBAP), inositol 1,4,5-trisphosphate receptor-interacting protein (ITPRIP, also DANGER), ITPRIP-like proteins 1 and 2 (ITPRIPL1 and ITPRIPL2) and an uncharacterized protein C2orf54. The Mab-21 family proteins are poorly characterized, though some of their functions are linked to development and mitochondrial fission.

TMEM102 (CBAP) was reported to have a pro-apoptotic function under granulocyte-macrophage-colony-stimulating factor (GM-CSF) depletion. TMEM102 was shown to interact with GM-CSF/interleukin-3 (IL-3)/IL-5 receptor in the absence of the ligand and to induce mitochondrial dysfunction and therefore apoptotic cell death [353]. TMEM102 was also shown to regulate chemokine-promoted T-cell trafficking and T-cell adhesion, as well as to participate in T-cell receptor signaling and thus to modulate apoptosis of thymocytes during negative selection [354, 355]. Another member of the family ITPRIP – a membrane associated protein with partial Mab-21 domain – was reported to bind and co-localize with inositol 1,4,5-trisphosphate receptors (IP$_3$R) in neuronal membranes. ITPRIP allosterically enhances Ca^{2+} inhibition of IP$_3$R-mediated Ca^{2+} release and modulates Ca^{2+} dynamics [356]. ITPRIP was also shown to bind and inhibit death-associated protein kinase (DAPK). ITPRIP depletion resulted in increased sensitivity to cell death stimuli and brain damage after excitotoxicity revealing the function of ITPRIP to regulate neuron viability [357]. ITPRIP inhibition of DAPK also results in impairment of DAPK-p53 signaling axis leading to attenuation of anchorage-dependent apoptosis. High-glucose conditions lead to high ITPRIP expression levels and thus to decreased efficiency of apoptosis

and radioresistance of non-small cell lung cancer cells [358]. C2orf54 was also described in the context of cancer among genes with down-regulated expression in esophageal squamous cell carcinoma [359].

Mitochondrial dynamics proteins of 49 and 51 kDa (MiD49 and MiD51, respectively, or SMCR7L) were first described to cause altered mitochondrial distribution in random cellular localization screen [360]. Both proteins are anchored in the outer mitochondrial membrane and have a cytosolic Mab-21 domain (Figure 12). MiD49 and Mid51 were shown to form rings around mitochondria and recruit the fission mediator GTPase dynamin-related protein 1 (Drp1) to mitochondrial surface independently from other factors, such as mitochondrial fission protein 1 (Fis1) or mitochondrial fission factor (Mff) [361, 362]. However, MiD49 and MiD51 are involved into the cross-talk between Drp1 and Mff and suppress Mff-dependent enhancement of GTPase activity of Drp1[363]. Knock-out of MiD proteins results in irregular distribution and fusion of mitochondria, whereas their overexpression sequesters Drp1 function and blocks mitochondrial fission [361, 364]. Tough these two proteins are redundant and have similar functions, MiD49 and MiD51 are differentially expressed in human tissues and during development and have different oligomerization states suggesting the possibility for distinct though similar functions of these proteins [365]. MiD51 forms mostly dimers by interaction between cytosolic domains, whereas MiD49 forms higher oligomers through interactions between its transmembrane domains and can even form heterodimers with MiD51 (Figure 12) [365, 366]. MiD proteins are involved in sensitivity to apoptotic stimuli and are controlled by E3 ubiquitin ligase MARCH5 targeting MiD49 for degradation and reducing cell sensitivity to stress-induced apoptosis [363, 367-369].

A number of Mab-21 family members are involved in cell fate determination and development. The first described Mab-21 gene was found in *Caenorhabditis elegans* responsible for cell fate determination of several cells in *C. elegans* mail tail and rays formation that lead to family name male-abnormal 21 (Mab-21) [370]. Mab-21 mutants had also pleotropic effects on movement, body shape and fecundity of *C. elegans* [311]. Further analysis revealed that Mab-21 acts downstream of transforming growth factor-β-like (TGF-β-like) small (sma) pathway regulating nematode body length and ray formation [371]. Components of this pathway are homologs to *Drosophila* and vertebrate SMAD proteins participating in signal transduction from TGF-β receptor [372, 373]. Mab-21 is negatively regulated by TGF-β member cet-1, which is a ligand of the sma pathway, and may participate in chromatin remodeling via regulation of histone deacetylase activity [371, 374]. Vertebrate homologs of *C. elegance* Mab-21 gene - Mab21L1 and Mab21L2 - were found to be involved in eye, preputial gland and ventral body wall development [375-379]. Mab21L1 and Mab21L2 have most probably evolved by ancestral Mab21 gene duplication prior to vertebrate divergence [380]. Both proteins have strongly intercepting expression pattern and are present in retina, midbrain, spinal cord, branchial arch and limb bud during development, though sub-localization of Mab21L1 and Mab21L2 differ [376, 381-383]. Mab21L1 and Mab21L2 are also expressed in adult tissues such as cerebellum and eyes [376, 384]. Depletion of Mab21L1 or Mab21L2 was reported to arrest axial turning and resulted in severe defects in neural tube and other organs' development in mouse embryos [350]. Mice lacking Mab21L1 revealed eye defects, and homozygous Mab21L2-knock-out embryos were not viable, though eye and ventral body wall defects could be observed during development [378, 379]. Mutation in R51G in one copy of Mab21L2 is associated with ocular coloboma and cataracts caused by protein instability and dysfunction during development [385]. Another mutation in Mab21L2 R51C was reported in two patients revealing a range of eye malformations, abnormal shortness and joint formation and intellectual disability [386]. More mutations in Mab21L2 (e.g. R51H, E49K and R247Q) were reported to influence eye development by impairing ssRNA binding and

Mab21L2 stability [387]. The third member of Mab21-like proteins – Mab21L3 – was first found in *Xenopus laevis* to mediate dorsoventral patterning in embryos, and its knock-down inhibits dorsalization of gastrula-stage mesoderm and neurula-stage ectoderm [388]. Mab21L3 is also required for formation of multiciliate cells and ionocytes responsible for generating extracellular fluid flow and regulating ionic homeostasis, respectively, during development of embryonic epidermis. The study showed that Mab21L3 acts downstream of Notch pathway that downregulates its expression in *Xenopus* and mouse embryos [389].

All Mab-21 proteins consist of an N-terminal NTase core followed by PAP/OAS1 substrate binding domain. Sequence analysis reveals that among human proteins only cGAS, Mab21D2 and Mab21-like group members have some conservation in catalytic residues, though only cGAS and Mab21D2 have a full catalytic triad. *C. elegans* Mab-21 has similar pattern in catalytic site, though its third catalytic residue is substituted to glutamine. Thus, only several group members seem to have catalytic activity similar to cGAS.

So far structures of human Mab21L1 and cytosolic domains of mouse MiD49 and human and mouse MiD51 were solved [348, 390-392]. All Mab-21 protein structures represent a bilobal fold with a mixed α/β topology very similar to cGAS (Figure 12). Mab21L1 structure is very similar to apo-state of cGAS with intact spine helix. MiD49 and MiD51 structures, on the other side, have a brake in activation spine helix that refers to an active state of cGAS. Unlike cGAS, other Mab-21 proteins do not have a Zn-thumb and thus are not able to bind dsDNA. Interestingly, all crystallized Mab-21 proteins have positively charged patches

Figure 12 Mab-21 proteins.
MiD49 (PDB: 4WOY, yellow) and MiD51 dimer (PDB: 4OAG, green and orange) are anchored in mitochondrial membrane and recruit Drp1 (green oval) to induce mitochondrial fission. Mab21L1 (teal) is involved in embryonic development. Different nucleotide (red) binding modes are illustrated for CTP-bound Mab21L1 and ADP-bound MiD51.

on the opposite side of the catalytic pocket in the region corresponding to DNA-binding platform of cGAS [348, 390-392]. Since Mab21L1 and Mab21L2 were reported to bind ssRNA, this could be explained as a potential ssRNA binding site, however, the same positively charged platform could be found in bacterial homolog DncV that doesn't bind nucleic acids and is constitutively active [322, 348, 387]. Differences between the proteins can be observed in their oligomeric states. Though Mab21L1 forms two stacked pentameric rings, it seems to be a monomer in solution [348]. In contrast, MiD51 was reported to form dimers in solution that are also observed in crystals [365, 390]. Unlike cGAS dimers that are held together by DNA-mediated interactions in spine helix and additional short helix between the
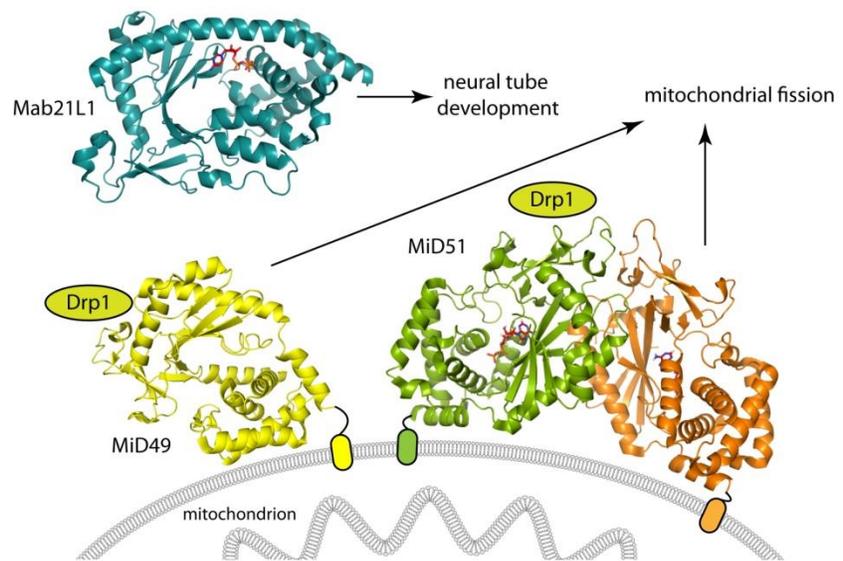
lobes, mouse MiD51 monomers make a direct protein-protein interaction using N-terminal α-helical segment (Figure 12). However, such dimerization was not necessary for Drp1 binding and function of MiD51 in the cells [390]. Surprisingly, human homolog of MiD51 does not make analogical dimeric packing in the crystal [391]. Though MiD49 structure is highly similar to MiD51, MiD49 lacks conservation of residues involved in MiD51 dimerization and is present in monomeric form without its transmembrane domain [392]. Mab21L1 and MiD51, but not MiD49 were reported to bind nucleotides, though unlike cGAS none of these proteins coordinates Mg^{2+} ions in their NTase core [348, 390-392]. Instead the binding of the nucleotides – CTP, CDP, ATP and ADP for Mab21L1, and ADP and GDP for MiD51 – occurs by direct interaction of proteins' side chains with the nucleotide base and phosphate groups [348, 390]. ADP binding by MiD51is facilitated by the residues corresponding to the catalytic triad of cGAS, but Mab21L1 seems to have a completely different nucleotide binding site close to its "spine" helix (Figure 12). Though the nucleotide binding was shown to stabilize the proteins *in vitro*, it does not introduce dramatic conformational changes in the structure and the functional importance of the nucleotide binding remains unclear.

5. Regulation of cGAS-STING pathway

The cGAS-STING pathway is a key platform not only for bacterial and viral infection detection, but also for recognition of cellular damage and tumor-derived DNA. Since the pathway plays the major role in cytosolic DNA recognition, many viral mechanisms counteracting type I IFN induction by cGAS-STING axis exist. Dysregulation of DNA sensing via cGAS-STING and its hyperactivation is associated with a number of autoimmune and inflammatory diseases. Recent discoveries described cGAS and STING as potential targets for cancer therapies. cGAS-STING pathway is regulated by a variety of posttranslational modifications, direct degradation of second messenger 2'3'-cGAMP and cellular co-factors and is tightly connected to other innate immune programs. cGAS-STING regulation mechanisms in health and disease will be addressed in detail in this chapter.

5.1. Post-translational modifications regulating cGAS-STING DNA-sensing axis

Except for regulation of cGAS activation by type and structure of the nucleic acid, cGAS is controlled by such post-translational modifications as phosphorylation, glutamylation, ubiquitination and SUMOylation affecting its DNA-binding, enzymatic activity and degradation rate (Figure 13). Akt-kinase that was identified in relation to cell proliferation and survival, as well as to metabolism, was reported to phosphorylate cGAS at S305 (corresponds to S291 in mouse protein) in cells and *in vitro* suppressing its enzymatic activity [393, 394]. Interestingly, cGAS has an Akt-kinase recognition motif [RK]x[RK]xx*[ST] (x - any amino acid) that is highly conserved in vertebrates, suggesting an evolutional importance of this regulation pathway. Overexpression of wild type cGAS but not S305A mutant incapable of phosphorylation resulted in lower levels of type I IFN production in response to transfected DNA or HSV-1 infection in mouse L929 cGAS-KO cell line, though in *in vitro* system enzymatic activity was not altered by this point mutation [394]. Akt-kinase was also shown to be activated by type I IFN in mouse fibroblasts and its depletion resulted in a dramatic reduction of IFN-induced antiviral responses, since it regulates mTOR pathway responsible for translation of IFN-inducible mRNA [395]. Moreover, Akt is activated by TBK1 kinase which apparently initiates negative feedback loop of DNA-sensing [396]. Though two opposite functions of Akt-kinase were proposed, it may act on early stages of cGAS

-31-

activation inhibiting excessive type I IFN production and supporting antiviral response. Furthermore, enzymatic activity of cGAS is regulated by mono- and polyglutamylation by tyrosine ligase-like enzyme (TTLL) 4 and 6, respectively [397]. TTLL4 attaches one glutamate group to E314 (E302 in mice) of cGAS at the entrance to the catalytic pocket rendering it catalytically inactive, whereas TTLL6 polyglutamylates cGAS at E286 (E272 in mice) and suppresses its DNA-binding ability. Glutamylation of cGAS was shown to be reversed by cytosolic carboxypeptidases (CCP) 5 and 6 that hydrolyze the glutamate chain on mono- and polyglutamylated cGAS, respectively. Thus, depletion of CCP5 or CCP6 led to higher susceptibility to DNA viruses *in vivo* and lower IFN-β production in cells [397].

Positive regulation of cGAS was recently found to be achieved by K27-linked polyubiquitination of cGAS presumably at K173 and K384. ER ubiquitin ligase RNF185 was discovered to interact with cGAS during HSV-1 infection and potentiate its catalytic activity [398]. cGAS is constitutively negatively regulated by K48-linked ubiquitination of K414 that targets cGAS to autophagic degradation [399]. However, viral infection triggering type I IFN production increases E3 ligase tripartite motif 14 (TRIM14) expression which causes cGAS stabilization. TRIM14 was shown to recruit ubiquitin carboxyl-terminal hydrolase 14 (USP14) and induce cleavage of the ubiquitin chains on cGAS at K414. Thus TRIM14/USP14 blocks cGAS degradation and generates a positive feedback loop of cGAS signaling [399]. cGAS ubiquitination was recently shown to be counteracted by SUMOylation [400]. This process is driven by ubiquitin ligase TRIM38 that dynamically SUMOylates cGAS in uninfected cells and during the early phase of viral infection attaching one or two SUMO1 moieties to K217 residue of cGAS. A mutation of this residue abolished SUMOylation and dramatically decreased cGAS protein level and IFN response to HSV-1 infection. The same study showed that at the late phase of viral infection SUMOylation of cGAS was decreased by desumoylating enzyme Senp2, that together with TRIM38 provides a mechanism for dynamic regulation of cGAS by post-translational modification [400]. Another SUMOylation event seems to occur at the late stages of infection. cGAS was found to undergo SUMOylation at residues K335, K372 and K382 of DNA-binding site which abolish DNA-binding and cGAS enzymatic activity. Sentrin/SUMO-specific protease 7 (SENP7) was found to counteract such SUMOylation and restore cGAS activity [401].

Downstream of cGAS adaptor STING is also strongly controlled by post-translational modifications (Figure 13). The necessity of STING phosphorylation by TBK1 for downstream signaling was described earlier [143]. This is further confirmed by the fact that protein phosphatase magnesium-dependent 1A (PPM1A) dampens the whole downstream phosphorylation cascades by removing S358 phosphorylation from STING [402]. Another kinase - UNC51-like kinase-1 (ULK1) – was also described to phosphorylate STING at the same position; however, it unexpectedly suppressed IRF3 activation by STING [403]. According to this study, 2'3'-cGAMP initiates a negative feedback loop triggering dissociation of ULK1 from its repressor AMP activated protein kinase (AMPK) that results in STING phosphorylation and inactivation. These conflicting results may indicate that other modifications and other players may be responsible for such opposite effects on STING activation.

STING is dynamically regulated by polyubiquitination that depending on the type of polyubiquitin and modified STING residues either activates STING or targets it to degradation. E3 ubiquitin ligases TRIM32 and TRIM56 were reported to ubiquitinate STING and facilitate its interaction with TBK1 [404, 405]. TRIM56 introduces K63-linked polyubiquitin chains to K150 and TRIM32 additionally modifies residues K20, K224 and K236, positively regulating IFN-β expression. Overexpression of these ubiquitin

ligases enhanced IFN-β promotor activation, whereas knockdown of either protein abrogated IFN signaling through STING in presence of cytosolic dsDNA or poly(I:C). Similar to TRIM32 and TRIM56 ER complex consisting of ubiquitin ligase autocrine motility factor receptor (AMFR) and its adaptor insulin-induced gene 1 (INSIG1) introduces K27-linked polyubiquitin chains to K137, K150, K224 and K236 STING residues and facilitates TBK1 recruitment and type I IFN production [406]. Deficiency of AMFR resulted in dramatic decrease of IRF3-activated genes expression in response to cytosolic dsDNA, 2'3'-cGAMP treatment or infection with HSV-1 or *Listeria monocytogenes,* and the lack of INSIG1 resulted in higher susceptibility to HSV-1 infection in mice. In contrast, K48-linked polyubiquitination of STING at the same residue K150 by E3 ubiquitin ligase RING finger protein (RNF) 5 inhibits STING-mediated signaling [407]. Overexpression of RNF5 dramatically decreased activation of NF-κB and IRF3 transcription factors upon ssRNA Sendai virus (SeV) infection and reduced IFN-β promotor activation upon transfection with poly(I:C) or dsDNA. Fluorescent microscopy and fractionation assays revealed that RNF5 ubiquitinates STING at mitochondria, where it targets STING for proteasomal
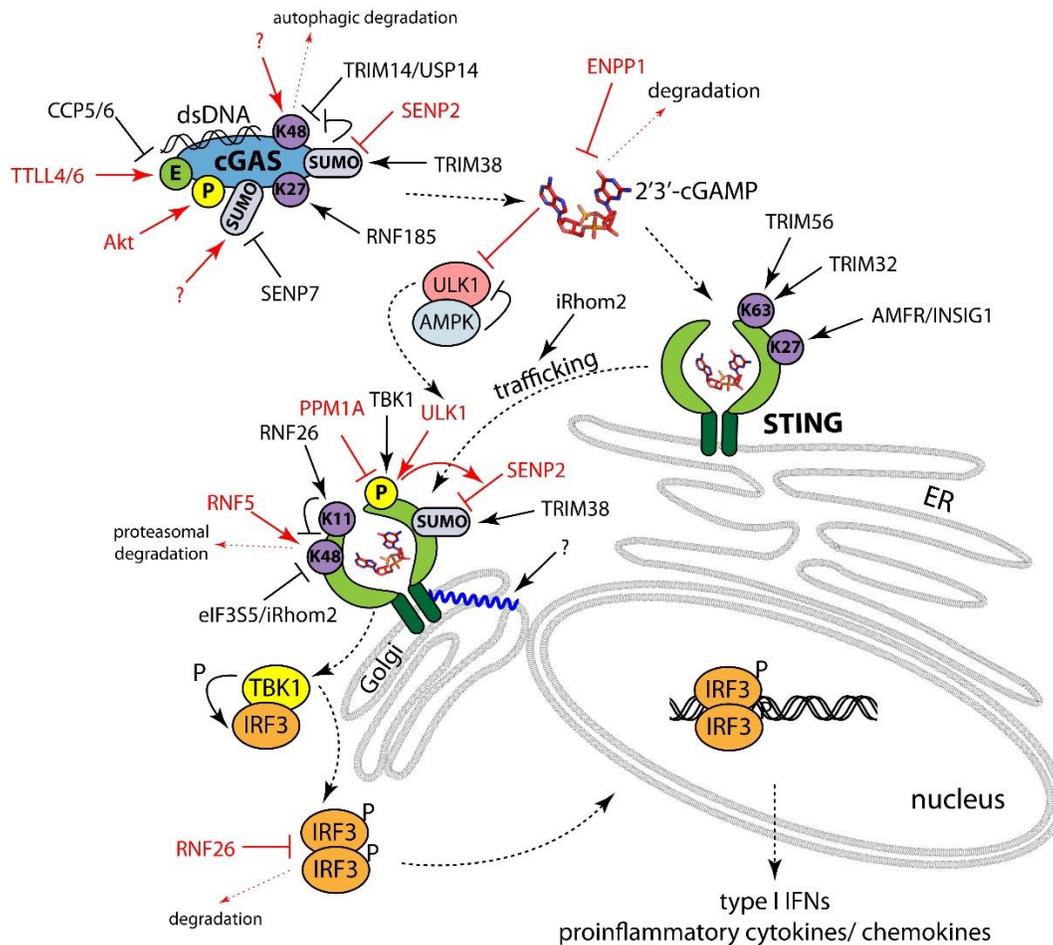
Figure 13 Regulation of cGAS/STING signaling pathway.
cGAS and STING are regulated by posttranslational modifications: phosphorylation (yellow circle, P), ubiquitination (purple circle with indicated type of linkage between polyubiquitin chains), glutamylation (green circle, E), SUMOylation (light blue oval, SUMO), palmitoylation (blue waved line). Positive regulation of cGAS/STING signaling is indicated in black, negative – in red. Arrows indicate activation, blunt lines – inhibition. Components of the pathway are connected with black dashed lines.

degradation [407]. K11-linked polyubiquitination of STING at K150 by RNF26 was described to counteract RNF5-mediated STING ubiquitination and its degradation upon viral infection [408]. In agreement to this, RNF26 knockdown inhibited activation of NF-κB and IFN-β promoters by HSV-1 and SeV infection. However, the overexpression of RNF26 promoted IRF3 degradation and reduced SeV-induced activation of the IFN-β promoter. Though no direct interaction between RNF26 and IRF3 or polyubiquitination of IRF3 was detected, RNF26 was suggested to mediate autophagic degradation of IRF3, since autophagy inhibitors or expression of catalytically inactive RNF26 mutant prevented IRF3 degradation. Apparently, RNF26 stimulates type I IFN production by STING ubiquitination at the early stages of infection and prevents excessive IFN response at the late phase of viral infection by targeting IRF3 [408]. Additionally to differential functions of STING polyubiquitination, deubiquitination was found to regulate STING activation. For example, eukaryotic translation initiation factor 3, subunit 5 (eIF3S5) was reported to be recruited to STING by inactive rhomboid protein 2 (iRhom2) and act as deubiquitinase to remove K48-linked polyubiquitin chains and thus to stabilize STING [409]. STING-stabilizing effect was observed as well by its SUMOylation. TRIM38 reported to modify cGAS and counteract its degradation can also SUMOylate STING at K338 increasing STING stability and mediating IRF3 recruitment and activation. DeSUMOylation at this position is performed by SENP2 and is facilitated by ULK1-dependent phosphorylation, giving a possible explanation for the negative regulatory effect of ULK1 [400].

STING function is also strongly regulated by its trafficking from ER to the perinuclear compartments as inhibition of such trafficking by brefeldin A abolishes STING phosphorylation and activation (Figure 13) [403]. iRhom2 apart for protecting STING from polyubiquitin-mediated degradation positively regulates type I IFN signaling by facilitating STING association with the translocon-associated protein TRAPβ and promoting STING trafficking [409]. STING was also shown to undergo palmitoylation at the in mammals conserved residues C88 and C91 taking place at the trans-Golgi network that was found to be necessary for type I FN production [410]. Palmitoylation inhibitor 2-BP abolished the expression of IFN-stimulated genes upon transfection of dsDNA or STING agonists. Furthermore, disruption of lipid rafts on Golgi inhibited STING-dependent phosphorylation of TBK1 and IRF3, suggesting that palmitoylation of STING facilitates STING clustering and thus affects TBK1 recruitment and IRF3 activation [410].

In addition to post-translational modification, cGAS and STING functions are controlled by microRNAs. MiR-576-3p expressed in response to STING and IRF3 activation was found to target STING, MAVS and TRAF3 reducing IFN response and generating a negative feedback mechanism for IFN signaling [411]. cGAS is downregulated in hypoxic tumors by miR-25/93 that target epigenetic factor maintaining cGAS basal expression level. Such cGAS repression results in immunological escape of tumor cells and tumor progression [412].

5.2. cGAMP degradation and transport

Besides STING activation and negative regulation of cGAS-STING pathway within the infected cells, 2'3'-cGAMP has a number of functions outside the cell where it was produced. For example, cGAMP was found to be and adjuvant of immune effects and was shown to boost antibody production in mice immunized with a model protein antigen ovalbumin [213, 413]. Moreover, such co-immunization resulted

in CD8 T cells expansion and higher expression levels of IFNγ and IL-2 in response to the antigen [213]. cGAMP was also shown to prolong cellular and humoral immune responses induced by influenza vaccines by intradermal immunization [414].

Additionally to its effects on T and B cell activation, cGAMP was discovered to be transferred through gap junctions in murine and human cells and participate in paracrine-like activation of the neighboring cells. Distinct connexin proteins forming gap junctions were shown to be redundant for cGAMP transfer and the depletion of one connexin member resulting in the impairment of STING activation in bystander cells was rescued by overexpression of other connexins [241]. Gap junctions were also reported to participate in tumor growth and chemoresistance. Brain metastatic cancer cells were described to form protocadherin 7 (PCDH7)-induced connexin 43 gap junctions with astrocytes and activate production of inflammatory cytokines and IFNs by transferring cGAMP. IFN-α and TNF expressed by astrocytes in turn activate STAT1 and NF-κB signaling in cancer cells leading for their expansion and survival during chemotherapy [415]. Since cGAS and STING are not always co-expressed in different cell types, the horizontal transfer of cGAMP allows not only to amplify inflammatory signals and escape viral mechanisms counteracting cell defense, but also to elicit IFN response in the cells lacking cGAS but competent for STING activation [416]. For example, T cells were found to be capable of cGAMP production but fail to activate IFN response [417]. In this case cGAMP transfer through direct cell contact is based on membrane fusion between HIV infected T cells expressing HIV envelope and primary macrophages. Such membrane fusion is triggered by interaction between the HIV-1 envelope on T cells with the CD4/co-receptor residing on primary macrophages resulting in STING-dependent type I IFN production in macrophages [418].

Another way of cGAMP transport between the cells includes its incorporation into the viral particles of herpesviruses, poxviruses and lentiviruses including HIV-1 [242, 243]. cGAMP-loaded viral particles were shown to trigger stronger cGAS-STING activation and IFN induction in newly infected cells [242]. Extracellular vesicles that are also involved in cellular mRNA and microRNA, as well as viral RNA transport between the cells, were also shown to contain cGAMP, however, their stimulation of uninfected cells was rather inefficient [243, 419, 420].

Unique 2'-5' phosphodiester linkage of 2'3'-cGAMP results not only in distinct thermodynamic properties and conformation leading to high-affinity binding to STING, but also provides a better stability towards degradation. There is a number of bacterial phosphodiesetases (PDs) known to degrade 3'-5'-phosphodiester linkages, however, only few 2'-5' PDs were discovered. For example, _Vibrio cholerae_ cGAMP PDEs (V-cGAPs) selectively hydrolyze 3'3'-cGAMP to the linear pApG product but fail to hydrolyze metazoan 2'3'-cGAMP [421]. One 2'-5'-specific phosphodiesterase – 2'-PDE (also PDE12) - was discovered to disrupt 2'-5' polyadenylates produced by OAS and its suppression resulted in reduction of vaccinia virus replication [422]. Another 2'-5'-PDE was discovered to specifically degrade 2'3'-cGAMP in human cell lines and murine liver and spleen [423]. Ecto-nucleotide pyrophosphatase/phosphodiesterase (ENPP1) is a plasma membrane and ER lumen protein that coordinates Ca^{2+} and Zn^{2+} ions and effectively digests 2'3'-cGAMP to nucleoside-monophosphates, but not 3'3'-cGAMP or 3'3'-c-diGMP of bacterial origin (Figure 13) [423-425]. Other ENPP1 substrates include ATP and diadenosine tetraphosphate (AP4A), whereas UTP and cAMP were hydrolyzed with less efficiency [425]. Tissues of ENPP1-knockout mice failed to degrade 2'3'-cGAMP efficiently [423]. ENPP1 can also be secreted into extracellular space, and in agreement to that, bovine and human sera

were shown to possess 2'3'-cGAMP hydrolase activity. In the same study 2'3'-cGAMP analog with phosphothioate diester linkages were synthesized and shown to have higher resistance to ENPP1 cleavage. Moreover, such 2'3'-cGAMP analogs were capable to inhibit 2'3'-cGAMP hydrolysis acting as a competitive inhibitor [423]. However, ENPP1 is a transmembrane protein, so it is not clear how it degrades 2'3'-cGAMP with a catalytic site located either in the extracellular space or inside the ER lumen. Interestingly, 2'3'-cGAMP levels were shown to be much higher in STING/Trex1-double knockout mice in comparison to a single Trex1-knockout suggesting that STING might facilitate 2'3'-cGAMP clearance and eventually participate in 2'3'-cGAMP transport into ER lumen [426, 427].

5.3. Cellular co-factors of cGAS-STING pathway

Though the presence of cGAS, STING and effector proteins TBK1 and IRF3 are the key players of type I IFN response to the cytosolic DNA, other cellular co-factors were found to contribute to cGAS/STING pathway (Figure 14). Ribosomal protein S6 kinase 1 (S6K1) was shown to interact with STING and to promote IFN signaling upon adenovirus infection [428]. S6K1 is an effector serine-threonine kinase downstream of mTOR that is involved in regulating cell growth, apoptosis, as well as transcription and translation in response to various cell stresses. Depletion of S6K1 in bone marrow–derived myeloid dendritic cells (BMDCs) resulted in a decreased IRF3 phosphorylation in response to adenovirus, HSV-1 or transfected DNA or cGAMP, but not to RNA virus VSV. S6K1 was shown to be required for the binding of IRF3 to the activated STING-TBK1 complex and to form a complex S6K1-STING-TBK1 in the presence of cytosolic DNA. Such tripartite complex was important for the early-phase expression of IRF3 target genes and induction of T cell responses in mice [428]. ZDHHC1 was found to be another co-factor of virus-triggered immune response through the cGAS/STING axis [429]. Zn-finger aspartate-histidine-histidine-cysteine (DHHC) domain-containing protein 1 (ZDHHC1) is an ER-associated protein and a member of the palmitoyl acytransferase family. ZDHHC1 knockdown abolished activation of IRF3 and NF-κB in response to HSV-1, but not SeV infection in mouse embryonic fibroblasts and in mice led to higher susceptibility to HSV-1 infection. Fluorescent microscopy revealed that ZDHHC1 co-localizes with STING independently of infection. ZDHHC1 was also discovered to support cGAMP-induced activation of STING^{H232} variant, otherwise incapable to stimulate IFN-β production. Surprisingly, though STING palmitoylation was shown to be necessary for STING activation, ZDHHC1 function in STING activation was described to be independent of its catalytic activity as palmitoyl transferase [429]. However, overexpression and transfection with high amounts of DNA might have resulted in too high level of IFN-β production and hidden any differences in ZDHHC1 mutants' activities.

Polyglutamine binding protein 1 (PQBP1) and NLRX1 were recently shown to regulate cGAS-STING pathway activation during HIV-1 infection (Figure 14). HIV-1 induces IFN innate immune response by cGAS-mediated recognition of HIV-1 reverse transcripts [220]. However, human dendritic cells (DCs) are resistant to HIV-1 infection because of SAM domain and HD domain containing protein 1 (SAMHD1) that depletes cellular deoxynucleoside triphosphate pools. Co-infection of HIV-1 with lentiviral protein VpX targeting SAMHD1 for degradation results in susceptibility of DCs to HIV-1 and elevated levels of type I IFNs [430-432]. In such system PQBP1 was defined as a co-factor of HIV-1 induced IFN signaling, since PQBP1 depletion resulted in reduced IFN response to HIV-1 and other retroviruses in monocyte-derived DCs and monocyte-like cell line THP-1 [219]. PQBP1 was discovered to associate with HIV-1 reverse-transcribed DNA product independently of cGAS and attracts cGAS to

HIV-1 DNA by direct interaction between proteins. PQBP1-deficient cells produced less cGAMP in response to HIV-1 infection showing that PQBP1 positively regulates cGAS activity in presence of retroviral products of reverse transcription. Interestingly, PQBP1 did not co-immunoprecipitate with cytosolic dsDNA from dsDNA viruses, making PQBP1 a unique sensor of retroviruses stimulating cGAS activation [219]. NLRX1, on the other side, was found to be a negative regulator of IFN signaling during HIV-1 or DNA virus infection [433]. NLRX1 is a member of NLRs usually associated with inflammasome formation. Unlike other NLRs, NLRX1 resides in mitochondria and induces ROS production and attenuates RIG-I/MAVS and TRAF/IKK signaling [434-437]. Additionally, NLRX1 downregulates cGAS/STING pathway by disrupting STING interaction with TBK1 [433]. NLRX1 was shown to directly interact with STING and to decrease STING-TBK1 interaction in a dose-dependent manner. Consistent to this, NLRX1 deficiency enhances immune responses to HSV-1 and HIV-1 viruses, as well as dsDNA and CDNs. Intriguingly, NLRX1 was demonstrated to support nuclear import of HIV-1 DNA promoting HIV-1 infection. However, STING-knockout cells showed the same HIV-1 infection rate independently of NLRX1 leading to conclusion that enhancement of HIV-1 infection by NLRX1 depends strongly on its function as STING inhibitor [433].

IFN-γ-inducible protein 16 (IFI16), initially proposed as DNA-sensor upstream from STING, was also reported to participate in HIV-1 immune recognition (Figure 14). HIV infection in primary human macrophages induces IFN response in IFI16- and STING-dependent manner and hence does not result in productive infection [237]. Permissive CD4+ T cells, on the other hand, are defective in DNA-signaling machinery, and despite IFI16 binding cytosolic DNA and STING recruiting TBK1 and IRF3 no IFN-induced genes are expressed leading to viral propagation [417]. Another role of IFI16 includes restriction of human papillomavirus and human cytomegalovirus replication by epigenetic modifications of viral promoters [438, 439]. Histone H2B was proposed to be a co-factor for IFI16-mediated recognition of HSV-1 and KSHV genomes in *de novo* infections leading to IFN production [440]. IFI16 was often linked to cGAS/STING pathway. IFI16 acetylation was discovered to drive IFI16 translocation into the cytosol where it engages STING for IFN-β production [237, 301]. On the other hand, nuclear localized cGAS was proposed to interact and stabilize the nuclear fraction of IFI16 [304]. Sensing of DNA from *L. monocytogenes* was shown to be dependent on cGAS and STING, as well as on IFI16 in human myeloid cells [222]. Recent data indicate that IFI16 positively regulates cGAS/STING axis by recruiting TBK1 to STING (Figure 14) [306, 307]. IFI16 knockout significantly reduced IFN response to transfected dsDNA and DNA viruses in human keratinocytes and macrophages. IFI16 forms a complex with TBK1 independently of infection and in presence of cytosolic dsDNA or 2'3'-cGAMP recruits TBK1 to STING to promote STING phosphorylation and activation. Though IFI16 involvement in STING activation seems to be robust, its influence on cGAS activity remains controversial. Some data suggest that IFI16 supports 2'3'-cGAMP production by cGAS and both domains of IFI16 are involved in this process [307]. However, another study failed to detect any decrease in cGAS-mediated 2'3'-cGAMP production in the absence of IFI16, though DNA-mediated interaction between IFI16 and cGAS was observed in co-immunoprecipitation studies, leaving the question whether IFI16 is a cGAS co-factor open [306].

Among recently discovered factors regulating cGAS/STING pathway are small GTPase RAB2B and its effector protein Golgi-associated RAB2B interactor-like 5 (GARIL5) (Figure 14). Upon DNA stimulation
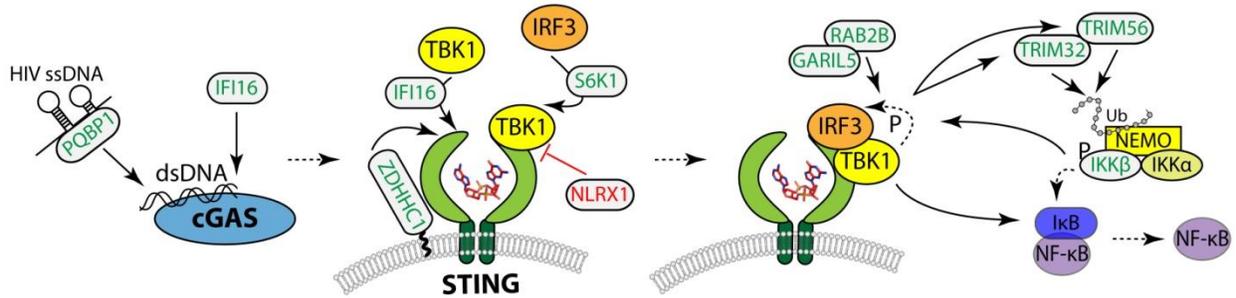
Figure 14 Cellular co-factors of cGAS/STING pathway.
cGAS/STING co-factors are indicated as grey ovals with green text for positive and with red text for negative regulation of the pathway. Black arrows stand for activation, red lines with blunt ends – for inhibition. Components within the pathway are connected with black dashed lines. Abbreviations used: ubiquitin (Ub) and phosphate (P).

RAB2B binds GTP and co-localizes with STING on the Golgi apparatus to recruit GARIL5. RAB2B/GARIL5 complex promotes IRF3 phosphorylation on STING and knockout of each protein leads to decreased IFN-β expression in presence of cytosolic dsDNA and cGAMP [441]. Another protein was found to link TBK1 and IKKβ into a positive feedback loop during cGAS/STING activation. Apart from IRF3, NF-κB is also activated by STING via an unknown mechanism. NF-κB is activated by the IKK complex consisting of kinases IKKα and IKKβ and regulatory subunit NF-κB essential modulator (NEMO). cGAS/STING activation was found to stimulate ubiquitin-ligases TRIM32 and TRIM56 to produce polyubiquitin chains that are bound by NEMO resulting in IKKβ activation. IKKβ was found to be essential for cytosolic DNA sensing and for TBK1, IRF3 and STING phosphorylation since its catalytically inactive mutant failed to restore phosphorylation upon STING activation. Interestingly, TBK1 was analogically required for NF-κB activation in response to cytosolic DNA, coupling IRF3 and NF-κB signaling during cGAS/STING activation (Figure 14) [442].

5.4. Cross-talk between cGAS/STING, autophagy, inflammasomes and apoptotic caspases

cGAS/STING DNA-sensing pathway is tightly linked to other innate immune pathways in cells. Crosstalk between DNA and RNA-sensing was first demonstrated by the discovery of RNA-polymerase III-driven synthesis of dsRNA from AT-rich dsDNA fragments. Resulting dsRNA activates the RIG-I pathway in parallel to dsDNA priming cGAS/STING activation [443]. HIV-1 with ssRNA genome produces a number of intermediates that are recognized by cGAS/STING axis, whereas STING in involved in RIG-I signaling upon RNA virus infection by forming RIG-I/MAVS/STING complex [220, 244, 245, 444]. Conversely, MAVS knockout in HeLa and HepG2 cells show reduced IFN response to cytosolic dsDNA [445].

Autophagy is an ancient mechanism counteracting infection and is induced upon bacterial infection under nutrient starvation [446]. Autophagy is integrated in nearly all innate immune pathways and is modulated by the majority of PAMPs and pathogen signaling pathways including cytosolic nucleic acid sensing pathways. For example, when MAVS signaling is triggered by viral invasion, autophagy protein ATG16L1 is recruited to the mitochondria by NLRX1, mediates mitophagy and eliminates the adaptor MAVS thus inhibiting cytosolic RNA sensing via RIG-I and MDA5 [447, 448]. Activation of DNA-

sensing cGAS/STING pathway can also initiate autophagy (Figure 15). *M. tuberculosis* infection was shown to be targeted by the components of ubiquitin-mediated autophagy in response to cGAS/STING activation resulting in delivery of bacteria to autophagosomes for degradation [449]. Activation of STING by DNA viruses HSV-1 and HCMV was also reported to elicit autophagy [450, 451]. Some autophagy-associated proteins like kinase ULK1 negatively regulating STING were described in the previous chapter. Moreover, upon dsDNA stimulation STING co-localizes with autophagy proteins microtubule-associated protein 1 light chain 3 (LC3) and autophagy-related gene 9a (Atg9a) [251]. Depletion of ATG9a results in increased STING trafficking to perinuclear region and enhances IFN production. Thus, ATG9a limits an excessive STING activation (Figure 15) [251]. Another autophagy protein – Beclin-1 – was found to be involved in negative regulation of cGAS. Beclin-1 is essential for autophagy induction and autophagosome maturation by forming complexes with its positive (ATG14 and UV irradiation resistance-associated gene, UVRAG) or negative (Bcl-2 and Rubicon) modulators [452-455]. Upon dsDNA stimulation cGAS was shown to bind Beclin-1 thus releasing it from the complex with Rubicon leading to Beaclin-1 mediated phosphatidylinositol 3-kinase class III (PI3KC3) activation and induction of autophagy for degradation of pathogen cytosolic DNA. Beclin-1 additionally suppresses cGAS enzymatic activity resulting in negative regulation of IFN signaling [456]. Negative regulation of IFN production by autophagy can be further illustrated by the observation that its inhibition in intestinal tissue and tumors leads to enhanced type I IFN signaling [457-459]. Autophagic elimination of DNA is also essential for preventing pathological STING activation: the leakage of chromatin DNA usually degraded by DNAse2a or autophagy results in elevated IFN levels and autoimmunity [460].

Crosstalk between cGAS/STING pathway and inflammasome activation, as well as activation of inflammatory and apoptotic caspases, also plays a crucial role in fine-tuning of cellular responses to PAMPs and DAMPs (Figure 15). Several inflammasome components were described to negatively influence innate immune responses like NLRX1 interfering with RIG-I/MAVS and TRAF6/NF-κB pathways, NLRP12 repressing NF-κB signaling and tumorigenesis in colon in mice, NLRP6 inhibiting TLR-induced NF-κB response to Gram-positive and -negative bacteria or NLRC5 interacting with IKKα and IKKβ and blocking NF-κB-dependent responses [101, 436, 461-463]. Some NLRs were reported to inhibit cGAS/STING activation through different mechanisms. A negative regulation of cGAS/STING by NLR protein NLRX1 was already described above. Another inflammasome component - NLRC3 - was found to directly bind STING and similar to NLRX1 to impede STING interaction with TBK1 [464]. NLRC3 associates with both ER-bound STING and TBK1 reducing their direct interaction and impairs STING trafficking to the perinuclear region attenuating type I IFN production in response to cytosolic DNA, cGAMP and DNA viruses *in vivo* and in the cells [464]. In addition, NLRC3 was reported to interact with TLR signaling adaptor TRAF6 blocking its ubiquitination and activation of NF-κB upon LPS treatment [465]. NLRP4, on the other hand, counteracts type I IFN production by targeting TBK1 for degradation [466]. NLRP4 was found to reduce IFN response to cytosolic dsDNA or dsRNA recruiting DTX4 to activated TBK1 after viral infection that results in TBK1 K63-linked ubiquitination and degradation. Importantly, NLRP4-mediated suppression of IFN signaling is specific towards TBK1-IRF3 type I IFN production and does not influence MyD88-IRF7 pathway triggered by CpG DNA [466]. Similarly to NLRP4, NLRP14 was discovered to target TBK1 for ubiquitination and degradation inhibiting RIG-I and cGAS RNA and DNA sensing pathways [467]. Another mechanism to inhibit IFN response was proposed for DNA-sensing AIM2 inflammasome. AIM2 activation was shown to counteract cGAS/STING pathway by promoting caspase-1-dependent cell death and AIM2 deficiency resulted in

elevated level of cGAMP production, STING activation and IFN-β synthesis in murine macrophages and dendritic cells [468]. In line with the assumption for AIM-2 inflammasome, a recent study suggested a general mechanism of cGAS/STING inhibition by canonical and non-canonical inflammasome activation. Caspase-1 or caspases 4, 5 and 11 activated upon DNA virus infection or by a non-canonical inflammasome formation, respectively, were found to cleave N-terminal domain of cGAS dramatically impairing its activity [469].

Unlike the most NLRs, IFI16 seems to have dual functions in regulating cGAS/STING and inflammasome pathways (Figure 15). IFI16 ability to positively regulate cGAS/STING pathway and facilitate type I IFN production was described in the previous chapter. On the other hand, IFI16 having the same domain organization as inflammasome-related AIM2 DNA sensor was proposed to assemble into inflammasome in response to Kaposi's sarcoma herpesvirus or Epstein-Barr virus infection and activate caspase-1 [67, 301]. IFI16 was also described to regulate inflammasome responses controlling Asc expression and to induce pyroptotic cell death in T cells and thus abolishing HIV replication [470, 471].

Interferon and inflammasomes cooperate during infection (Figure 15). Expression of inflammasome components and inflammatory caspases were found to be dependent on IFN-β [472-476]. The cGAS/STING pathway was reported to prime AIM2 inflammasome formation in human and mouse antigen-presenting cells in response to poxvirus [477]. STING was found to prime inflammasome formation upon infection with *Chlamydia trachomatis* which activated inflammasome cell death through its metabolic products [478]. On the other hand, IFN-β can induce IL-1R antagonist expression to attenuate IL-1 signaling and suppress IL-1β release from macrophages by inducing expression of anti-inflammatory cytokine IL-10 or by stimulating synthesis of inducible nitric oxide synthase which increases concentration of endogenous NO that inhibits NLRP3 oligomerization [479-482]. The latest research revealed that cGAS/STING activation can not only prime inflammasome formation, but also trigger NLRP3 inflammasome activation (Figure 15). AIM2 was found to be dispensable for DNA-induced inflammasome activation and IL-1β release in human myeloid cells. cGAS/STING activation was discovered to result in STING trafficking to lysosome and induction of lysosomal cell death. Activated STING was shown to trigger membrane permeabilization which caused K^+ efflux and NLRP3 activation [29]. According to another study 2'3'-cGAMP was found to prime and activate AIM2 and NLRP3 inflammasomes in mouse bone marrow–derived macrophages (BMDMs) and human primary macrophages and dendritic cells. Interestingly, 2'3'-cGAMP-induced inflammasomes were composed of both AIM2 and NLRP3 and led to STING-dependent IL-1β secretion, but not cell death [483]. Thus, cGAS/STING activation induces inflammasome formation differentially depending on organism and cell type. Taken together, inflammasome and cGAS/STING axis present a dynamic response to invading pathogens. Initial activation of IFN producing pathways leads to priming and activation of inflammasomes that in turn switch off IFN production and drive IL-1β and IL-18 secretion and pyroptotic cell death.

Apart from inflammatory caspases, apoptotic caspases were found to control IFN response. Defective enzymatic activity of caspase 8 was found to cause elevated levels of IRF3 and TBK1 activation and chronic inflammatory skin disorder in mice (Figure 15) [484]. Caspase 8 was found to negatively regulate RIG-I and TLR activation by cleaving polyubiquitinated RIPK1 thus inhibiting type I IFN production [485, 486]. Apoptotic caspases 3, 7 and 9 activated by intrinsic apoptotic pathway were found to inhibit
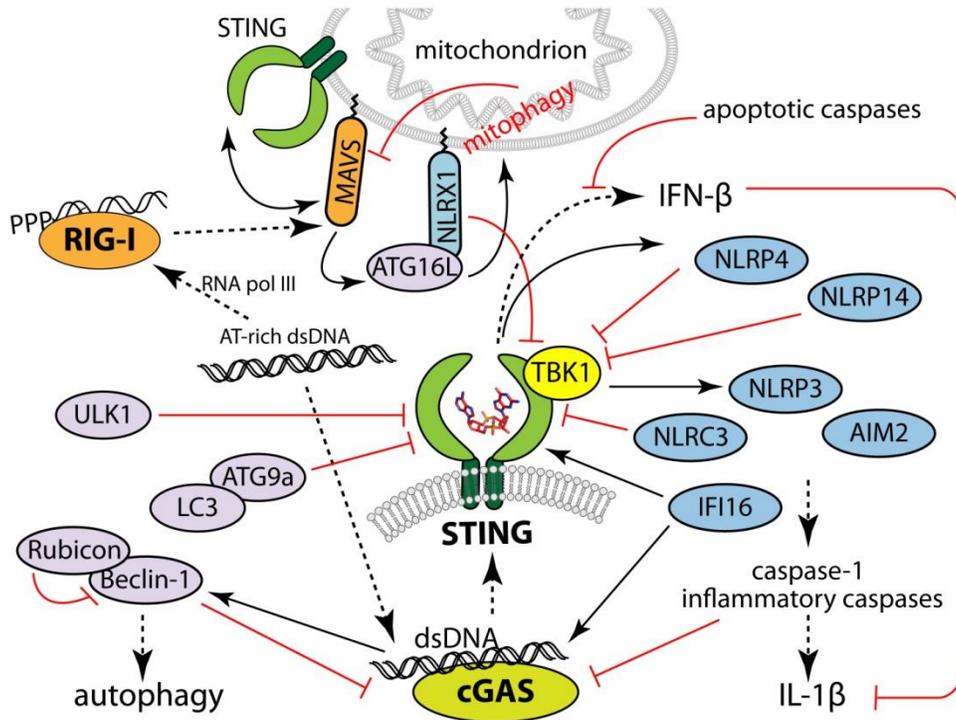
Figure 15 A cross-talk between cGAS/STING, autophagy and inflammasome pathways.
Components of cGAS/STING pathway (green) are interconnected with proteins involved in autophagy (purple),
RIG-I signaling (orange) and inflammasome formation (blue). Components within one pathway are connected with
dashed black arrows. Activation is indicated with black arrows, inhibition – with red lines with blunt ends.

cGAS/STING activation and type I IFN production during apoptosis. Knockout or pharmacological
inhibition of these caspases together with activation of apoptosis leads to Bak/Bax-mediated release of
mitochondrial DNA, cGAS/STING activation and IFN synthesis. Knockout of apoptotic initiator caspase
9 or double knockout of effector caspases 3/7 is sufficient for elevated type I IFN levels, and this effect is
further increased by treatment with pro-apoptotic drug ABT-737 [226, 227]. Since some viruses express
caspase inhibitor proteins to prevent apoptosis of the host cell, such mechanism of IFN response might
have evolved as counteracting viral defenses. However, if active apoptotic caspases target mitochondrial
DNA, cGAS/STING or downstream components of the pathway, remains to be uncovered.

5.5. Viral inhibition of cytosolic DNA sensing

During co-evolution bacteria and viruses evolved several mechanisms counteracting host defenses, some
of them inhibiting cGAS and STING (Figure 16). Group B *Streptococcus* expresses ectonuclease CdnP
which degrades extracellular c-di-AMP and limits STING activation. Though CdnP limits host IFN
response, it does not digest 2'3'-cGAMP and fails to abolish bacterial DNA sensing via cGAS [487].
Phosphodiesterase CdnP from *Mycobacterium tuberculosis*, however, is capable of hydrolysis of not only
bacterial derived c-di-AMP, but cGAS-generated 2'3'-cGAMP as well effectively inhibiting *M.
tuberculosis* detection by both cGAS and STING [488]. *Shigella flexneri* utilizes IpaJ and virA proteins to
inhibit STING transport from ER to Golgi and deletion of both proteins results in a robust IFN response
to *Shigella* infection [489].

STING is also a target for viral defensive programs (Figure 16). vIRF1 from KSHV was discovered to bind STING and prevent its association with TBK1 [217]. Similarly, ICP27 protein from HSV-1 interacts with activated STING and TBK1 preventing phosphorylation of IRF3 [490]. Oncogenes E7 and E1A from the DNA tumor viruses human papillomavirus (HPV) and adenovirus (AdV), respectively, were described to bind STING and abolish its activity [491]. Surprisingly, according to another study adenovirus E1A did not suppress TBK1/IRF3 activation through cGAS/STING in the early stage of infection and replication of adenovirus was not affected by the lack of cGAS or STING [492]. Such results can be explained with E1A blocking antiviral gene expression downstream of IFN-β by inhibition STAT1 activation [493]. Polymerase of Hepatitis B virus (HBV) was found to associate with STING and prevent its K63-linked ubiquitination that is necessary for STING activation [494].

The importance of STING for RNA viruses is supported by their STING-inhibiting strategies. HCV, usually recognized by RIG-I signaling pathway, encodes NS4B protein that inhibits STING either by intersecting its interaction with TBK1 or by disrupting STING-MAVS interactions [495, 496]. Dengue virus (DENV) encodes protease complex NS2B/3 that cleaves STING between its transmembrane helices inhibiting IFN response. Interestingly, NS2B/3 is not capable for murine STING degradation explaining the resistance of murine cells to DENV infection [497-499]. Human and porcine coronaviruses, as well as severe acute respiratory syndrome coronavirus (SARS-CoV) possess a membrane-anchored papaine-like protease (PLP) domain containing protein that co-localizes with STING, disrupts its interaction with TBK1 and repress K63-linked polyubiquitination of STING resulting in decreased IFN-β production [500-503]. However, whether or not these PLPs need their deubiquitinating activity remains debatable.

cGAS activity is also directly targeted by some pathogens (Figure 16). KSHV tegument protein ORF52 and its homologs in other gammaherpesviruses were discovered to bind both cGAS and DNA and to inhibit cGAS enzymatic activity [504]. Another protein from KSHV - latency-associated nuclear antigen (LANA) – was also reported to inactivate cGAS. N-terminally truncated cytosolic LANA was found to interact with cGAS and abolish 2'3'-cGAMP production promoting reactivation of KSHV from its latent state, otherwise suppressed by cGAS/STING IFN signaling [505]. LANA was also shown to recruit members of the MRN complex in the cytosol and to inhibit their function in cytosolic DNA sensing. Such inhibition results in decrease of NF-κB activation and KSHV progression into lytic replication cycle [275]. Furthermore, ORF64 tegument protein from murine gammaherpesvirus MHV68 was found to carry a deubiquitination activity which is necessary to antagonize cGAS/STING pathway and to facilitate MHV68 latent infection [506].

Another way to avoid immune recognition is to hide their genomes and replication products within a capsid. For example, reverse transcription of HIV-1 is covered from cytosol by viral capsid and only cDNA is later injected directly into nucleus thus avoiding activation of cGAS and STING. HIV-1 capsid regulates recruitment of the host factors cleavage and polyadenylation factor 6 (CPSF6) and the components of the nuclear pore complex cyclophilins (Nup358 and CypA) that participate in translocation of HIV reverse transcripts into the nucleus [507, 508]. However, depletion or impaired interaction of HIV capsid with these factors, or loss of other host factors like TREX1 and SAMHD results in detection of HIV by cGAS/STING pathway [430-432, 509, 510]. HIV capsid interacts with another host factor - prolyl-isomerase cyclophilin A (CypA) – that counteracts the masking of HIV cDNA and induces IFN response in dendritic cells. Mutations in capsid proteins causing an enhanced binding of CypA lead to a potent immune response to HIV-1 in human dendritic cells thus driving HIV-1 capsid
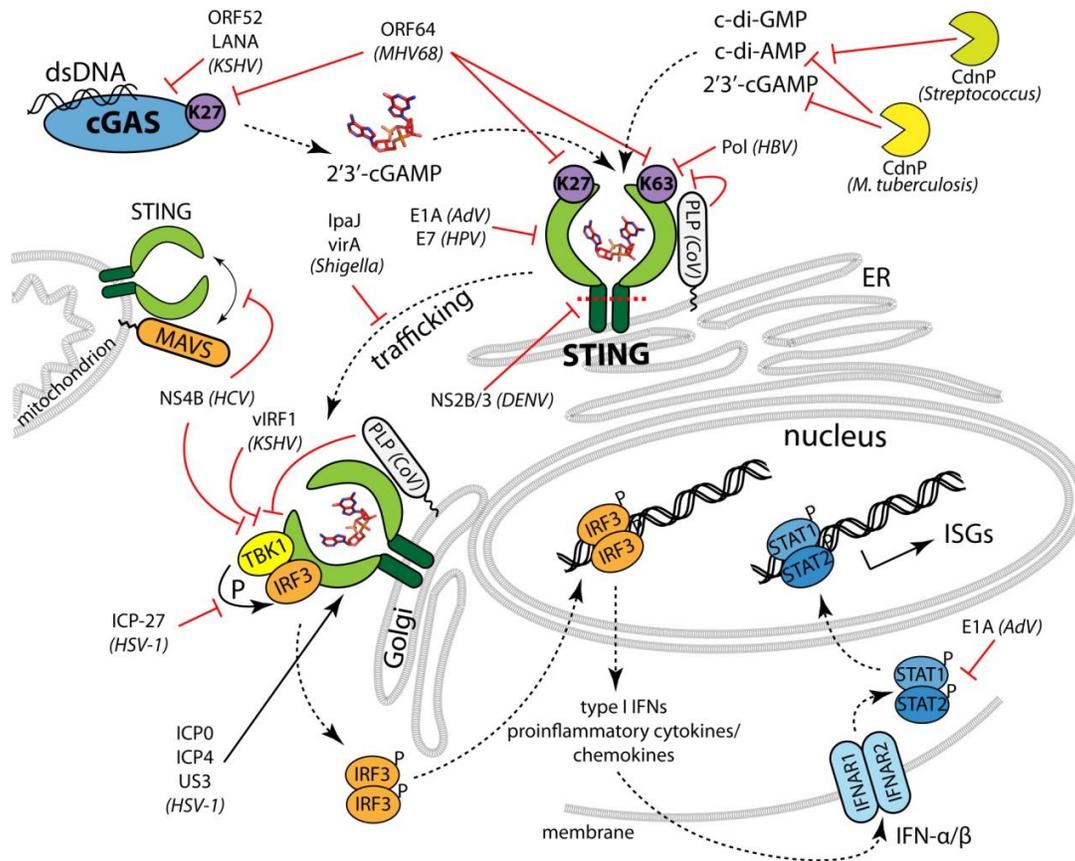
Figure 16 Inhibition of cGAS/STING pathway by pathogens.
Components of the pathway are connected with black dashed lines. Bacterial and viral proteins are indicated by abbreviations with an organism name given in brackets. Inhibition is depicted with red lines with blunt ends, activation with black arrows. Polyubiquitination is marked with purple circle (text inside indicates a type of linkage between ubiquitin monomers). "P" stands for phosphorylation.

evolution to escape interaction with CypA [511, 512].

Surprisingly, some viruses and bacteria profit from type I IFN production. In cancer-derived HEp-2 or HeLa cells STING activation was shown to be necessary for HSV-1 replication. HSV-1 proteins ICP0, ICP4 and protein kinase US3 stabilized STING in these cell lines and their deletion resulted in STING degradation and inhibited HSV-1 replication. Conversely, in human embryonic lung cells and HEK293T derived from normal tissues STING activation repressed HSV-1 replication, suggesting that HSV-1 interaction with the host are dependent on the cell type [513]. Unexpectedly, mice deficient in IRF3 were discovered to be resistant to *M. tuberculosis* infection. Activation of cGAS/STING in macrophages by cytosolic DNA emerging from *M. tuberculosis* was found to promote long-term infection and explains type I IFN signature in tuberculosis [221]. Similarly, *L. monocytogenes* requires STING and IFN-β expression in response to its cytosolic DNA and c-di-AMP for downregulation of cell-mediated immunity and STING knockout mice exhibited enhanced immunity against infection [514]. Type I IFN is beneficial for systemic infection with *L. monocytogenes* since the IFN receptor is required for cell-to-cell spread of bacteria in macrophages and in mouse liver [515].

5.6. cGAS/STING pathway in self-DNA recognition and disease

In addition to pathogen DNA and CDN recognition, cGAS/STING pathway also participates in type I IFN induction in response to self-DNA (Figure 17). Though cellular DNA is strictly compartmentalized and under normal circumstances is not present in cytosol, under certain conditions it can leak into the cytoplasmic compartment leading to cGAS/STING activation. Recent studies revealed that DNA emerging from mitochondrial nucleoids (mtDNA) is a potent activator of IFN response (Figure 17) [225]. Mitochondrial stress resulting from depletion of mtDNA-binding protein TFAM (transcription factor A, mitochondrial) or infection with herpesviruses HSV-1, HSV-2 or MHV-68 leads to hyperfused mitochondrial phenotype and, apparently, mitochondrial membrane rupture. As a result of such mitochondrial stress mtDNA leaks into the cytosol where it engages cGAS causing enhanced type I IFN response (Figure 17). Reduction of mtDNA copy number by dideoxycytidine (ddC) attenuated ISG expression [225]. Furthermore, inhibition of apoptosis was found to lead to mtDNA leakage into cytosol through Bax/Bak protein pores in mitochondrial membrane and to induce type I IFN production by activating cGAS/STING axis [226, 227]. Deficiency or pharmacological inhibition of apoptotic caspase-9, apoptotic protease- activating factor 1 (Apaf-1) or caspases-3/7 alone or together with activation of intrinsic apoptotic pathway by targeting Bax/Bak inhibitors Bcl-2 and Bcl-x_L in murine hematopoietic stem cells and mouse embryonic fibroblasts results in elevated IFN response (Figure 17). However, whereas importance of caspase catalytic activity for inhibiting cGAS signaling in apoptosis is clarified, the questions why activated caspase 9 in caspase-3/7 knockout cells is not sufficient for IFN inhibition and what is the exact mechanism of immunological silencing of apoptosis remain open. Additionally, mitochondrial damage and consequent mtDNA leakage was previously reported to induce NLRP3 activation and IL-1β release that according to the recent research can be linked to cGAS/STING activation [29, 516, 517]. Elevated mtDNA levels in serum are a hallmark of diabetes and mtDNA was shown to induce NLRP3 activation and increased expression of IL-1β in diabetic mice [518, 519]. Mitochondrial polymorphisms and dysfunction were also linked to the autoimmune disease systemic lupus erythematosus (SLE) that has a characteristic IFN signature. However, whether cGAS/STING recognition of mtDNA is responsible for elevated type I IFN levels in SLE remains unclear [520-523]. Involvement of mtDNA in SLE pathogenesis is further supported by the presence of anti-mtDNA antibodies in blood of SLE patients [524]. Moreover, neutrophils from SLE patients were found to accumulate oxidized mtDNA that upon extrusion activates type I IFN expression by plasmacytoid dendritic cells [525].

Another source of self-DNA emerges from nucleus (Figure 17). Chromatin DNA was recently found to form cytosolic micronuclei and activate the cGAS/STING pathway [228, 229]. Mitotic progression despite the presence of double-stranded DNA breaks caused by ionizing radiation was shown to result in chromosomal missegregation and formation of micronuclei – chromatin fragments surrounded by their own nuclear membrane. Micronuclei membrane having defective lamina organization undergoes rupture resulting in accessibility of DNA fragments for cytosol. cGAS was shown to co-localize with such structures and to trigger type I IFN production. Excessive micronuclei formation was also observed in cells deficient of RNase H2 responsible for ribonucleotide excision repair and genome stability [228, 526, 527]. Importantly, cGAS/STING-dependent elevated levels of IFN-stimulated genes and proinflammatory response were characteristic to mouse embryonic fibroblasts (MEFs) deficient in functional RNase H2 [528, 529]. Moreover, malfunction of RNase H2 is associated with the autoinflammatory disorders Aicardi-Goutières syndrome (AGS) and SLE bridging micronuclei formation to cGAS activation and

autoinflammation [529-532]. Activation of cGAS/STING by micronuclei formation plays an important role in cancer therapy. Indeed, depletion of STING in irradiated melanoma cells results in larger tumor volumes and smaller survival rates in mice [229].

Incomplete DNA damage repair caused by environmental stresses constantly results in generation of ssDNAs emerging from aberrant replication intermediates or endogenous retroelements that leak to the cytosol (Figure 17) [261, 533]. Normally these DNA fragments are cleaved by exonuclease TREX1 anchored to the nuclear membrane and TREX1 deficiency results in DNA accumulation in cytosol and cGAS/STING activation followed by type I IFN production [426, 534, 535]. Chronic cytokine and IFN production is characteristic for autoinflammatory diseases and mutations in exonuclease TREX1 are associated with AGS, SLE, familial chilblain lupus and retinal vasculopathy with cerebral leukodystrophy [536-540]. Remarkably, oxidative stress and UV irradiation result in enhanced levels of oxidized self-DNA that is resistant to TREX1-mediated degradation and therefore a very potent activator of cGAS/STING. DNA oxidation was found to be responsible for sun-induced skin lesion in SLE patients [224]. Another nuclease - DNase I – that is secreted into extracellular space to eliminate DNA fragments emerging from apoptotic and necrotic cells has also an important role in controlling autoimmunity [541, 542]. DNase I activity was decreased in SLE patients and DNase I deficiency in mice led to higher levels of antibodies against DNA [263, 543]. SLE patients were found to have mutations in DNaseI, and a polymorphism of this gene was associated with SLE susceptibility, though it does not seem to influence DNase I activity [544-546]. Unlike DNase I, DNase II is an endonuclease degrading lysosomal DNA fragments in macrophages. DNase II deficiency results in type I IFN overexpression caused by cGAS/STING pathway activation and development of polyartritis and anemia in mice [264, 547-549]. Genetic studies identified mutations in sterile α motif and histidine-aspartic acid (HD) domain–containing protein 1 (SAMHD1), adenosine deaminase acting on RNA 1 (ADAR1) and dsRNA sensor MDA5 to be associated with AGS [532]. SAMHD1 was first described as an antiviral GTP-dependent deoxynucleotide triphosphohydrolase that reduces dNTP levels below the concentration sufficient for HIV-1 replication [550, 551]. SAMHD1 was reported to maintain genome stability and its depletion results in impaired proliferation, DNA damage and increased IFN levels [552]. Moreover, bone marrow-derived macrophages (BMDMs) lacking SAMHD1 showed spontaneous IFN response that was dependent on cGAS/STING activation providing one more link between cGAS/STING axis and AGS [553]. Other nucleic acid sensing pathways can also be involved in autoimmunity and AGS. ADAR1 catalyzing deamination of adenosine in dsRNA destabilizes RNA secondary structures and prevents recognition of endogenous dsRNA derived from 5'-untranslated regions, introns and rRNA by MDA5 [554-556]. Thus, mutations in both ADAR1 and MDA5 can be linked to increased type I IFN response through MDA5/MAVS pathway and autoimmunity [557].

Gain of function mutations in human STING were found to correlate with inflammatory diseases and hyperproduction of cytokines and IFNs. V147L, N154S, V155M, G166E, and R284M point mutations in STING located close to membrane-embedded part of the protein were described to result in its autoactivated state presumably due to uncontrolled dimerization of STING, its traffic from ER and enhanced ability to bind TBK1 [489, 558-561].

Interferons play a crucial role in cancer surveillance [562]. They exhibit cytotoxic effects on cancer cells directly and promote maturation of dendritic cells (DCs) linking innate and adaptive immunity [563, 564]. Spontaneous T-cell priming against tumors is dependent on STING and its deficiency results in
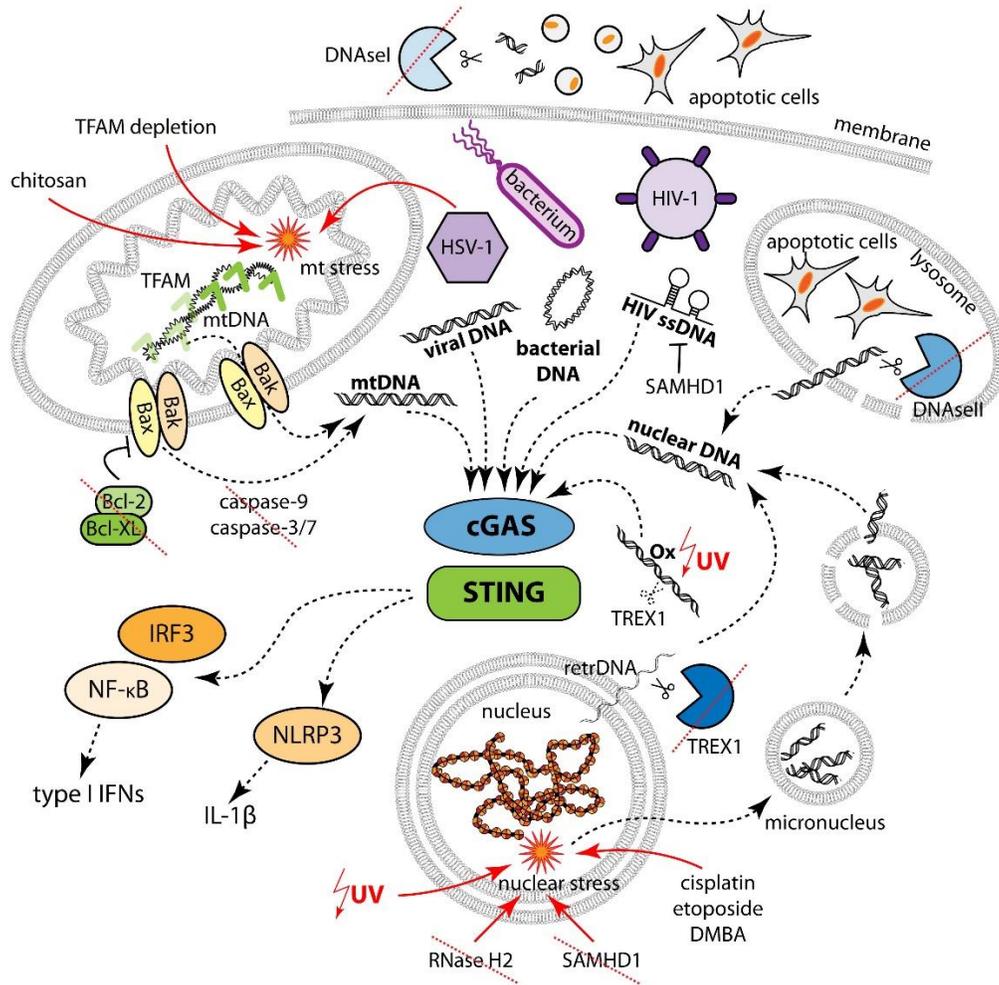
Figure 17 Main sources of DNA recognized by cGAS/STING pathway.
Components of one pathway are connected with black dashed lines. Inhibition is shown as black lines with blunt ends. Induction of mitochondrial (mt) and nuclear stresses (orange stars) is shown with red arrows. Depletion or dysfunction of the proteins is depicted as red dashed line. DNA digestion is shown with scissors.

higher tumor growth in mice and a decreased response to CD47-blocking antibody, counteracting CD47-mediated inhibition of phagocytosis [565-567]. Dying tumor cells are phagocytized by DCs where their DNA is detected by cGAS leading to type I IFN production [568]. Type I IFNs, in turn, facilitate cross-presentation and recruit effector T-cells to tumor microenvironment [566, 569]. Cleavage of genomic DNA in prostate cancer cells was found to lead to accumulation of cytosolic DNA and cGAS/STING activation promoting phagocytosis of tumor cells and T cell responses [570]. cGAS/STING activation plays important role in radio- and chemotherapy as well. DNA-damage induced by irradiation or chemotherapeutics like cisplatin and etoposide results in DNA leakage into cytosol and cGAS/STING-mediated IFN production [228, 229, 566, 571]. Treatment with chitosan - cationic polysaccharide - was shown to cause increased ROS production and mtDNA leakage into cytosol in dendritic cells and trigger type I IFN production via cGAS/STING pathway resulting in antigen-specific T helper 1 (Th1) activation [572]. Importance of cGAS/STING pathway in anti-tumor immunity is further illustrated by cGAMP functions as immune adjuvant. Efficiency of immune checkpoint therapy targeting PD-1/PD-L1 interaction that prevents tumor cells from T-cell mediated lysis was strongly improved by simultaneous

activation of STING pathway by cGAMP or its analogs [573-575]. Intratumoral injection of cGAMP or its analogs was shown to enhance antitumor immunity and to inhibit tumor growth in mouse models of colon, brain, skin, pancreatic, breast and B cell cancers revealing a new field of anti-cancer therapeutics [566, 575-580]. Interestingly, STING is not activated in animals with acute myeloid leukemia resulting in the absence of type I IFN response and poor survival rate, however, administration of STING agonist DMXAA induced expression of IFN-β and other inflammatory cytokines and resulted in leukemia-specific T cell activation and higher survival rates [566, 581].

Given its major role in DCs activation and T cells priming, cGAS/STING pathway is often suppressed in cancer cells. Colon adenocarcinoma and melanoma cells have impaired cGAS/STING signaling by hypermethylation of cGAS and STING promoter regions, decreased cGAS and STING protein expression and impaired STING trafficking to avoid immune surveillance [582, 583]. Furthermore, oncogenes from DNA tumor viruses like E1A from human papilloma virus are potent inhibitors blocking STING activation [491]. Several tumors were shown to accumulate extrachromosomal telomere repeats (ECTRs) that are used as a template for telomere length extension and can activate cGAS/STING pathway when introduced into human fibroblasts [584-587]. Such tumors, including osteosarcoma, neuroblastoma and adenocarcinoma, do not express STING presumably due to its epigenetic silencing in order to avoid immune surveillance [587].

Surprisingly, in some cases cGAS/STING activation can contribute to carcinogenesis and tumor propagation. For example, treatment with the polyaromatic hydrocarbon 7,12-dimethylbenz[α]anthracene (DMBA) inducing leakage of nuclear DNA into cytosol promotes proinflammatory cytokines and skin carcinogenesis in STING-dependent manner [571, 588]. Chronic infection with *Helicobacter pylori* results in STING upregulation and IFN-β expression and correlates with gastric cancer development, since STING was found to be crucial in controlling viability, migration and invasion of gastric cancer cells [589]. STING-dependent type I IFN production in response to genotoxic agents was also associated with survival and growth of breast cancer resulting in its resistance to chemotherapy [590]. STING-mediated DNA sensing leading to IFN-α/β production was found to activate indoleamine-2,3-dioxygenase (IDO) that limits T cell function and suppress Th1 responses [591, 592]. This mechanism was found to promote the growth of Lewis lung carcinoma, but not B16 melanoma suggesting that DNA sensing through cGAS/STING can promote growth of tumors with low antigenicity [593]. Moreover, a transfer of cGAMP via gap junctions between tumor cells and astrocytes was found to stimulate growth and chemoresistance of brain metastatic cells [415]. Taken together, these findings indicate that though STING activators may become a potent anti-tumor therapeutics, discriminating between STING-mediated anti-tumor effects and tumor-promoting inflammation is necessary for a good clinical outcome.

6. Objectives

cGAS/STING pathway is a major mechanism for cytosolic DNA detection and is well characterized. A huge range of biochemical and structural studies on cGAS revealed its functions and detailed molecular mechanisms of DNA recognition and 2'3'-cGAMP synthesis. A strong conservation of the Mab-21 domain of cGAS on sequence and structural levels throughout evolution and structural similarity to bacterial CDN-synthases is remarkable. Despite of this, the Zn-thumb element and the DNA-sensing function of cGAS homologs seems to be a rather novel feature that developed in vertebrates.

The majority of data obtained about cGAS are based on the studies of its conserved Mab-21 domain-containing part comprising main DNA binding and catalytic functions. Though the mode of DNA binding and a mechanism of 2'3'-cGAMP synthesis are well understood, several questions about cGAS Mab-21 domain remain open. As such, according to the published structures it can accommodate 14-20 bp long dsDNA that introduces a "spine" helix break and active site rearrangement. According to this, the 20 bp DNA ligand should be in principle sufficient for cGAS activation since it spans the entire length of the catalytic domain, however, it fails to fully activate cGAS *in vivo* and can stimulate cGAS *in vitro* only in artificially high concentrations of all reaction components. A second question arises from the finding that cGAS has two DNA-binding sites and requires dimerization for its activity. In this case two DNA molecules held by two cGAS protomers are pointing to each other in such manner that prolonging these DNA fragments will result in steric clashes between DNA stands leading to an assumption that cGAS might preferentially bind DNA ends. This hypothesis is, however, contradicted by the fact that long DNA shows much higher potential for cGAS activation than the shorter fragments with higher number of DNA ends. The instability of such $cGAS_2:DNA_2$ complex with 20 bp DNA is another point that needs further clarification. Since such dimer was not found for OAS proteins and seems to be a unique feature of cGAS, it is puzzling why such peculiar and important for cGAS activity dimerization step is so sensitive and whether such instability was specifically selected during evolution.

Since recent studies discovered an involvement of cGAS in the recognition of self-DNA emerging from mitochondria and nucleus, as well as in HIV-1 detection, the search for specific cGAS ligands and their structures seems to be a very important direction of research. Moreover, since several cross-talks between cGAS activation and other cellular pathways and proteins such as Mre11 and IFI16 were proposed, a finding of cGAS co-factors is of an utmost importance and might shed light on the mechanism of cGAS activation *in vivo*.

The aim of this thesis was to investigate cGAS activation mechanism by long stimulating dsDNA ligands. For this purpose structural and biochemical *in vitro* studies, as well as cGAS stimulation in human cell lines were performed using DNA ligands of different lengths. In order to evaluate cGAS activity in a high-throughput manner a novel fluorescence-based cGAS activity assay was developed. A relevance of the obtained cGAS crystal structure with 39 bp stimulator DNA was validated using isothermal titration calorimetry (ITC) and right-angle light scattering. Moreover, cGAS mutagenesis studies *in vitro* and in cell lines were performed. Based on the crystal structure of cGAS active complex and the studies of cGAS enzymatic activity additional cGAS co-factors were proposed to facilitate cGAS function *in vivo*. To validate this hypothesis mutational analysis of these co-factors together with co-immunoprecipitation assays in human cell lines were performed. Further studies included three-dimensional structured

illumination and wide-field fluorescent microscopy to confirm co-localization of cGAS with its co-factors and DNA in cell lines.

Additionally, biologically relevant cGAS nucleic acid ligands were studied. As such, reverse-transcribed ssDNA resulting from HIV-1 infection was analyzed. For this purpose stimulating effect of modified HIV-1 hairpins and synthetic DNA constructs on cGAS was investigated by cell line experiments and further validated with *in vitro* cGAS activity assays and co-immunoprecipitation studies. Moreover, another potential cGAS ligand - RNA:DNA hybrids – was evaluated by stimulation of wild type and knock-out human cell lines. Cells were transfected with different synthetic constructs and cGAS activity was measured with luciferase reporter assays, IFN-β mRNA expression levels or ELISA. The results were confirmed in primary peripheral blood mononuclear cells and with *in vitro* cGAS activity assays.

Publications

1. cGAS senses long and HMGB/TFAM-bound U-turn DNA by forming protein–DNA ladders

Summary

This publication provides the first crystal structure of cGAS in complex with immunostimulatory 39 bp DNA, as well as a mechanism for a cooperative DNA recognition by cGAS. cGAS is a central sensor of cytosolic DNA, which upon DNA binding produces a second messenger 2'3'-cGAMP. cGAMP triggers type I IFN response through activation of adaptor protein STING. Previously known structures explained a molecular mechanism of cGAS dimerization, activation and 2'3'-cGAMP synthesis, however, they contain only short 12-20 bp DNA fragments that fail to activate cGAS *in vivo* and stimulate 2'3'-cGAMP production *in vitro* only under non-physiologically high DNA and protein concentrations. In this study we offer a link between a well-established cGAS-dimerization and a necessity of long dsDNA strands for a full cGAS activation. According to our hypothesis, cGAS forms a DNA-protein ladder consisting of a row of cGAS dimers as "rungs" arranged on two long DNA "ladder sides". Observed not only in crystal packing, but also in solution studies, such ladders introduce a cooperative DNA-binding of cGAS and explain its DNA length-dependent activation in *in vitro* and cell-based assays. We propose that the binding of the first cGAS dimer is rather unstable, however, if formed, it parallelizes two long DNA strands resulting in efficient binding of a subsequent cGAS dimer. cGAS dimers within such ladder are mutually stabilized, which explains a higher affinity and stronger cGAS activation by long DNA fragments. Moreover, previously published structures of DNA U-turns introduced by high-mobility group box (HMGB) proteins, such as mitochondrial transcription factor A (TFAM), HMGB-1 and bacterial nucleoid packaging HU proteins, share a remarkable similarity in DNA arrangement with discovered in this work cGAS-DNA ladders. Indeed, these DNA-bending proteins were shown to dramatically increase cGAS activity *in vitro* and competed with cGAS activation at high concentrations. Normally present in mitochondria and nucleus, TFAM and HMGB1 were found to relocalize to a cytosolic compartment under mitochondrial stress conditions or DNA transfection, respectively, giving a possibility for their engagement with cGAS *in vivo*. Furthermore, HMGB1 was shown to co-localize in cytosol with DNA and cGAS further supporting the hypothesis of its involvement in cGAS activation. Combined with cGAS-DNA protein ladder model these findings introduce a nucleation-oligomerization mechanism for cGAS activation similar to proposed for other innate immune sensors. Though nucleation of cGAS filament by the first cGAS dimer is rather inefficient, other co-factors like DNA-bending proteins might successfully prearrange DNA in nearly parallel manner and thus efficiently nucleate cGAS-DNA ladder

formation. This hypothesis explains instability of a single cGAS dimer that might have emerged to counteract an aberrant cGAS activation by short non-pathogenic DNA fragments and offers U-shaped structured DNA as a preferable cGAS ligand. In biological sense it provides a mechanism for specific recognition of partly dissociated mitochondrial and bacterial nucleoids, as well as nuclear DNA fragments and introduces HMGB proteins as a class of potent cellular cGAS co-factors.

Author contribution

The author of this thesis performed crystallographic and biochemical studies. She crystallized and collected diffraction data of cGAS in complex with 39 bp DNA. Together with Prof. Dr. K.-P. Hopfner she invented a novel fluorescence-based assay for cGAS activity measurement. The author of this thesis performed various *in vitro* cGAS activity assays for studying cGAS stimulation by DNA fragments of different lengths and cGAS co-factors. With help from C. Isakaj and O. Fettscher she cloned and purified cGAS point mutants and tested their activity *in vitro*. Assisted by D.J. Drexler she performed isothermal titration calorimetry (ITC) and determined stoichiometry of cGAS:DNA complexes and thermodynamic parameters of cGAS binding to various DNA constructs. She also optimized size exclusion chromatography coupled to right-angle light scattering (SEC-RALS) and successfully used it for composition analysis of several cGAS:DNA complexes. She supervised an advanced research practical course in Biochemistry, where T. Schaller and her designed, cloned and purified HMGB1 point mutants and studied their DNA binding efficiency with electromobility shift assays (EMSA) and their cGAS-activating capacity in different *in vitro* cGAS activity assays. Together with Prof. Dr. H. Leonhardt and Dr. C. Lässig she designed and interpreted localization studies of HMGB1 and cGAS with fluorescence microscopy. With contributions from all other co-authors she assisted Prof. Dr. K.-P. Hopfner in manuscript writing.

cGAS senses long and HMGB/TFAM-bound U-turn DNA by forming protein–DNA ladders

Liudmila Andreeva[1,2], Björn Hiller[1,2], Dirk Kostrewa[1,2], Charlotte Lässig[1,2], Carina C. de Oliveira Mann[1,2]†, David Jan Drexler[1,2], Andreas Maiser[3], Moritz Gaidt[1,2], Heinrich Leonhardt[3,4], Veit Hornung[1,2,4] & Karl-Peter Hopfner[1,2,4]

[1]Department Biochemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany

[2]Gene Center, Ludwig-Maximilians-Universität München, 81377 Munich, Germany.

[3]Department Biology, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany.

[4]Center for Integrated Protein Science Munich, 81377 Munich, Germany.

†Present address: Dana-Farber Cancer Institute, Harvard Medical School, MA 02115-5418 Boston, Massachusetts, USA.

Cytosolic DNA arising from intracellular pathogens triggers a powerful innate immune response[1,2]. It is sensed by cyclic GMP–AMP synthase (cGAS) which elicits the production of type I interferons (IFN) by generating the second messenger 2′3′-cyclic-GMP–AMP (cGAMP)[3-5]. Endogenous nuclear or mitochondrial DNA can also be sensed by cGAS under certain conditions, resulting in sterile inflammation. The cGAS dimer binds two DNA ligands shorter than 20 base pairs side-by-side[6-9], but 20-base-pair DNA fails to activate cGAS *in vivo* and is a poor activator *in vitro*. Here we show that cGAS is activated in a strongly DNA length-dependent manner both *in vitro* and in human cells. We also show that cGAS dimers form ladder-like networks with DNA, leading to cooperative sensing of DNA length: assembly of the pioneering cGAS dimer between two DNA molecules is ineffective; but, once formed, it prearranges the flanking DNA to promote binding of subsequent cGAS dimers. Remarkably, bacterial and mitochondrial nucleoid proteins HU and mitochondrial transcription factor A (TFAM), as well as high-mobility group box 1 protein (HMGB1), can strongly stimulate long DNA sensing by cGAS. U-turns and bends in DNA induced by these proteins pre-structure DNA to nucleate cGAS dimers. Our results suggest a nucleation-cooperativity-based mechanism for sensitive detection of mitochondrial DNA[10] and pathogen genomes[11], and identify HMGB/TFAM proteins as DNA-structuring host factors. They provide an explanation for the peculiar cGAS dimer structure and suggest that cGAS preferentially binds incomplete nucleoid-like structures or bent DNA.

cGAS detects a broad range of intracellular viral and bacterial pathogens[11-16] but also senses mitochondrial (mt)DNA in the cytosol upon mitochondrial stress[10]. cGAS dimers bind two DNA ligands side-by-side, whereby each cGAS monomer directly recognizes fewer than 20 base pairs (bp). In principle, 20 bp DNA should be sufficient to fully activate cGAS; however, it does not activate cGAS in human cells[17]. Longer DNA ligands were even proposed to lead to steric clashes unless cGAS binds near DNA ends[8]. Thus, the current model fails to explain the detection of immunostimulatory DNA. Keeping the overall amount of DNA constant, we stimulated a human monocyte cell line (BLaER1 cells) with DNA of different lengths and observed a striking concentration- and length-dependent activation of cGAS by measuring C-X-C motif chemokine 10 (CXCL10) production as an appropriate surrogate parameter of cGAS activity[18], as well as interferon-β (IFN-β) mRNA expression levels (Fig. 1a and Extended Data Fig. 1a-c). Long herring testis DNA robustly activated cGAS at all tested DNA amounts, while shorter cGAS ligands required increasing amounts of DNA. Consistent with previous studies, DNA of ~45 bp constituted a 'length' threshold, below which no activation was observed.

Since other cellular factors could contribute to the length-sensitive detection, we analysed the intrinsic capability of mouse (m)cGAS and human (h)cGAS to sense DNA length *in vitro*, using a new fluorescent-based assay (Extended Data Fig. 2 and Methods). Both hcGAS and hcGAS[cd] (catalytic domain) showed a remarkable DNA length-dependent activation, which appears to be independent of the cGAS N-terminal part (Fig. 1b). Like the ~45 bp threshold *in vivo*, robust activation at physiologically relevant cGAS concentrations *in vitro* required DNA >40 bp. Plasmid DNA was the most potent activator. Mouse mcGAS[cd] exhibited a comparable length-dependent activation, with a gradual increase in activity until about 75 bp (Fig. 1c and Extended Data Fig. 1d). At higher DNA concentrations and short DNA lengths, mcGAS was activated in an almost stepwise fashion, with a first plateau between 30 and 45 bp and a second above 50 bp (Fig. 1c inset), perhaps reflecting stepwise binding along DNA. DNA <20 bp can also activate cGAS but requires 10–20 times higher cGAS and 50–250 times higher DNA amounts to induce activities similar to 50 bp DNA (Extended Data Fig. 1e, f). Finally, circular and linearized plasmid DNA activated cGAS equally well, ruling out the possibility that cGAS needs DNA ends (Extended Data Fig. 1g).

To determine a mechanism for the DNA length sensing, we crystallized mcGAS[cd] in complex with 39 bp DNA. The crystals diffracted anisotropically to 3.6–4.8 Å resolution and we determined the structure by molecular replacement. In the crystal, two cGAS dimers and two

DNA 39-mers form a $cGAS_4–DNA_2$ complex (Fig. 1d and Extended Data Fig. 3a). Each of the two cGAS dimers in our $cGAS_4–DNA_2$ complex is similar to the dimers seen in $cGAS_2–DNA_2$ complexes determined with <20 bp DNA (Extended Data Fig. 3b, c), indicating that long DNA does not induce a substantially different structural state in cGAS than short DNA. The two cGAS dimers are arranged in a 'head-to-head' orientation along the DNA. DNA between the cGAS dimers is slightly curved (Fig. 1d), avoiding the proposed clashes and showing that cGAS does not need DNA ends (Extended Data Fig. 3c, d).

The asymmetric unit contains one full $cGAS_4–DNA_2$ complex along with one half $cGAS_4–DNA_2$ complex situated on a twofold crystallographic symmetry axis (Extended Data Fig. 4a). Remarkably, these $cGAS_4–DNA_2$ complexes are further stacked into fibrils with alternating 'head-to-head'- and 'tail-to-tail'-oriented cGAS dimers (Fig. 1e), forming a DNA–protein ladder with rungs (cGAS dimers) and two ladder sides (DNA). The DNA is continuous between 'head-to-head'-oriented cGAS dimers and quasi-continuous (stacked 3′ to 3′ and 5′ to 5′) between the 'tail-to-tail'-oriented cGAS dimers. A ~5 bp (half helical turn) increased spacing of 'tail-to-tail'-oriented cGAS dimers, however, would allow an energetically favourable continuous assembly of alternatingly oriented cGAS dimers between two long DNA elements (Extended Data Fig. 3e, f).

Neighbouring cGAS dimers along the DNA barely interact in either 'head-to-head' or 'tail-to-tail' orientations in the crystal lattice (Extended Data Fig. 4a). Consistently, mutational analysis *in vitro* and *in vivo* showed that while protein–DNA contacts are critical, protein–protein contacts between adjacent cGAS dimers in either orientation are not (Extended Data Fig. 4b-d)[8,9]. Accordingly, the DNA–protein ladder does not require direct interactions between neighbouring cGAS dimers to be stable, because cGAS–DNA contacts and cGAS–cGAS dimer contacts between the DNA already generate a mesh-like structure.

To see whether ladder-like structures exist in solution, we used isothermal titration calorimetry (ITC) and size-exclusion chromatography coupled to right-angle light scattering (SEC–RALS). ITC showed a DNA length-dependent increase in affinity and at the same time an increase in cGAS:DNA molar ratio (Fig. 2a and Extended Data Fig. 5). DNA of ~45 bp confered a threshold, resulting in higher affinity and at the same time binding of two cGAS molecules per DNA ligand. The affinity increase was driven by a decrease in binding enthalpy that also compensated a decreasing entropy (Fig. 2b and Extended Data Fig. 6a, b). The latter suggests the

formation of ordered structures on DNA >45 bp. Together with ITC, SEC–RALS showed that cGAS forms single homogeneous complexes with DNA of different length: a $cGAS_1$–DNA_1 complex for 20-mer DNA, a $cGAS_4$–DNA_2 complex for 50-mer DNA, and $cGAS_6$–DNA_2 complexes for 70/80-mer DNA (Fig. 2c-f and Extended Data Fig. 6c). Of note, $cGAS_2$–DNA_2 complexes were instable in the presence of 20 bp DNA and only the $cGAS_1$–DNA_1 complex robustly formed with 20 bp DNA that poorly activates cGAS (Fig. 1a-c). Given that activation of cGAS requires additional dimerization[8], this explains why 20 bp DNA is not a good cGAS activator.

The DNA–protein ladders offer an intriguing mechanism for the DNA length-dependent activation of cGAS. Formation of a pioneering $cGAS_2$–DNA_2 complex is highly unfavourable, but once formed, it prearranges the two DNA molecules in its vicinity to promote formation of adjacent cGAS dimers with increased affinity (Fig. 3a). Multiple cGAS dimers in $cGAS_{2n}$–DNA_2 ($n \geq 2$) ladders stabilize each other by cooperatively holding together the two DNA ladder sides. An analytical equation for this 'DNA–protein ladder' model can be derived (see Supplementary Methods) and comprises a V_{max} (maximal reaction velocity); three DNA constants K_1, K_2^2 and K_3 that describe interaction of monomeric cGAS with DNA (K_1), interaction of cGAS dimers with two DNA ligands (K_2^2), and the association/dissociation of cGAS–DNA and $cGAS_2$–DNA_2 states (K_3); and a parameter s that is a measure of cooperative binding sites of cGAS dimers along DNA (equivalent to a Hill coefficient).

To test this model, we measured cGAS activity by varying both $mcGAS^{cd}$ concentration and DNA length. We observed a dramatic length-dependent activation that was not due to an increase in V_{max}, consistent with our model but was due to a DNA length-dependent increase in binding affinity, as seen with ITC (Fig. 3b). The data can be globally fitted with a simplified equation derived from the mathematical model (Supplementary Methods equation 3.6 and sections 4.1 and 4.2) and can be explained by a model in which V_{max} and the DNA binding constants do not depend on the length of DNA, while s increases as a function of DNA length (Fig. 3c). Parameter s measures cooperativity between adjacent cGAS dimers owing to the DNA arranging activity[19]. The transitions of s around 40–50 bp and $2s \approx 2.8$ for plasmid DNA suggest that length sensing can be explained by cooperative binding of two or more adjacent cGAS dimers. Furthermore, titrating catalytically inactive $mcGAS^{cd}$(D307N) to low amounts of active $mcGAS^{cd}$ markedly increased cGAS activity about 200-fold, providing direct evidence for cooperative sequestering of cGAS into oligomeric structures on DNA (Fig. 3d). Higher amounts

of the inactive mutant gradually competed mcGAS[cd] away from DNA, as expected. These data can be fitted well ($R^2 = 0.96$) with the 'DNA–protein ladder' equation (Supplementary Methods) and result in a cooperativity parameter $2s = 3.6$, similar to that described in Fig. 3c. In summary, cGAS length sensing emerges from cooperative formation of $cGAS_{2n}–DNA_2$ ($n \geq 2$) complexes.

Cellular factors could enhance detection of DNA by cGAS if they suitably structure the DNA. We noticed that the HMGB proteins HMGB1/2 and mitochondrial nucleoid organizing protein TFAM could be well suited to nucleate cGAS dimers because they properly prearrange DNA by forming U-turns[20,21] (Extended Data Fig. 7a, b) and because mtDNA activates cGAS[10] and HMGB1/2 facilitate cytosolic nucleic-acid sensing[22]. Bacterial HU proteins, for example from *Listeria monocytogenes* (lHU), also bend DNA to form bacterial nucleoids. Indeed, adding increasing amounts of mTFAM, mHMGB1, and lHU robustly activated mcGAS[cd] *in vitro* up to ~25-fold (Fig. 4a-c and Extended Data Fig. 7c). A similar activation was also seen for hTFAM and full-length hcGAS (Fig. 4d). cGAS activation required the DNA-bending ability of HMGB1, since mutations that decreased DNA-bending capacity without significantly affecting DNA binding[23] reduced or nearly abolished cGAS activation (Extended Data Fig. 8a-d). TFAM activates cGAS both on circular and on linear plasmid DNA (Extended Data Fig. 8e), works without direct TFAM–cGAS interactions (Extended Data Fig. 8f), and even does not require proteins from the same species (Extended Data Fig. 8g, h). Thus, these proteins presumably activate cGAS not by protein–protein interactions but rather by prestructuring DNA to nucleate or stabilize cGAS dimers. At higher concentrations, TFAM and lHU sharply abolished cGAS activity, presumably because of cooperative formation of nucleoid-like structures[24]. HMGB1 inhibited cGAS more gradually, probably because it does not cooperatively form compact nucleoids. Strong stimulation of cGAS activity by TFAM required DNA >100 bp (Extended Data Fig. 7d), suggesting that robust activation still involves $cGAS_{2n}–DNA_2$ complexes ($n \geq 2$).

TFAM normally resides in mitochondria; however, under mitochondrial stress conditions, induced by a combination of caspase and Bcl-2 inhibitors[25,26], we saw increased TFAM presence in the cytosol (Extended Data Fig. 9a-d) where it can assist detection of leaked mtDNA. Consistently, cytosolic hcGAS and hTFAM co-immunoprecipitated (Extended Data Fig. 9e). Importantly, we found that upon DNA transfection of mouse embryonic fibroblasts (MEFs) endogenous HMGB1 co-localized with cytosolic DNA and cGAS in 85% of the observed cytosolic DNA loci (Fig. 4e and Extended Data Fig. 10). This is consistent with previous findings that HMGB1 is involved in cytosolic DNA sensing[22] even though it normally

resides in the nucleus. However, HMGB1 shuttles between nucleus and cytoplasm[27], enabling it to encounter cytosolic DNA where it could assist cGAS.

In summary, we provide a molecular mechanism for the sensitive detection of long DNA by cGAS (Fig. 4f). $cGAS_1$–DNA_1 complexes appear to be inactive and $cGAS_2$–DNA_2 complexes unstable, requiring DNA of sufficient length to form stable $cGAS_{2n}$–DNA_2 ladders ($n \geq 2$). Nucleic-acid-stress HMGB proteins and nucleoid-structuring proteins (TFAM, HU) can additionally nucleate and stabilize cGAS-ladders by prearranging DNA. Thus, cGAS preferentially senses structured DNA ligands and DNA with residual nucleoid proteins bound, rather than naked DNA. In terms of recognizing danger or pathogen-associated molecular patterns in the form of mtDNA and bacterial nucleoids, these findings make biological sense and provide a plausible rationale for the evolution of the peculiar cGAS dimer structure. TFAM/HMGB1/HU-enhanced stimulation of cGAS could be especially helpful for the initial detection of long cytosolic DNA with low amounts of cGAS present. In this context, it should be noted that short ~20-mer DNA can also strongly activate cGAS if it additionally contains G-rich ssDNA Y overhangs[17]. In principle, this flanking G-rich DNA could also stabilize cGAS dimers, but the precise mechanism, as well as that of other postulated host proteins[28,29], remains to be uncovered. In any case, the nucleation- and cooperativity-based mechanism imposes a threshold-like response that conceptually unifies DNA sensing by the cGAS–STING axis with other oligomerization-based nucleic-acid-sensing pathways[30].

REFERENCES

1. Rodero, M. P. & Crow, Y. J. Type I interferon-mediated monogenic autoinflammation: the type I interferonopathies, a conceptual overview. *J. Exp. Med.* **213**, 2527–2538 (2016).

2. Barber, G. N. STING: infection, inflammation and cancer. *Nat. Rev. Immunol.* **15**, 760–770 (2015).

3. Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786–791 (2013).

4.   Zhang, X. et al. Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol. Cell* **51**, 226–235 (2013).

5.   Chen, Q., Sun, L. & Chen, Z. J. Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing. *Nat. Immunol.* **17**, 1142–1149 (2016).

6.   Gao, P. et al. Cyclic [G(2□,5□)pA(3□,5□)p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. *Cell* **153**, 1094–1107 (2013).

7.   Civril, F. et al. Structural mechanism of cytosolic DNA sensing by cGAS. *Nature* **498**, 332–337 (2013).

8.   Li, X. et al. Cyclic GMP-AMP synthase is activated by double-stranded DNA-induced oligomerization. *Immunity* **39**, 1019–1031 (2013).

9.   Zhang, X. et al. The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. *Cell Reports* **6**, 421–430 (2014).

10.   West, A. P. et al. Mitochondrial DNA stress primes the antiviral innate immune response. *Nature* **520**, 553–557 (2015).

11.   Hansen, K. et al. Listeria monocytogenes induces IFN□ expression through an IFI16-, cGAS- and STING-dependent pathway. *EMBO J.* **33**, 1654–1666 (2014).

12.   Gao, D. et al. Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science* **341**, 903–906 (2013).

13.   Reinert, L. S. et al. Sensing of HSV-1 by the cGAS–STING pathway in microglia orchestrates antiviral defence in the CNS. *Nat. Commun.* **7**, 13348 (2016).

14.   Collins, A. C. et al. Cyclic GMP-AMP synthase is an innate immune DNA sensor for Mycobacterium tuberculosis. *Cell Host Microbe* **17**, 820–828 (2015).

15.   Wassermann, R. et al. Mycobacterium tuberculosis differentially activates cGAS- and inflammasome-dependent intracellular immune responses through ESX-1. *Cell Host Microbe* **17**, 799–810 (2015).

16.   Watson, R. O. et al. The cytosolic sensor cGAS detects Mycobacterium tuberculosis DNA to induce type I interferons and activate autophagy. *Cell Host Microbe* **17**, 811–819 (2015).

17.   Herzner, A. M. et al. Sequence-specific activation of the DNA sensor cGAS by Y-form DNA structures as found in primary HIV-1 cDNA. *Nat. Immunol.* **16**, 1025–1033 (2015).

18.   Holm, C. K. et al. Virus-cell fusion as a trigger of innate immunity dependent on the adaptor STING. *Nat. Immunol*. **13**, 737–743 (2012).

19.   Weiss, J. N. The Hill equation revisited: uses and misuses. *FASEB J*. **11**, 835–841 (1997).

20.   Ngo, H. B., Kaiser, J. T. & Chan, D. C. The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nat. Struct. Mol. Biol*. **18**, 1290–1296 (2011).

21.   Rubio-Cosials, A. et al. Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat. Struct. Mol. Biol*. **18**, 1281–1289 (2011).

22.   Yanai, H. et al. HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses. *Nature* **462**, 99–103 (2009).

23.   Allain, F. H. T. et al. Solution structure of the HMG protein NHP6A and its interaction with DNA reveals the structural determinants for non-sequence-specific binding. *EMBO J*. **18**, 2563–2579 (1999).

24.   Ngo, H. B., Lovely, G. A., Phillips, R. & Chan, D. C. Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. *Nat. Commun*. **5**, 3077 (2014).

25.   Rongvaux, A. et al. Apoptotic caspases prevent the induction of type I interferons by mitochondrial DNA. *Cell* **159**, 1563–1577 (2014).

26.   White, M. J. et al. Apoptotic caspases suppress mtDNA-induced STING-mediated type I IFN production. *Cell* **159**, 1549–1562 (2014).

27.   Bonaldi, T. et al. Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. *EMBO J*. **22**, 5551–5560 (2003).

28.   Yoh, S. M. et al. PQBP1 is a proximal sensor of the cGAS-dependent innate response to HIV-1. *Cell* **161**, 1293–1305 (2015).

29.   Jønsson, K. L. et al. IFI16 is required for DNA sensing in human macrophages by promoting production and function of cGAMP. *Nat. Commun*. **8**, 14391 (2017).

30.   Sohn, J. & Hur, S. Filament assemblies in foreign nucleic acid sensors. *Curr. Opin. Struct. Biol.* **37**, 134–144 (2016).

Supplementary Information is available in the online version of the paper.

Author Contributions L.A. performed crystallographic and biochemical studies. B.H. performed enzyme-linked immunosorbent assay (ELISA) assays and IFN-β mRNA expression analysis. D.K. built and refined the structure. C.L. performed co-immunopurification studies. C.C.O.M. established staining protocols for three-dimensional structured illumination microscopy (3D SIM). D.J.D. performed luciferase reporter assays and analysed cGAS products. A.M. performed microscopy. M.G. generated cGAS-deficient BLaER1 cells. H.L., C.C.O.M., C.L., and L.A. designed and interpreted microscopy experiments. V.H., B.H., and C.L. designed and interpreted cell-based experiments. K.-P.H. designed the study, derived the mathematical model, and analysed data. K.-P.H. and L.A. wrote the paper with contributions from all other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to K.-P.H. (hopfner@genzentrum.lmu.de).

Reviewer Information *Nature* thanks O. Nureki and the other anonymous reviewer(s) for their contribution to the peer review of this work.

FIGURES

Figure 1 cGAS shows DNA length-dependent activity and forms DNA–protein ladders. a, cGAS activation in trans-differentiated BLaER1 cells. Cells were stimulated with 20, 40, and 60 ng DNA of different length (20–100 bp in 5 bp intervals) and herring testis DNA. First two bars each represent unstimulated cells and lipofectamine controls. Shown are mean values ± s.d., $n = 3$. **b**, *In vitro* activity of hcGAS, hcGAS^{cd}, and (**c**) mcGAS^{cd} in the presence of DNA of increasing length (20–100 bp in 5 bp intervals) or plasmid DNA (last bar). Mean values of initial cGAS reaction rates ($\Delta F/\Delta t$, relative fluorescence units (RFU) min^{-1}), measured by the rate of fluorescent ATP analogue (fATP) incorporation into fluorescent cGAS product (fGAMP, see Extended Data Fig. 2) are plotted against DNA length ± s.d., $n = 3–5$. **d**, Crystal structure of cGAS in complex with 39 bp DNA. Two cGAS dimers (green and orange) assemble on DNA in

'head-to-head' orientation. Zinc-thumb dimerization elements and the nucleotidyl-transferase

(NT) sites are indicated. **e**, Ladder-like cGAS assembly along a quasi-continuous DNA in the

crystal lattice through alternating 'head-to-head' or 'tail-to-tail' orientations of cGAS dimers.
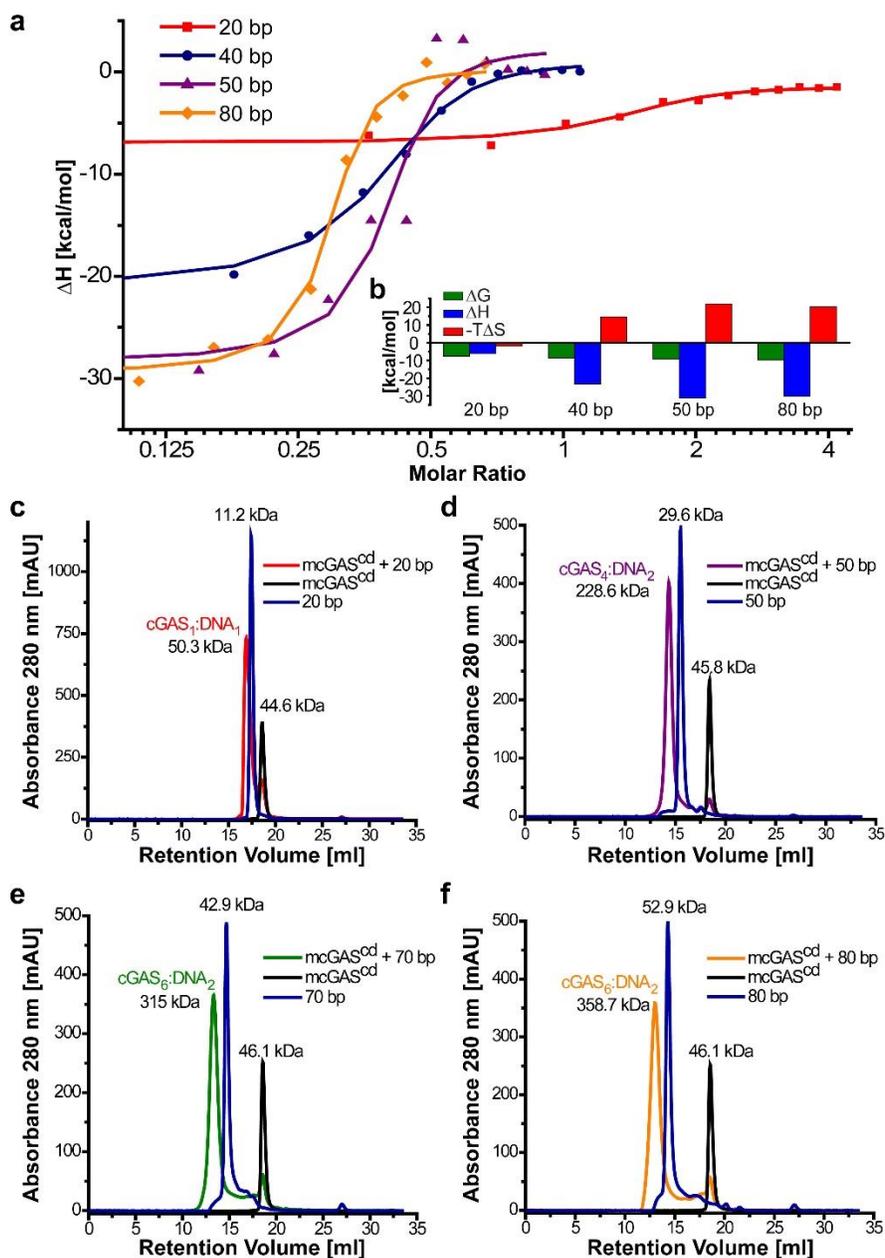
Figure 2 cGAS and DNA assemble into cGAS$_{2n}$–DNA$_2$ complexes in solution. a, ITC of

mcGAS^{cd} and DNA. Enthalpy (ΔH) values are plotted against the molar ratio of DNA:mcGAS^{cd}.

Obtained dissociation constant (K_d) values for 20, 40, 50, and 80 bp DNA are

$2.4 \times 10^{-6} \pm 1.89 \times 10^{-6}$, $507 \times 10^{-9} \pm 128 \times 10^{-9}$, $165 \times 10^{-9} \pm 166 \times 10^{-9}$, and

$73.3 \times 10^{-9} \pm 25.5 \times 10^{-9}$ M, respectively. Molar ratios are 1.28 ± 0.125, 0.364 ± 0.01, 0.371 ± 0.027, and 0.274 ± 0.006, respectively. **b**, Thermodynamic parameters obtained with ITC. **c**–**f**, SEC–RALS analysis of mcGAS^{cd} with 20 bp (**c**), 50 bp (**d**), 70 bp (**e**), and 80 bp (**f**) DNA; mAU, milli absorption units. Estimated molecular masses are indicated. Molecular masses 12.2, 30.8, 43.1, and 49.3 kDa were used as reference for 20, 50, 70, and 80 bp DNA, respectively; mcGAS^{cd} was estimated with 43 kDa (see also Extended Data Fig. 6c). Some molecular mass deviation from reference values occurred because of limitations of measurement accuracy.
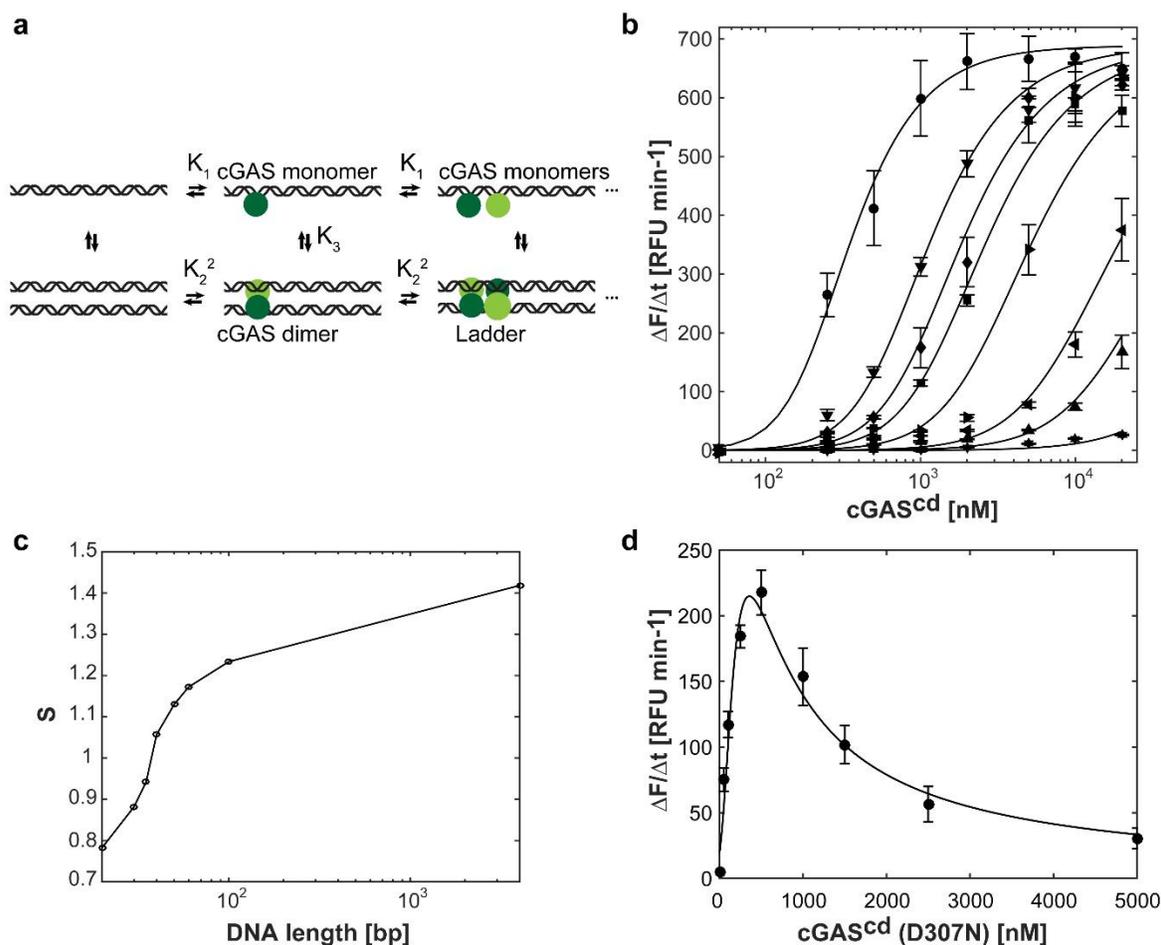
Figure 3 DNA–protein ladder model. a, Kinetic scheme for the DNA–protein ladder model. cGAS binds as monomer to DNA (dissociation constant K_1). These $cGAS_1$–DNA_1 complexes can assemble into $cGAS_2$–DNA_2 structures with parallelized DNA (K_3). Finally, two cGAS molecules can directly bind as dimer to the parallelized DNA ligands (K_2^2). Overall, this scheme describes a cooperative, DNA length-dependent interaction of cGAS with DNA (see Supplementary Methods). **b**, Activity of mcGAS^{cd} on different DNA ligands as a function of protein concentration (symbols: circle, plasmid; downward arrowhead, 100 bp; diamond, 60 bp; square, 50 bp; rightward arrowhead, 40 bp; leftward arrowhead, 35 bp; upward arrowhead, 30 bp; star, 20 bp). The overall concentration of base pairs in each reaction is kept constant.

Mean values of cGAS activity (see Fig. 1c legend) are plotted against mcGAS^{cd} concentrations $\pm$ s.d., $n = 3$. Superposed is a global least square minimization of the data with an equation describing the DNA–protein ladder model ($R^2 = 0.988$) (see Supplementary Methods). **c**, Plot of the cooperativity parameter s obtained from **b** as a function of DNA length. **d**, Titration of catalytically inactive mcGAS^{cd} (D307N) into a sub-active solution of mcGAS^{cd}. Mean values of cGAS activity (see Fig. 1c legend) are plotted against mcGAS^{cd} (D307N) concentrations $\pm$ s.d., $n = 3$. Superposed is a least square minimization of the data with the DNA–protein ladder model ($R^2 = 0.96$) (see Supplementary Methods).

Figure 4 DNA-bending proteins enhance cGAS activity by prearranging DNA into U-shaped structures. a–d, cGAS activities (see Fig. 1c legend) in the presence of different DNA-bending proteins. The last value corresponds to a control without DNA. Mean values of initial cGAS reaction rates ($\Delta F/\Delta t$) are plotted against increasing concentrations of DNA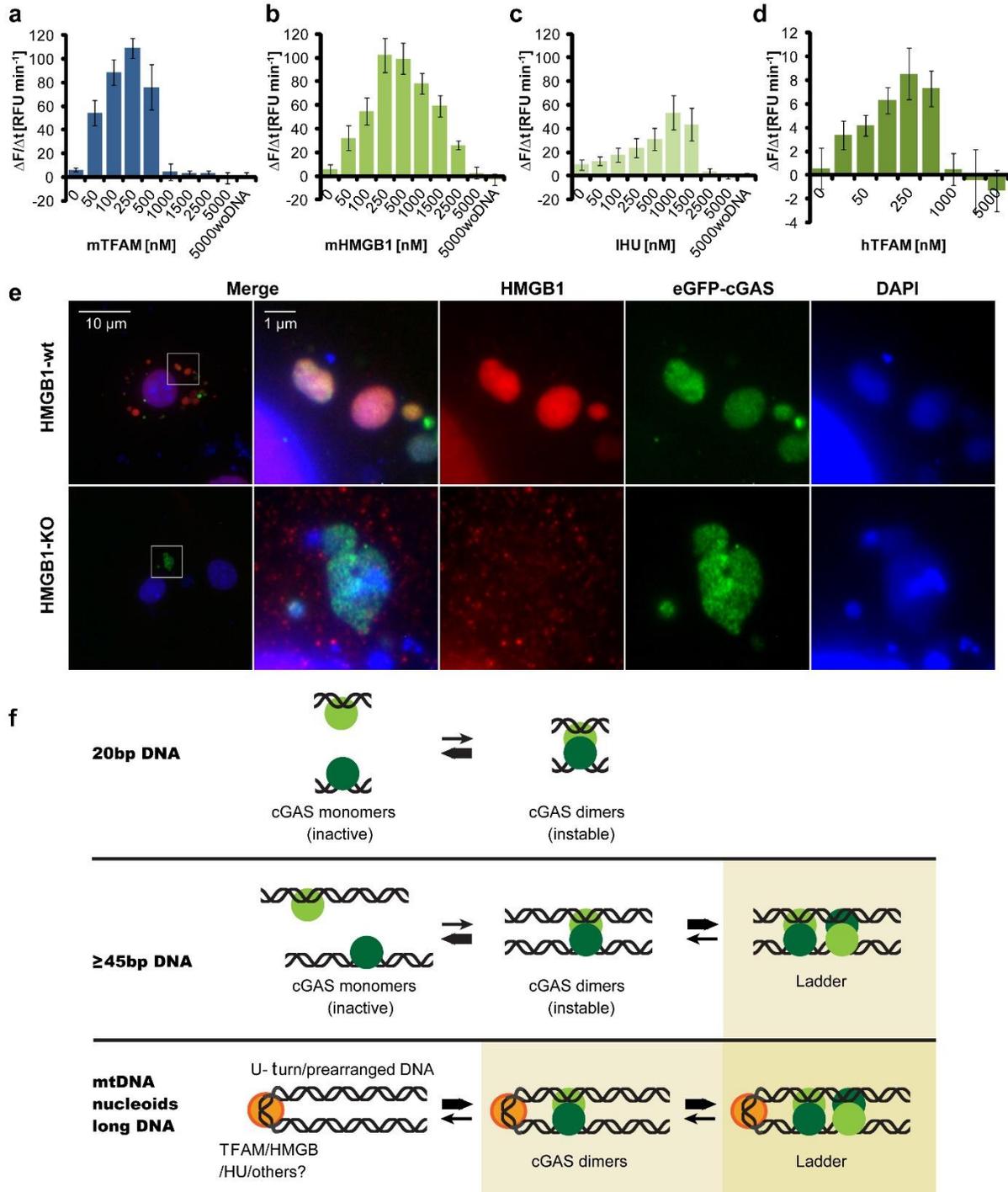-bending proteins ± s.d., $n = 3$–6. **a**, mTFAM, (**b**) mHMGB1, and (**c**) lHU robustly enhance mcGAS[cd] activity in a dose-dependent manner until they eventually compete cGAS away. **d**, hTFAM activates hcGAS. **e**, HMGB1 (red), cGAS (green), and cytosolic DNA (blue) co-localization. Wild-type (WT) or HMGB1-knockout (KO) MEFs were transfected with an enhanced green fluorescent protein (eGFP)–cGAS-expressing vector and stained for wide-field fluorescence microscopy (pDV) (compare with Extended Data Fig. 10). DNA was stained with 4,6-diamidino-2-phenylindole (DAPI, blue). Fluorescent signal within enlarged images is enhanced until saturation of the nuclear signal to better visualize cytoplasmic structures. **f**, Model for DNA length- and structure-dependent cGAS activation. Activation of cGAS requires the formation of $cGAS_2$–DNA_2 dimers, which are very unstable. Multiple cGAS dimers along DNA ≥ 45 bp stabilize each other by cooperatively holding DNA together, leading to stable, active ladders (shaded panels). DNA-bending/U-turn-inducing proteins prearrange DNA, nucleating the formation of cGAS dimers and ladders (darker shaded panel).

METHODS

Constructs and cloning

The plasmids encoding full-length (amino acids (aa) 1–522) and truncated (catalytic domain, 'cd') *Homo sapiens* (h) (aa 155–522) and *Mus musculus* (m) (aa 141–507) cGAS for N-terminal His_6–MBP (maltose-binding protein) fusion protein expression were described before[7].

The sequence encoding mTFAM without a mitochondrial localization signal (aa 43–243) optimized for *Escherichia coli* expression was synthesized by Eurofins Genomics and cloned for N-terminal His_6 fusion protein expression into modified pET28a vector (Novagen), where the thrombin cleavage site was exchanged with a tobacco etch virus (TEV) protease cleavage site.

hTFAM (aa 43–246) sequence was obtained from total cDNA and inserted into pET28a for His_6 fusion protein expression in *E. coli*. Full-length mHMGB1 (aa 1–215), as well as *L. monocytogenes* HU (lHU) (aa 1–121) sequences were purchased from Eurofins Genomics and cloned into pET28a for His_6 fusion protein expression. Truncated mHMGB1 (aa 1–185) (mHMGB1dCTT) was cloned analogously. cGAS and mHMGB1dCTT point mutants were

generated by QuikChange site-directed mutagenesis with Pfu-Ultra polymerase (Agilent)

followed by DpnI (Fermentas) digestion. For expression in human cells, Flag/haemagglutinin

(HA)-tagged hcGAS (aa 1–522) and HA-tagged hTFAM (aa 43–246) were cloned into

pcDNA5/FRT/TO (Invitrogen, Thermo Fisher Scientific). For localization studies hcGAS was

cloned into pEGFP-C1vector (Clontech) for N-terminal eGFP-tagged cGAS expression.. All

protein constructs abbreviations and descriptions are listed in Supplementary Table 1.

Cell lines and reagents

All DNA oligonucleotides were purchased from Metabion. The exact sequences of stimulatory

DNAs are listed in Supplementary Table 2. Linearized plasmid DNA was obtained by digestion

of pET28M–SUMO1–GFP vector (EMBL) with BamHI (NEB). A 200 bp PCR fragment was

amplified from MBP sequence. Anti-hcGAS (catalytic domain) and a control antibody were

produced by The Service Unit Monoclonal Antibodies (German Research Center for

Environmental Health, Helmholtz Zentrum München). The following antibodies were purchased

commercially: anti-HA–HRP (Cell Signaling, clone 6E2), mouse anti-hTFAM (Abnova), rabbit

anti-hTOM20 (FL-145) (Santa Cruz Biotechnology), mouse anti-HMGB1 (Sigma, clone 2F6),

goat anti-mouse Alexa Fluor 594 (Life Technologies), goat anti-rabbit Alexa Fluor 488 (Life

Technologies), donkey anti-rabbit Alexa Fluor 594 (Life Technologies), and donkey anti-mouse

Alexa594 and Alexa Fluor 488 (Life Technologies). The following cell lines were used:

HEK293T (American Type Culture Collection, CRL-11268), HEK293T STING-KI[31], primary

human fibroblasts (provided by M. Cremer, T. Cremer's group (Ludwig-Maximilians-University

Munich, Biocentre Martinsried)), BLaER1[32] (provided by T. Graf's group (Center for Genomic

Regulation, Universidad Pompeu Fabra and Institució Catalana de Recerca i Estudis Avançats,

Barcelona)), and BLaER1 cGAS-KO, HMGB1-KO and HMGB1-WT MEF (HMGBiotech, HM-

221).

For generation of cGAS-deficient BLaER1 cells (BLaER1 cGAS-KO), a single-guide

(sg)RNA targeting the sequence GAACTTTCCCGCCTTAGGCA**GGG** (protospacer adjacent

motif is in bold type) of the human *MB21D1* gene was cloned via ligation-independent cloning

into pR-U6-gRNA to yield pR-U6-MB21D1 as previously described[33]. BlaER1 cells were

electroporated with pR-U6-MB21D1 and pCMV-mCherry-T2A-Cas9 (ref. 31) expression

plasmids using a GenePulser device (Biorad), and 2 days later FACS-sorted mCherry-positive

cells were sub-cloned by limiting dilution. Monoclonal cell lines were rearranged and duplicated

for genotyping. The genomic locus surrounding the sgRNA binding site was PCR amplified

(cGAS forward primer:

ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTTTTGGCGCGGGCCCCAGTTG;

cGAS reverse primer:

TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGGCCATGCAGAGAGCTTCCGA)

and subjected to deep sequencing using a MiSeq platform (Illumina) as previously described[34].

KO cell clones contained all-allelic frame shift mutations without any wild-type reads. The

deficiency for cGAS was not validated at protein level.

All bought cell lines were kept at low passages to maintain their identity. Non-

commercially available HEK293T STING-KI, BLaER1, and BLaER1 cGAS-KO cell lines were

not authenticated.

The female primary human fibroblast cell line was authenticated as follows. DNA was

isolated separately from the samples. Genetic characteristics were determined by PCR-single-

locus-technology. Twenty-one independent PCR-systems (Amelogenin, D3S1358, D1S1656,

D6S1043, D13S317, Penta E, D16S539, D18S51, D2S1338, CSF1PO, Penta D, TH01, vWA,

D21S11, D7S820, D5S818, TPOX, D8S1179, D12S391, D19S433, and FGA) were investigated

(PowerPlex 21 PCR Kit, Promega). In parallel, positive and negative controls were performed

yielding correct results.

No mycoplasma contamination of the used cells was detected in regular screenings.

Cell culture

Wild-type and cGAS KO BLaER1 cells were cultivated in RPMI medium containing heat-

inactivated 10% FCS, penicillin (100 U ml^{-1}), streptomycin (100 µg ml^{-1}) (Thermo Fisher

Scientific), and 1 mM sodium pyruvate (Thermo Fisher Scientific). For trans-differentiation,

5×10^4 cells were seeded per well of a flat bottomed 96-well plate and cultivated in the presence

of β-oestradiol (100 nM, Sigma-Aldrich), hr-IL-3 (10 ng ml^{-1}), and M-CSF (10 ng ml^{-1}) (both

PeproTech) for 5 days before the experiment[35]. HEK293T, HEK293T STING-KI, and human

fibroblasts were cultured in DMEM (Thermo Fisher Scientific or Sigma-Aldrich, respectively)

supplemented with 10–20% heat-inactivated FBS (Thermo Fisher Scientific or Biochrom,

respectively). All cells were incubated at 37 °C with 5% CO_2.

HMGB1-KO and WT MEFs were cultured in DMEM (Thermo Fisher Scientific)

supplemented with 15% heat-inactivated FBS (Thermo Fisher Scientific), 2 mM L-glutamine

(Sigma, G7513), 1% non-essential amino acids (Sigma, M7145), 0.1 mM β-mercaptoethanol

(Sigma, M3148), and 0.1 mg ml^{-1} penicillin–streptomycin (Sigma, P4333) and incubated at 37 °C with 5% CO_2.

Protein expression and purification

All proteins were overexpressed in *E. coli* Rosetta (DE3) for 16–18h at 18 °C after induction with 0.2 mM IPTG. Cells were lysed by sonication in 50 mM Tris, 500 mM NaCl, 5 mM $MgCl_2$, 10 mM imidazole, 10% glycerol, pH 7.5, supplemented with 2 mM β-mercaptoethanol and protease inhibitor cocktail (Sigma-Aldrich) and purified with Ni-NTA agarose resin (Qiagen). For truncated cGAS proteins and human full-length cGAS without MBP-tag, His$_6$–MBP tag was removed with tobacco etch virus protease (1:50 mass ratio) during 16 h dialysis against 30 mM Tris, 100 mM NaCl, 2 mM DTT, pH 7.0. cGAS proteins were further purified by cation-exchange chromatography (30 mM Tris, 100 mM/1 M NaCl, 2 mM DTT, pH 7.0) on HiTrap SP HP columns (GE Healthcare) followed by size exclusion chromatography (SEC) on HiLoad S200 16/60 column (GE Healthcare) equilibrated with 20 mM Tris, 100 mM NaCl, pH 7.5. Full-length hcGAS with or without N-terminal His$_6$–MBP tag (hcGAS) and mcGAS proteins was concentrated to 8–12 mg ml^{-1}. Truncated hcGAS was concentrated to 4 mg ml^{-1}. All proteins were flash-frozen in liquid nitrogen and stored at −80 °C mTfam and mHMGB1 variants were purified as described for full-length hcGAS with His$_6$-MBP tag, except the cation-exchange chromatography step was omitted and after dialysis against 20 mM Tris, 300 mM NaCl, pH 7.5 SEC on HiLoad S75 16/60 (GE Healthcare) column was performed. Proteins were concentrated to 10–13 mg ml^{-1}.

lHU was purified as described for full-length hcGAS with His$_6$-MBP tag, except after dialysis affinity chromatography on a HiTrap Heparin HP (GE Healthcare) column was performed (20 mM Tris, 100 mM/1 M NaCl, 2 mM DTT, pH 7.5). HiLoad S75 16/60 (GE Healthcare) equilibrated with 20 mM Tris, 100 mM NaCl, pH 7.5 was used for SEC. Protein was concentrated to 7 mg ml^{-1}.

Crystallization of cGAS–DNA complex

For crystallization, purified mcGAS (aa 141–507) 6 mg ml^{-1} was mixed with 39 bp DNA (39 bp-s: AGATCTACTAGTGATCTATGACTGATCTGTACATGATCT; 39 bp-as: AGATCATGTACAGATCAGTCATAGATCACTAGTAGATCT) in a molar ratio of 1:0.6 protein:DNA in a buffer containing 20 mM Tris pH 7.5, 300 mM NaCl and 20 mM $MgCl_2$. Crystals were obtained by hanging-drop vapour diffusion in 0.1 M Tris pH 8, 0.2 M ammonium

citrate pH 7, 27.5% w/v PEG3350 after 1 month at 20 °C. The crystals were soaked in 25%

glycerol diluted in reservoir solution, flash-frozen, and stored in liquid nitrogen.

Data collection and refinement

X-ray diffraction data were collected at the beamline PXI (X06SA) at the Swiss Light Source,

Switzerland, using a Pilatus 6M detector. Data sets were processed with XDS and merged with

XSCALE[36]. The STARANISO server[37] was used to generate structure factor amplitudes and

their s.d. with a local $I/\sigma(I)$ cut-off of 1.9. The resulting data set showed varying high-resolution

cut-offs between 4.8 Å and 3.6 Å and an effective resolution of ~4.2 Å. The crystal structure was

solved by molecular replacement with PHASER[38,39] using six copies of a search model based on

the published structure of mouse cGAS[cd] in complex with 18 bp DNA[8] (Protein Data Bank

(PDB) accession number 4LEY). Iterative model building and refinement was done with the

molecular graphics program MOLOC[40] and the CCP4 suite refinement program REFMAC5 (ref.

41). Owing to the rather low resolution, we could not identify the DNA sequence. Thus, sense

and anti-sense strands, as well as the starting nucleotide, were chosen arbitrarily. At the

symmetry contact between two neighbouring cGAS dimers, we saw continuous electron density

for a 34 bp DNA passing through the crystallographic twofold axis. Since the DNA sequence

was not palindromic, we modelled the DNA with two 17 bp long symmetry-halves. Except for

removal of a steric clash with the bound DNA by choosing a different rotamer for Arg244, we

did not attempt to re-model the cGAS protein. Data collection and refinement statistics are listed

in Supplementary Table 3.

Fluorescence-based cGAS activity assays

In fluorescence-based cGAS activity assays, a fluorescent analogue of ATP (2-aminopurine

riboside-5′-*O*-triphosphate (Biolog) (fATP)) was used. DNA (6.5 ng µl⁻¹) of different lengths

(20–100 bp in 5 bp intervals, and pET28M–SUMO1–GFP vector (EMBL) (6.2 kbp, plasmid))

corresponding to roughly 0.5 µM 20 bp (approximate length of cGAS binding site) was

premixed with 0.5 µM cGAS in 40 mM Tris pH 7.5 and 100 mM NaCl. Alternatively,

2.6 ng µl⁻¹ DNA (~0.2 µM binding sites) and 1 µM cGAS, 13 ng µl⁻¹ DNA (~1 µM binding

sites), and 0.2 µM cGAS or 2.6 ng µl⁻¹ DNA (~0.2 µM binding sites) and 0.2 µM cGAS were

used. To compare cGAS mutants, the same 2.6 ng µl⁻¹ DNA (~0.2 µM binding sites) was mixed

with 1 µM cGAS. The reaction was started by adding 5 mM MgCl₂ with 500 µM GTP and

50 µM fATP and performed at 32 °C. Fluorescence decrease was measured in 96-well black

non-binding PS plates (Greiner Bio-One) on Tecan infinite M1000 (λ_{ex} = 305 nm, λ_{em} = 363 nm, gain 100, settle time 10 ms, kinetic interval 2 min). In cGAS titration experiments, 2.6 ng µl^{-1} DNA (~0.2 µM binding sites) was mixed with increasing cGAS concentrations from 0.05 to 20 µM.

cGAS stimulation assays induced by DNA-bending protein were performed analogously. Briefly, 13 ng µl^{-1} DNA (~1 µM binding sites) was premixed with DNA-bending proteins or cGAS inactive mutant in concentrations of 0–5 µM, after which 100 nM or 50 nM cGAS in 40 mM Tris pH 7.5 and 100 mM NaCl was added. The reaction was started by adding 5 mM MgCl$_2$ together with 500 µM GTP and 50 µM fATP and performed at 32 °C for mcGAS[cd] or 37 °C for hcGAS. Fluorescence measurement was made as described above.

Data were analysed with OriginPro 8G (OriginLab).

Radiolabelled cGAS activity assays

Radiolabelled cGAS activity assays were performed analogously as previously described[7]. Briefly, 13 ng µl^{-1} DNA (~1 µM binding sites) were mixed with 2 µM mcGAS[cd] and the reaction started by adding 50 µM ATP, 500 µM GTP, 5 mM MgCl$_2$ in 40 mM Tris pH 7.5, and 100 mM NaCl containing 1:800 [α^{32}P]ATP (3,000 Ci mmol^{-1}, Hartman Analytic). Samples were incubated at 35 °C and the reactions were stopped by plotting on PEI-Cellulose F plates (Merck) and analysed by thin-layer chromatography with 1 M (NH$_4$)$_2$SO$_4$/1.5 M KH$_2$PO$_4$ as running buffer. The radiolabelled products were visualized with a Typhoon FLA 9000 phosphor imaging system. For testing of fATP incorporation into cGAS enzymatic activity product, 1 µM mcGAS[cd] was mixed with 13 ng µl^{-1} (~1 µM binding sites) 55 bp, 500 µM fATP/ATP, 500 µM GTP, 5 mM MgCl$_2$, and 1:160 [α^{32}P]ATP or [α^{32}P]GTP (3,000 Ci mmol^{-1}, Hartman Analytic) in the same condition. For testing TFAM-dependent increase in cGAS activity, 250 nM mcGAS[cd] was mixed with 13 ng µl^{-1} (~1 µM binding sites) plasmid in the same condition and 1:1,000 instead of 1:160 [α^{32}P]ATP was added. The reaction was incubated at 32 °C and the radiolabelled products were separated and visualized as described previously. For testing cGAS activation by short DNAs, the previously described protocol was used[8]. Briefly, 10 µM or 5 µM cGAS was mixed with 650 ng µl^{-1} or 325 ng µl^{-1} DNA, respectively, in buffer containing 5 mM MgCl$_2$, 2 mM ATP and GTP, and 1:400 [α^{32}P]ATP. The reaction mixture was incubated at 37 °C and the radiolabelled products were separated and visualized as described previously.

Anion-exchange chromatography of cGAS reaction products

Ten micromolar mcGAS^{cd} or hcGAS^{cd} and 195 ng µl^{-1} (~15 µM binding sites) plasmid DNA

were incubated at 35 °C for 2 h in buffer containing 40 mM Tris pH 7.5, 100 mM NaCl, 10 mM

MgCl$_2$, 2 mM fATP, and 2 mM GTP. Reaction mixtures were centrifuged for 10 min at 16,100

relative centrifugal force, and the supernatant was separated from cGAS by ultrafiltration

(30 kDa, Amicon). Resulting flow through was diluted in 50 mM Tris pH 9.0 and loaded on a

Mono Q 5/50 GL (GE Healthcare) for anion exchange chromatography (50 mM Tris, 0 M/1 M

NaCl, pH 9.0). Control runs with fATP, GTP, 2′3′- and 3′3′-cGAMP were made analogously to

validate the resulting peaks of cGAS reaction products.

Electrophoretic mobility shift assay

DNA (2.6 ng µl^{-1}) of length 200 bp (~0.2 µM binding sites) was incubated with 0–2.5 µM WT

or mutant mHMGB1dCTT for 30 min on ice in the buffer containing 50 mM Tris pH 7.5 and

100 mM NaCl. Samples were separated in 0.6% agarose gel prepared with Gel-Red (Biotium) in

40 mM Tris pH 9.2 as running buffer. The gel images were obtained with Gel iX Imager (Intas).

ITC

The calorimetric titration of DNA (60–400 µM in the syringe) to mcGAS^{cd} (20 µM in the

reaction cell) was performed with MicroCal PEAQ-ITC (Malvern) in buffer containing 30 mM

HEPES, 100 mM NaCl, pH 7.5. The following parameters were used: 25 °C, 1×0.4 µl +

13×3 µl injections, 10 µcal s^{-1} reference offset, 750 revolutions per minute syringe stirring

speed, 60 s initial delay, 6 s injection duration, 150 s spacing between injections, high feedback.

To study TFAM direct interaction with cGAS, 50 µM human full-length cGAS without MBP-tag

in the reaction cell and 400–500 µM hTFAM in the syringe were used. ITC was performed with

the same parameters as above at 25 °C, 35 °C, and 15 °C. K_d, molar ratio, Gibbs free energy,

enthalpy, and entropy of the binding were calculated with MicroCal PEAQ-ITC analysis

software (Malvern).

SEC–RALS

SEC of cGAS–DNA complexes was performed on a Superose 6 increase 10/300 column (GE

Healthcare). mcGAS^{cd} (4 mg ml^{-1}) was mixed with DNA in 1:0.6 molar ratio for 50 bp DNA or

1:0.3 molar ratio for 70 and 80 bp DNA in binding buffer containing 20 mM Tris, 243 mM

NaCl, pH 7.5. Alternatively, 6 mg ml^{-1} mcGAS^{cd} and 20 bp DNA in 1:1.1 molar ratio were used.

One hundred microlitres of the mix were loaded to Superose 6 increase 10/300 column (GE

Healthcare) equilibrated with 20 mM Tris, 100 mM NaCl, pH 7.5 buffer and separate peaks were
analysed with RALS using Viscotek 270 Dual Detector (Malvern) and Viscotek VE3580
Refractive Index Detector (Malvern). Molecular masses of the complexes were calculated with
OmniSEC 4.7.0 software (Malvern).

ELISA

Fifty thousand trans-differentiated BLaER1 cells were transfected with 20, 40, or 60 ng DNA of
different lengths (20–100 bp in 5 bp intervals) and herring testis DNA using 0.5 µl
Lipofectamine 2000 Transfection Reagent (Thermo Fischer Scientific) in 50 µl Opti-MEM
Reduced Serum medium (Thermo Fischer Scientific). CXCL10 expression in the supernatants
was quantified 8 h after transfection using ELISA (BD OptEIA, human IP-10 ELISA Set). Cells
stimulated with transfection reagent only and unstimulated cells served as control.

IFN-β mRNA expression analysis

Fifty thousand trans-differentiated WT and cGAS-KO BLaER1 cells were transfected with 40 ng
DNA of different lengths (20, 40, 60, 80, 100 bp and herring testis DNA). Eight hours after
stimulation, total RNA was isolated from 7.5×10^5 pooled cells using an RNeasy Mini Kit
(Qiagen) following the vendor's recommendations. Subsequently, RNA was digested with
DNase I (Thermo Fisher Scientific) to remove residual DNA. Five hundred nanograms of RNA
were reverse transcribed using poly $(dT)_{18}$ oligonucleotides according to the manufacturer's
instructions (RevertAid cDNA Synthesis Kit). IFN-β expression levels were analysed by
quantitative PCR, using gene-specific primers (IFN-β forward primer:
CAGCATCTGCTGGTTGAAGA; reverse primer: CATTACCTGAAGGCCAAGGA),
normalized to GAPDH expression measured analogously (GAPDH forward primer:
GAGTCAACGGATTTGGTCGT; reverse primer: GACAAGCTTCCCGTTCTCAG) and the
fold change was calculated on the basis of the unstimulated control.

Luciferase reporter assays

All immunostimulatory assays were performed in HEK293T STING-KI cells[31]. Five hundred
thousand cells were seeded on 24-well plates and transfected with 100 ng p-125Luc[42], 10 ng
pGL4.74 (Promega), 50 ng Flag/HA–cGAS plasmids, and a total of 500 ng DNA per well (filled
up with an empty vector pcDNA5) using Lipofectamine 2000 (Invitrogen, Thermo Fisher
Scientific) as transfection reagent according to the vendor's protocol. After 14 h, cells were lysed

in 200 µl passive lysis buffer (Promega). Immunoactivity experiments using a Dual-Glo luciferase assay system (Promega) were performed as previously described[43].

Co-immunopurification of cGAS and TFAM from HEK293T cells

Ten million HEK293T cells were transfected with 10 µg Flag–HA–cGAS and 10 µg HA–TFAM expression vectors and harvested 24 h after transfection. Cell pellets were flash-frozen in liquid nitrogen and stored at −20 °C. For immunoprecipitation, cells were incubated in Nonidet P-40 lysis buffer (50 mM HEPES, 150 mM KCl, 1 mM NaF, 0.5% NP-40, 2 mM DTT, 10 µM $ZnCl_2$, protease inhibitor (Sigma Aldrich), pH 7.5) for 10 min on ice. Lysates were cleared by centrifugation for 30 min. Proteins were immunoprecipitated for 1.5 h with anti-hcGAS or a control antibody bound to magnetic protein G Dynabeads (Novex, Life Technologies). Beads were washed four times with Nonidet-P40 lysis buffer and subjected to SDS–polyacrylamide gel electrophoresis and immunoblotting.

Fluorescence microscopy

HEK293T cells and primary human fibroblasts were grown at 40–45% and 50–60% confluency, respectively, on coverslips in a six-well plate overnight and treated the next day with the inhibitors Q-VD-OPH (MP Biomedicals) and ABT-737 (Santa Cruz Biotechnology) at a final concentration of 10 µM for 0 h or 6 h. Cells were fixed with 3% paraformaldehyde and 0.1% glutaraldehyde in PBS for 12 min, reduced with 1 mg ml^{-1} $NaBH_4$ for 7 min, permeabilized with 0.25% (v/v) Triton X-100, and blocked with 3% (w/v) BSA in PBS for 3 h.

MEF and MEF HMGB1-KO cells were grown at 40–50% confluency on coverslips in a six-well plate overnight and transfected with 2 µg pcDNA5/FRT/TO (Invitrogen, Thermo Fisher Scientific) or eGFP–cGAS expression vectors overnight. Cells were then fixed with 2% paraformaldehyde in PBS for 10 min at room temperature, permeabilized with 0.5% (v/v) Triton X-100, and blocked with 2% (w/v) BSA in PBST for 10 min.

Immunostaining was performed for 1 h at room temperature with primary antibodies against TFAM (1:50), TOM20 (1:250), or HMGB1 (1:100). The primary antibodies were detected with the secondary anti-mouse or anti-rabbit goat or donkey antibodies (1:500) by incubation for 1 h. After immunostaining, samples were post-fixed in 4% paraformaldehyde for 10 min, stained with 1 µg ml^{-1} (HEK293T and primary human fibroblasts) or 10 µg ml^{-1} DAPI (MEF) for 10 min, and mounted in VECTASHIELD (Vector Laboratories). Cells were washed three times with PBST after each step.

Three-dimensional structured illumination microscopy (3D SIM) was performed with a DeltaVision OMX V3 microscope (GE Healthcare), equipped with a ×100/1.40 numerical aperture PlanApo oil immersion objective (Olympus), Cascade II:512 EMCCD cameras (Photometrics), and lasers for 405 nm, 488 nm, and 594 nm. Image stacks were first reconstructed and corrected for colour shifts with softWoRx 6.0 Beta 19 (unreleased) software. After establishing composite tiff stacks with a custom-made macro in Fiji, the data were subsequently aligned again and maximum intensity projections were used.

Wide-field fluorescent microscopy was performed with a personal DeltaVision (pDV) microscope (GE Healthcare) equipped with a ×60/1.42 oil-immersion objective PlanApo U (Olympus), Cool-Snap camera (12 bit, 1024 pixels × 1024 pixels, Photometrics) by acquiring one focal plane. Insight SSI LEDs (GE Healthcare) for 405 nm, 488 nm, 594 nm were used.

Statistical analysis

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments or accessing the outcome.

All experiments were conducted at least three times and the mean values and s.d. of technical or biological replicates (BLaER1 stimulation experiments) were calculated. Statistical significance, if applicable, was calculated on the basis of unpaired one-sided *t*-tests.

For the data fitting in Fig. 3, the description is included in the figure legend and in Supplementary Methods.

For quantification of TFAM signal in microscopy images presented in Extended Data Fig. 9c, maximum intensity projections were subsequently loaded as RGB into Volocity calculation software (Volocity 6.1.2 (Perkin Elmer)). Areas corresponding to mitochondria, nucleus, and cytosol were obtained, segmented, and measured in all channels by defining intensity threshold and minimum object size on each channel. Cytosolic TFAM signal was quantified according to number of counts and signal area, and additionally corrected and normalized to the cytosolic area in an image. For each calculation, three different cells and three segments per cell were used.

Figure preparation

Figures showing protein structures and electron densities were prepared with PyMOL Molecular Graphic Systems[44]. Sequence alignment (Extended Data Fig. 8a) was prepared with Jalview[45]. All other figures were prepared with Microsoft Excel, OriginPro 8G (OriginLab), or Matlab_R2015a (The MathWorks).

Data availability

The coordinates and structure factors have been deposited in the PDB under accession number 5N6I. All other data are available from the corresponding author upon reasonable request.

Code availability

Matlab code used in this study is available from the corresponding author upon reasonable request.

REFERENCES

31.   Ablasser, A. et al. Cell intrinsic immunity spreads to bystander cells via the intercellular transfer of cGAMP. *Nature* **503**, 530–534 (2013).

32.   Rapino, F. et al. C/EBP☐ induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Report*s **3**, 1153–1163 (2013).

33.   Schmidt, T., Schmid-Burgk, J. L. & Hornung, V. Synthesis of an arrayed sgRNA library targeting the human genome. *Sci. Rep*. **5**, 14987 (2015).

34.   Schmid-Burgk, J. L. et al. OutKnocker: a web tool for rapid and simple genotyping of designer nuclease edited cell lines. *Genome Res*. **24**, 1719–1723 (2014).

35.   Gaidt, M. M. et al. Human monocytes engage an alternative inflammasome pathway. *Immunity* **44**, 833–846 (2016).

36.   Kabsch, W. XDS. *Acta Crystallogr*. *D* **66**, 125–132 (2010).

37.   Tickle, I. J. et al. STARANISO (Global Phasing, 2016).

38.   Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr*. *D* **50**, 760–763 (1994).

39.   Read, R. J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr*. *D* **57**, 1373–1382 (2001).

40.   Gerber, P. R. & Müller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aided Mol. Des*. **9**, 251–268 (1995).

41.   Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr*. *D* **67**, 355–367 (2011).

42.   Yoneyama, M. et al. Autocrine amplification of type I interferon gene expression mediated by interferon stimulated gene factor 3 (ISGF3). *J. Biochem*. **120**, 160–169 (1996).

43.   Lässig, C. et al. ATP hydrolysis by the viral RNA sensor RIG-I prevents unintentional recognition of self-RNA. *eLife* **4**, e10859 (2015).

44.   Schrödinger. The PyMOL molecular graphics system, version 1.8 (2010).

45.   Waterhouse, A. M. et al. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191 (2009).

EXTENDED DATA FIGURES
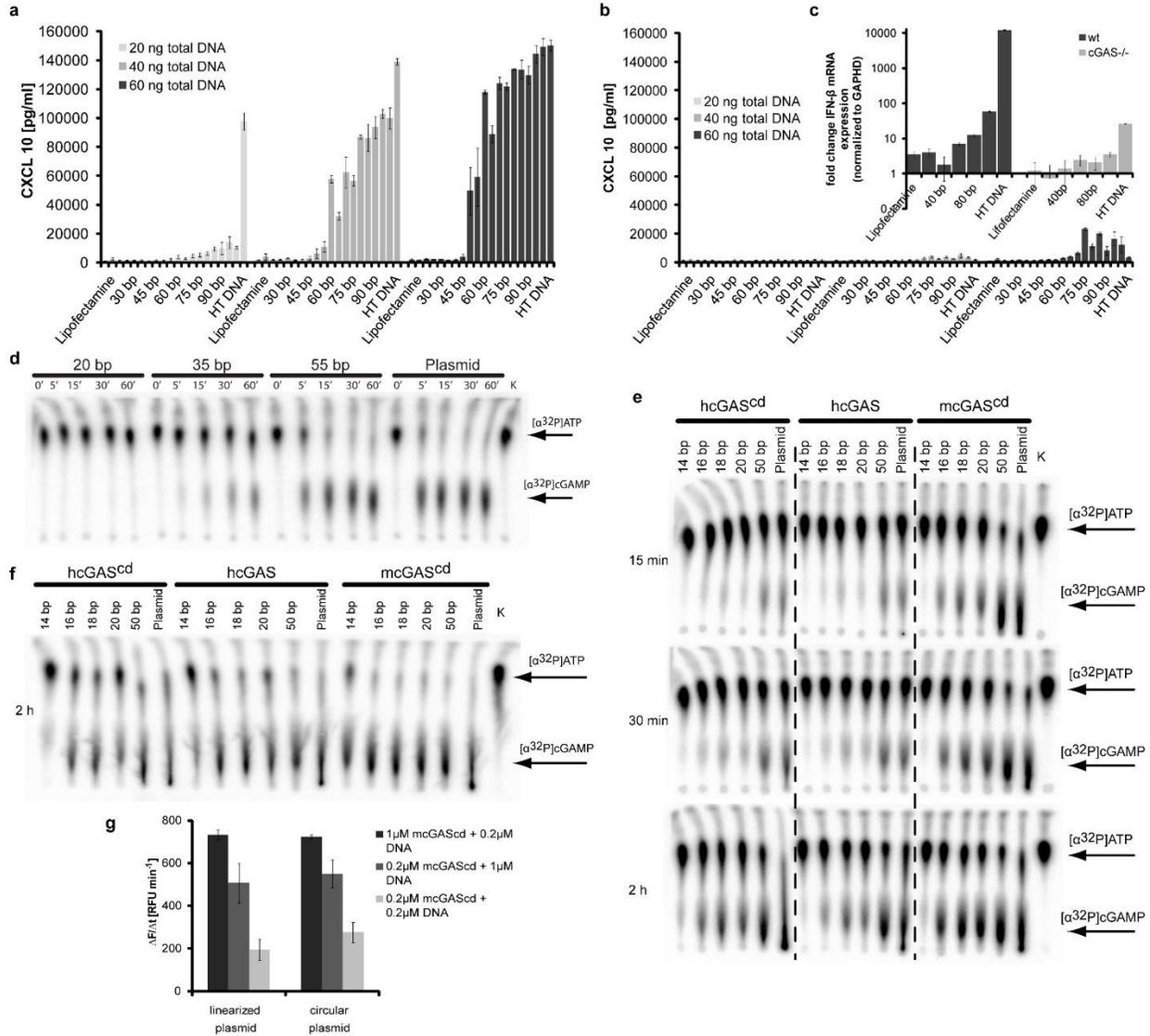
Extended Data Figure 1. cGAS activity increases with DNA length *in vitro* and *in vivo*. a, **b**, CXCL 10 cytokine production of WT (**a**) and cGAS KO (**b**) transdifferentiated BLaER1 cells. Cells were transfected with 20, 40, and 60 ng DNA of increasing length (20–100 bp in 5 bp intervals) and herring testis (HT) DNA, and CXCL 10 concentration in the supernatant was measured by ELISA. The first two bars of each series represent unstimulated cells and lipofectamine controls. Shown are mean values ± s.d., *n* = 3. **c**, Fold change of IFN-β mRNA expression in WT and cGAS-KO BLaER1 cells transfected with 20–100 bp and herring testis

DNA. IFN-β mRNA expression levels were normalized to GAPDH mRNA and the fold change was calculated on the basis of the unstimulated control. Shown are mean values ± s.d., $n = 3$. **d**, Radiolabelled cGAMP production of cGAS stimulated with DNA of different lengths (20, 35, 55 bp, and plasmid). cGAS reactions in the presence of [α^{32}P]ATP were stopped at the indicated time points and radiolabelled compounds (shown with black arrows) were visualized. **e**, **f**, Radiolabelled cGAMP production by (left to right) hcGAS^{cd}, hcGAS, and mcGAS^{cd} stimulated with 14, 16, 18, 20, 50 bp and plasmid DNA. cGAS (**e**, 5 µM; **f**, 10 µM) was incubated with 325 ng µl^{-1} (**e**) or 650 ng µl^{-1} (**f**) DNA of indicated length in the presence of ATP, GTP, and [α^{32}P]ATP at 37 °C. The reactions were stopped at the indicated time points and radiolabelled compounds (shown with black arrows) were visualized. **g**, mcGAS^{cd} activity measured by the rate of fATP incorporation into fGAMP (see Extended Data Fig. 2 for the assays) in the presence of linearized or circular plasmid DNA. Mean values of initial cGAS reaction rates ($\Delta F/\Delta t$) are plotted against DNA constructs ± s.d., $n = 3$. No significant difference between linearized and circular plasmid could be detected.
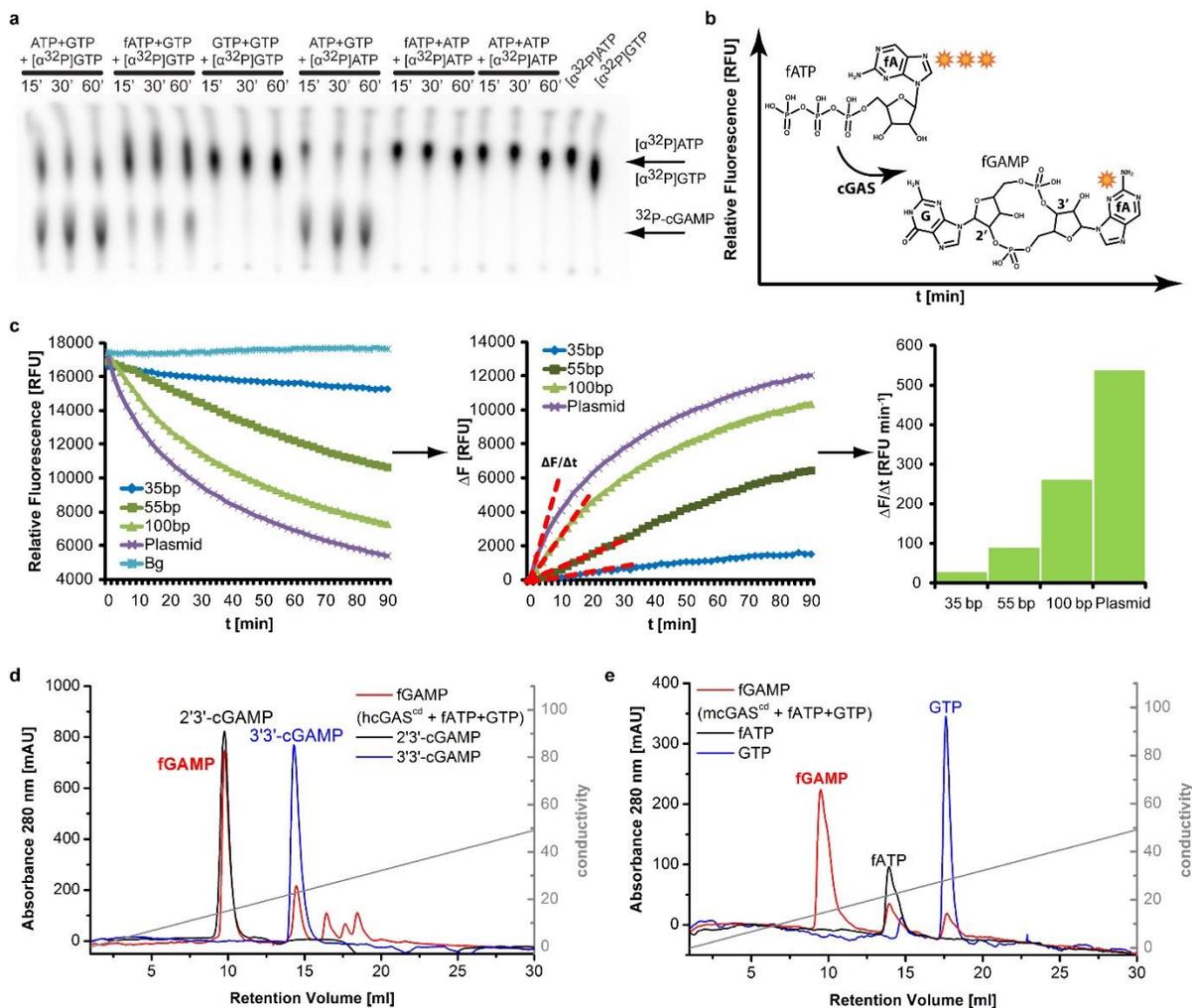
Extended Data Figure 2. Description of the fluorescence-based cGAS activity assay. a, Radiolabelled cGAS products obtained with different NTP combinations. Reactions were stopped at the indicated time points, and products separated by thin layer chromatography and visualized by radiography. **b**, The principle of the fluorescence-based cGAS activity assay. cGAS catalyses the conversion of fluorescent ATP analogue fATP (2-aminopurine riboside triphosphate, three orange stars) into less fluorescent fGAMP (fluorescent cGAS product, one orange star), resulting in a gradual decrease in fluorescence intensity during the reaction. **c**, General workflow for calculating the initial cGAS reaction rates. From initial fluorescence curves (left), the background fluorescence was subtracted and the resulting curve was inverted for better visualization (ΔF). Initial rates were calculated as a slope of the linear intervals (red dashed lines) and defined as $\Delta F/\Delta t$ (relative fluorescence units per minute) (right). **d**, **e**, fGAMP mobility in anion-exchange chromatography on a MonoQ 5/50 GL column. **d**, Comparison of

fGAMP, produced by cGAS from fATP and GTP (red), 2′3′-cGAMP (black), and 3′3′-cGAMP (blue) mobilities. **e**, Comparison of fGAMP (red), fATP (black), and GTP (blue) mobilities.
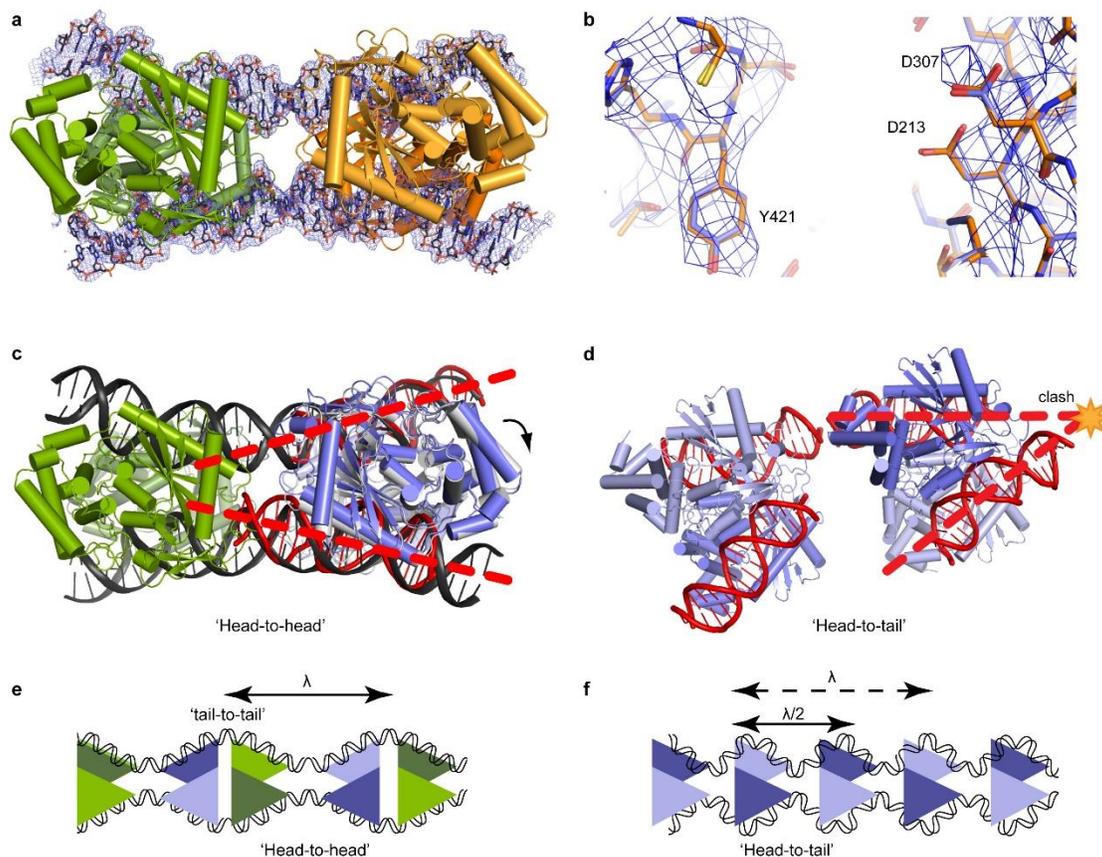
Extended Data Figure 3. Electron densities and comparison with cGAS in complex with 18 bp DNA. a, $2F_o − F_c$ electron density of 39 bp DNA within the complex at a contour level of 1σ. **b**, Close-up view of the cGAS active site (orange) superposed with a previously published structure of cGAS bound to 18 bp DNA (PDB accession number 4LEY, blue). Shown is the $2F_o − F_c$ electron density of the active-site residues at a contour level of 1σ. **c**, Superposition of the cGAS complex with 39 bp DNA (green and grey) with 18 bp DNA-bound cGAS (PDB accession number 4LEY, blue). A slight rotation of superposed 18 bp-cGAS (blue) relative to 39 bp-cGAS (grey) is shown with a black arrow. The superposition shows the difference between a hypothetical straight (red dashed line) DNA, leading the previously proposed DNA end preference of cGAS, and the curved DNA observed in our crystal structure. **d**, Binding mode of the DNA strands to cGAS within the previously published structure (PDB accession number 4LEY). Two neighbouring crystallographic asymmetric units represent a 'head-to-tail' cGAS dimer orientation. Elongation of both strands (red dashed line) leads to a steric clash (orange

star). **e**, Schematic model of cGAS binding to continuous DNA in alternating 'head-to-head' and 'tail-to-tail' arrangements. DNA curves over λ bp (black arrows). **f**, Schematic model of cGAS binding to continuous DNA in a (not observed) 'head-to-tail' arrangement. DNA curves over $\lambda/2$ bp (black arrows), if the same density of cGAS per DNA is assumed. Thus DNA must be bent twice more often than in model **e** to sustain the parallel DNA arrangement and is energetically less favourable.
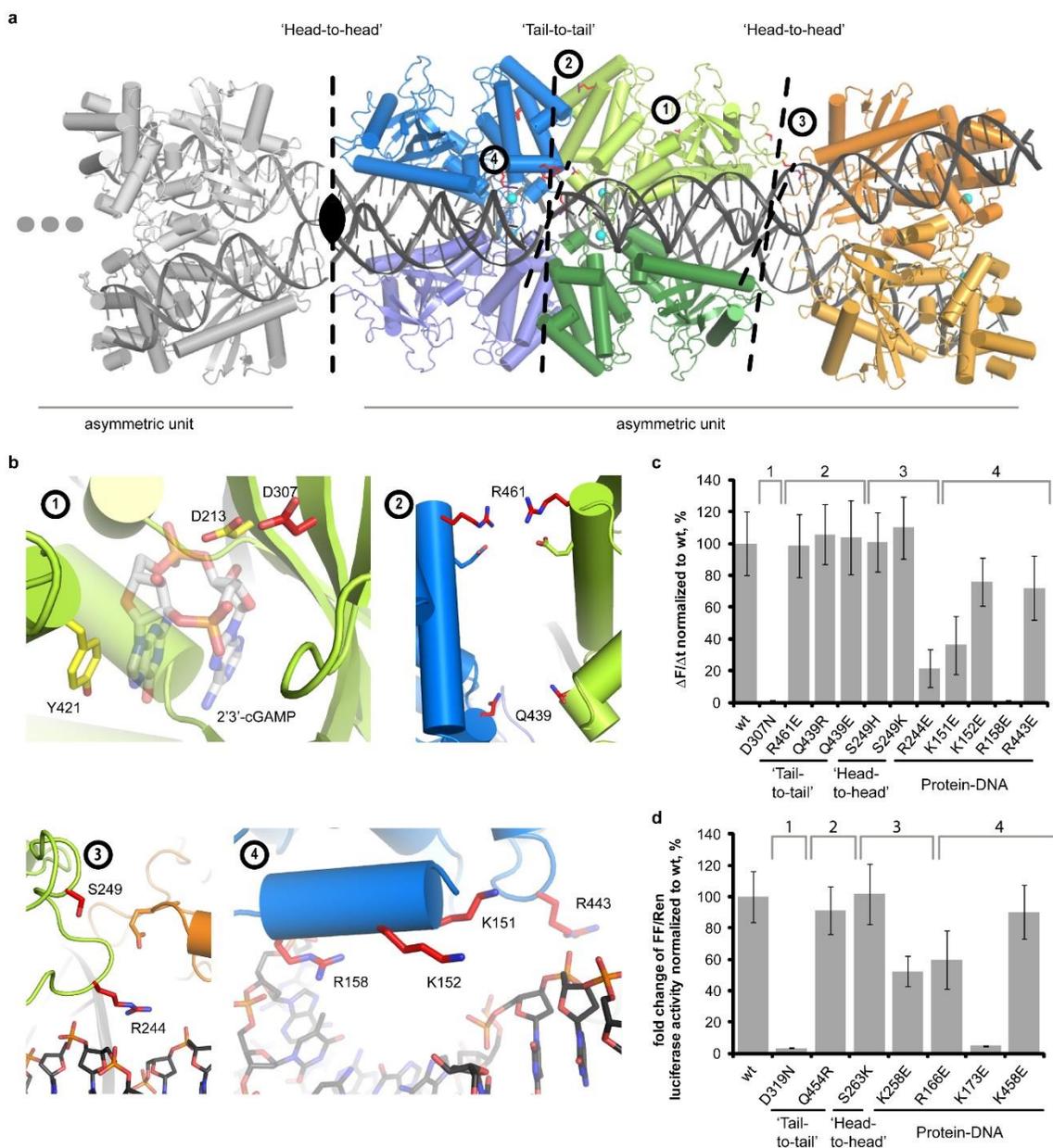
Extended Data Figure 4. Asymmetric unit of the cGAS–39 bp complex and details of protein–protein and protein–DNA contacts. a, An overview of the asymmetric unit of 39 bp DNA-bound cGAS. Filled oval-like symbol represents a twofold crystallographic symmetry axis;

black dashed lines represent non-crystallographic twofold symmetry axes. The asymmetric unit

contains one full and one half 'head-to-head'-oriented $cGAS_4$–DNA_2 complex. Residues mutated

to examine four areas (encircled 1-4) are shown as red sticks. **b**, Close-up view of mutated

interfaces: 1, cGAS active site (yellow) with superimposed cGAMP (from PDB accession

number 4LEZ, grey) and labelled active-site residues D213 and D307; 2–4, potential protein–

protein or DNA–protein interaction sites, respectively. Mutated residues are in red. **c**, **d**,

Mutational analysis of the described regions of $mcGAS^{cd}$ *in vitro* and hcGAS *in vivo*,

respectively. D307N and D319N correspond to active-site mutations in mouse and human cGAS,

respectively. **c**, cGAS activity measured by the rate of fATP incorporation into fGAMP ($\Delta F/\Delta t$,

see Extended Data Fig. 2) and normalized to WT. Mean values represent percentage of WT

activity ± s.d., $n = 3$. **d**, Percentage change of IFN-β promoter-driven luciferase activity upon the

expression of WT or mutant cGAS. IFN-β response was measured as a proportion of firefly (FF)

to *Renilla* (REN) luciferase activity in HEK293T STING-KI cells upon Flag/HA–hcGAS

overexpression. All ratios were normalized to to WT. Mean values represent the percentage of

WT activity ± s.d., $n = 3$. Mutants are named according to their position in hcGAS: D319N

(active site), Q454R, S263K, R166E, K173E, K458E, K258E mutants refer to D307N, Q439R,

S249K, K151E, R158E, R443E, and R244E in mcGAS.
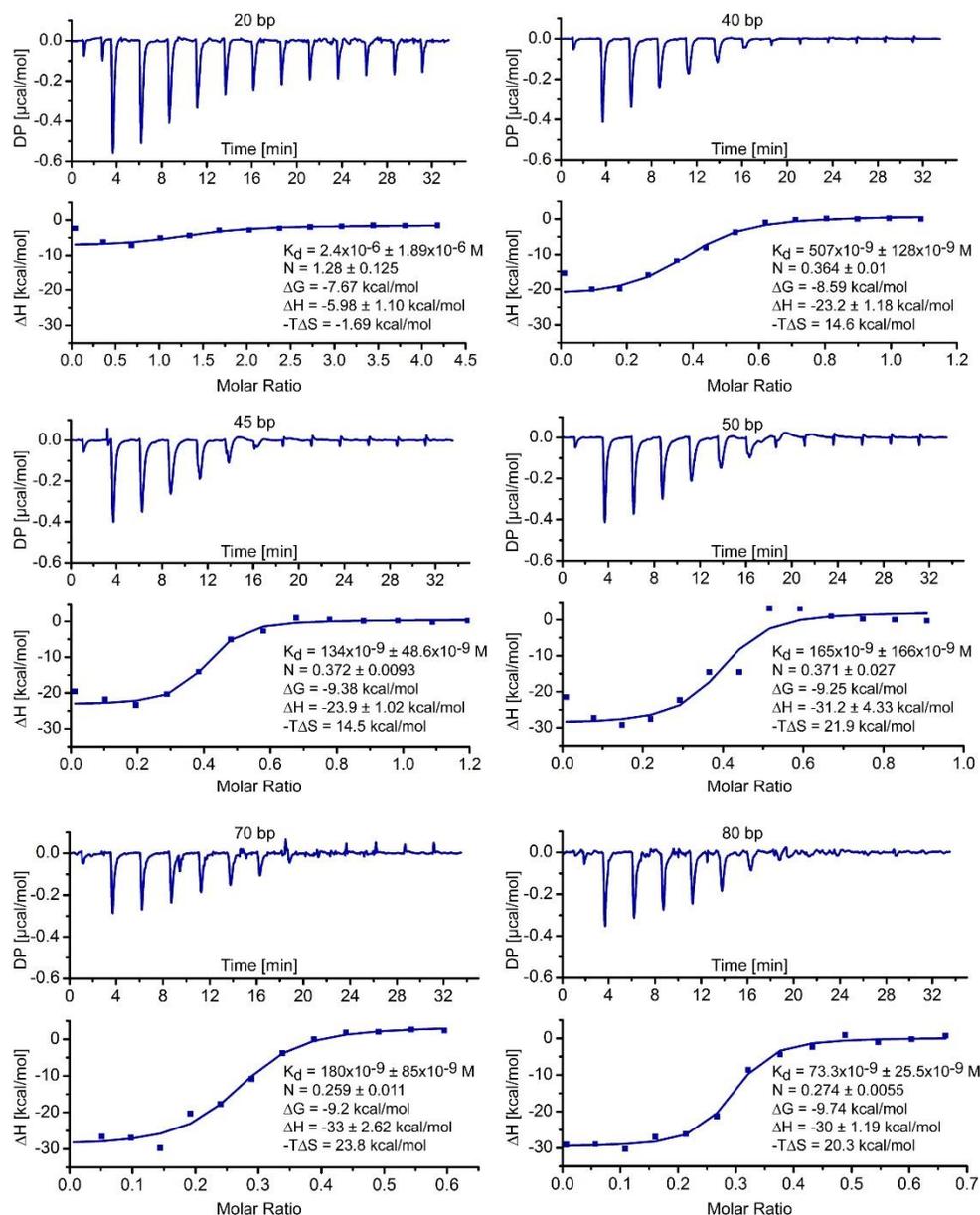
Extended Data Figure 5. cGAS affinity to DNA increases with DNA length accompanied by increase in number of cGAS binding sites along the DNA. ITC measurements of mcGAS[cd] binding to 20, 40, 45, 50, 70, and 80 bp DNA. For each DNA, the power differential (DP) is plotted against time and ΔH is plotted against the molar ratio of DNA:cGAS. Calculated binding parameters are given on each graph.
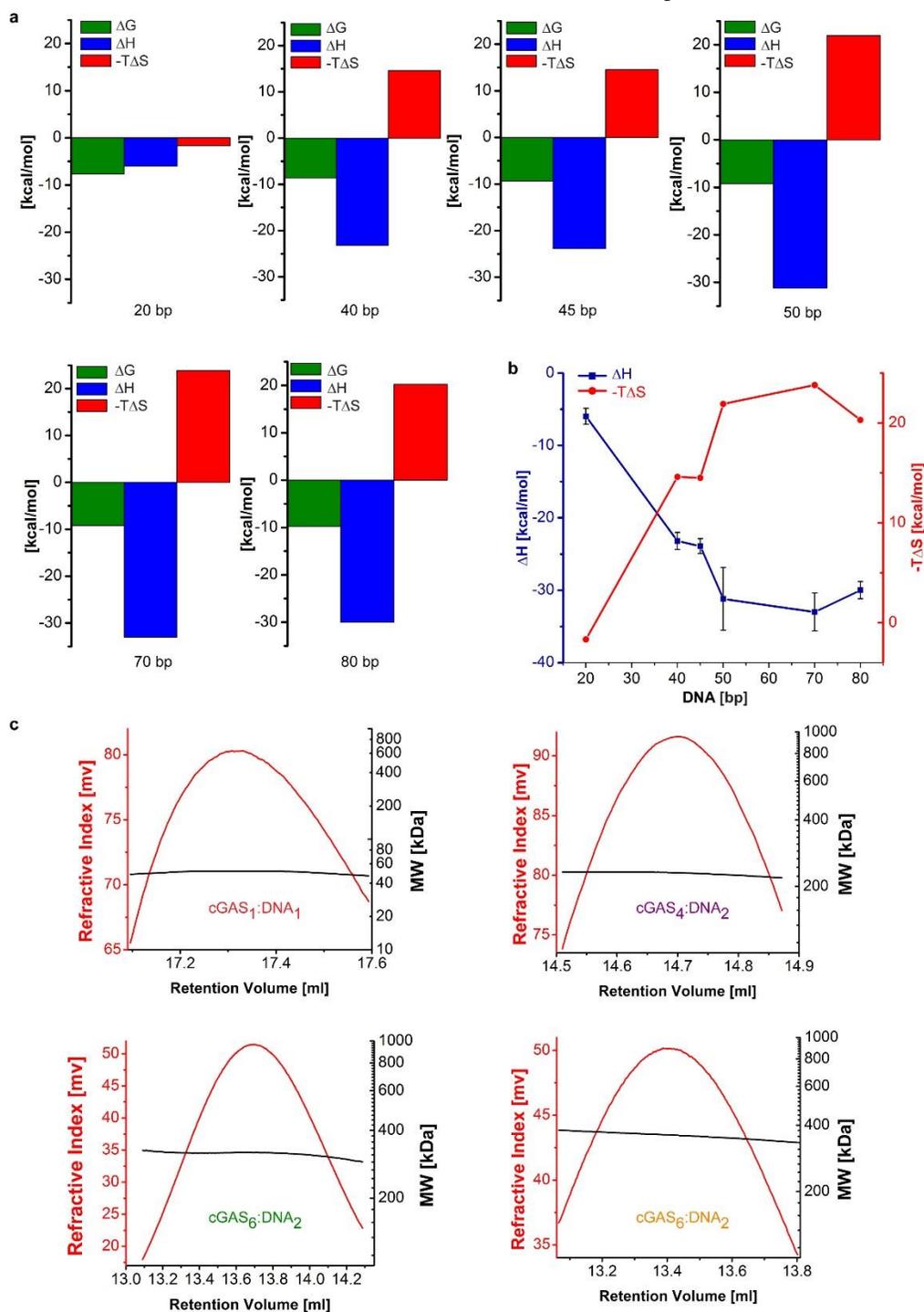
Extended Data Figure 6. cGAS forms higher-ordered structures on long DNA. a, Signature plots for each ITC measurement (Extended Data Fig. 5) showing ΔG (green), ΔH (blue), and $-T\Delta S$ (red) for each binding reaction. **b**, ΔH and $-T\Delta S$ components of Gibbs free energy measured with ITC for mcGAS[cd] binding to 20–80 bp DNA are plotted against DNA length ± s.d. of the measured values from the fit. **c**, Molecular mass distribution within SEC

peaks (Fig. 2c-f) containing mcGAS$^{cd}_1$–DNA$_1$ (with 20 bp DNA), mcGAS$^{cd}_4$–DNA$_2$ (with 50 bp DNA), and mcGAS$^{cd}_6$–DNA$_2$ (with 70 and 80 bp DNA) calculated with RALS. Refractive index (red) and estimated molecular mass (black) are plotted against retention volume.
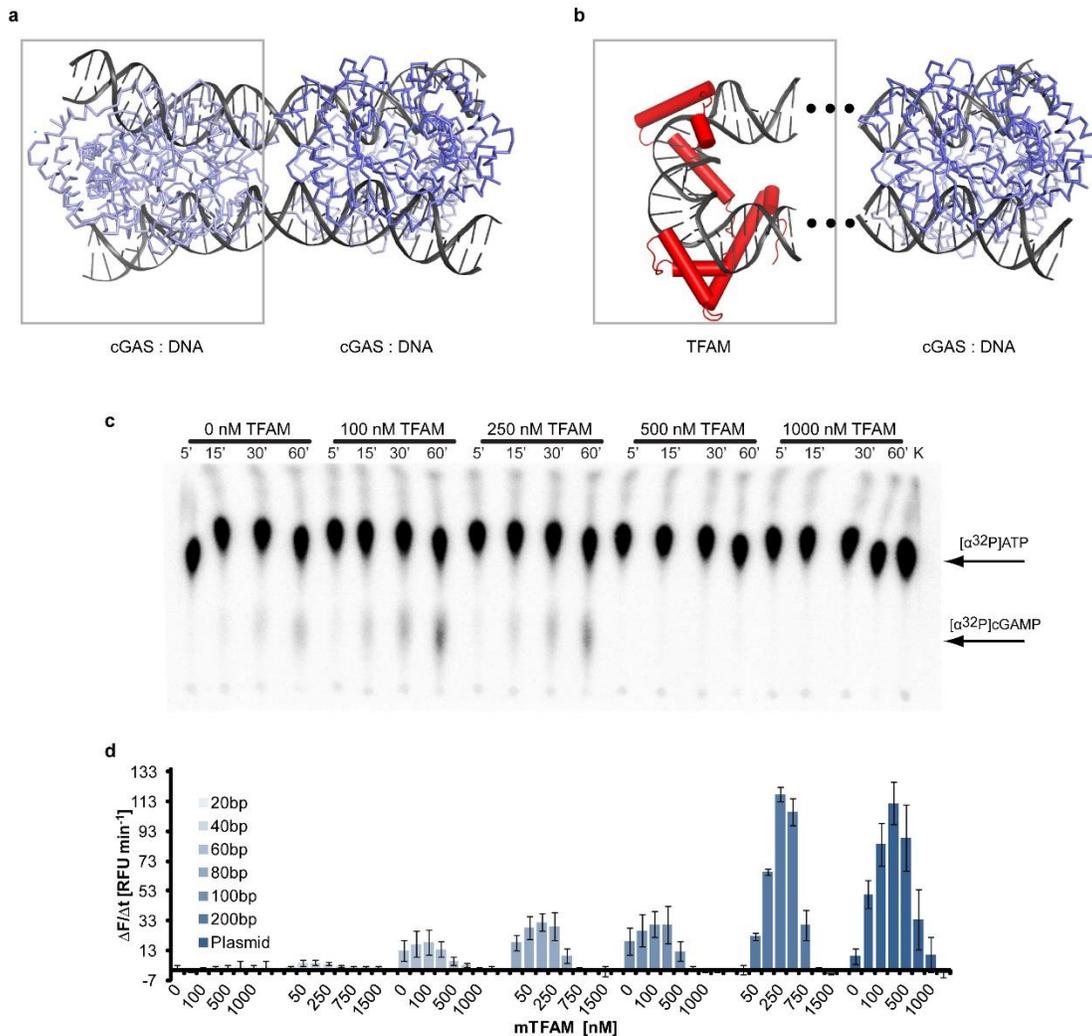
Extended Data Figure 7. TFAM enhances cGAS activity by prearranging DNA into U-shape. a, DNA ladders (blue) and (**b**) DNA-bound TFAM structure (PDB 3TMM, red) arrange DNA in a remarkably similar fashion. The two similarly spaced DNA strands flanking cGAS ladders or the TFAM U-turn (black dotted lines) provide a possible explanation for the effect of TFAM on cGAS activity. **c**, Radiolabelled cGAMP production in the presence of increasing TFAM concentrations. cGAS reactions with ATP, GTP, and [α^{32}P]ATP were stopped at the indicated time points and the radiolabelled compounds (shown with black arrows) were visualized. **d**, Activation of mcGAS^{cd} by mTFAM and DNA of increasing length (20–200 bp or

plasmid DNA). Mean values of initial cGAS reaction rates ($\Delta F/\Delta t$, see Extended Data Fig. 2) are

plotted against increasing concentrations of mTFAM $\pm$ s.d., $n = 4–8$.
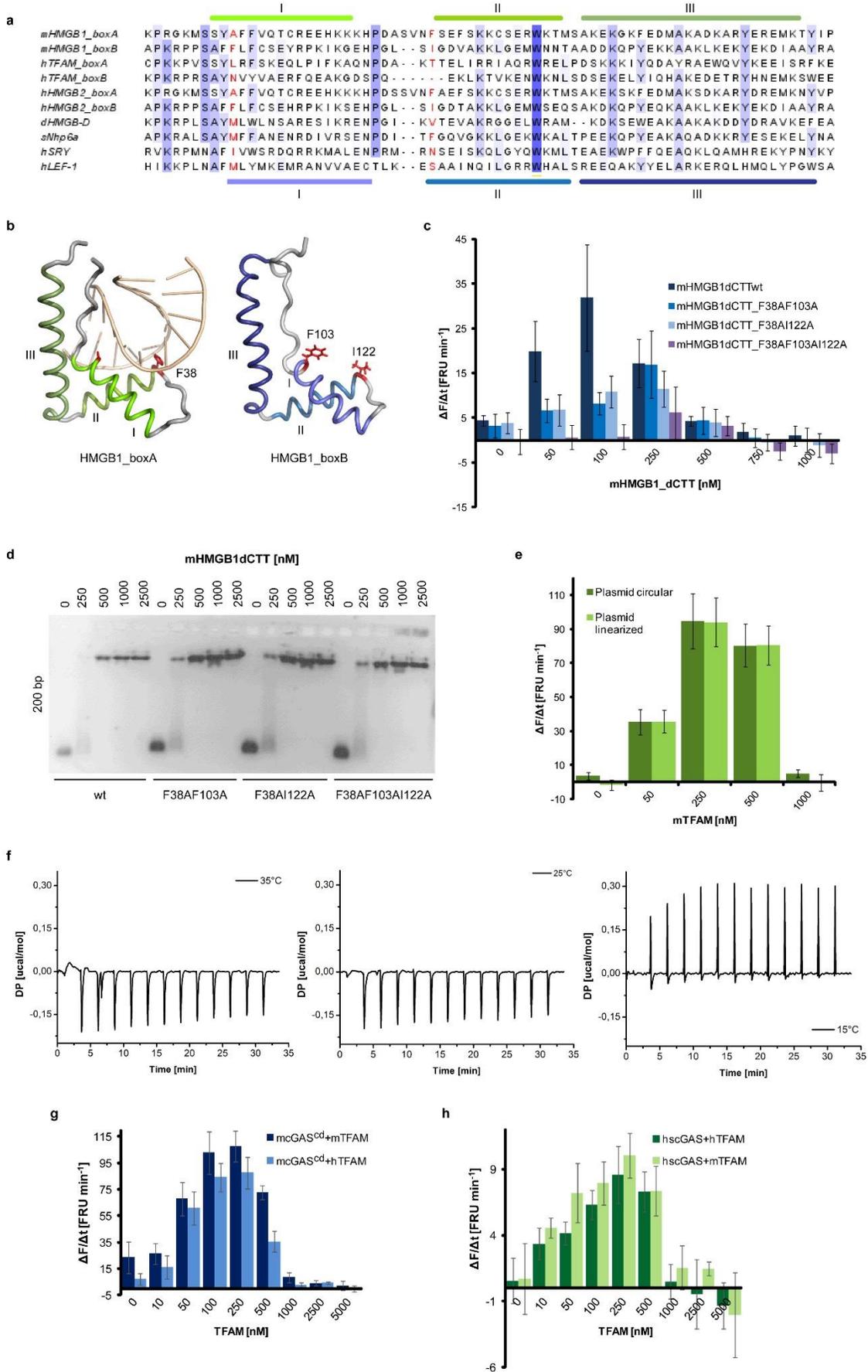
Extended Data Figure 8. HMGB proteins activate cGAS through DNA bending. a,

Sequence alignment of HMG boxes from different HMG proteins. Positions of intercalating

residues, responsible for the DNA-bending activity (red), are shown with respect to their location

within helices I, II, or III of HMGB1_boxA (green lines) and HMGB1_boxB (blue lines). **b**,

Structures of box A in complex with DNA (PDB accession number 4QR9) and box B (PDB

accession number 1HME) of rat HMGB1 with indicated intercalating residues (red). **c**, mcGAS^{cd}

activity measured by the rate of fATP incorporation into fGAMP (see Extended Data Fig. 2) in

the presence of increasing concentration of mHMGB1dCTT WT and intercalating residues

mutants. Double mutations (F38A in box A, and F103A or I122A in box B) or triple mutants

(F38A in box A, and both F103A and I122A in box B) were used. Mean values of initial cGAS

reaction rates ($\Delta F/\Delta t$) are plotted against increasing concentrations of mHMGB1dCTT $\pm$ s.d.,

$n = 7$. **d**, Electrophoretic mobility shift assay of mHMGB1dCTT WT and point mutants with

2.6 ng μl^{-1} (~200 nM binding sites) 200 bp DNA. Introduced mutations do not reduce DNA

binding of mHMGB1 mutants under conditions used for cGAS activity assays (**c**). **e**, mcGAS^{cd}

activity measured by the rate of fATP incorporation into fGAMP (see Extended Data Fig. 2) in

the presence of mTFAM with circular or linearized plasmid. Mean values of initial cGAS

reaction rates ($\Delta F/\Delta t$) are plotted against increasing concentrations of mTFAM $\pm$ s.d., $n = 3$. **f**,

ITC of hcGAS with hTFAM at 35, 25, and 15 °C. Power differential (DP) is plotted against time

of the experiment. hTFAM (~530 μM) was titrated to ~50 μM hcGAS. No binding is observed,

indicating $K_d > 100$ μM. **g**, **h**, cGAS activity measured by the rate of fATP incorporation into

fGAMP (see Extended Data Fig. 2) in the presence of increasing TFAM concentrations. Mean

values of initial cGAS reaction rates ($\Delta F/\Delta t$) are plotted against increasing concentrations of

TFAM $\pm$ s.d., $n = 3$ or 4. mcGAS^{cd} (**g**) and hcGAS (**h**) are activated by both human and mouse

TFAM.
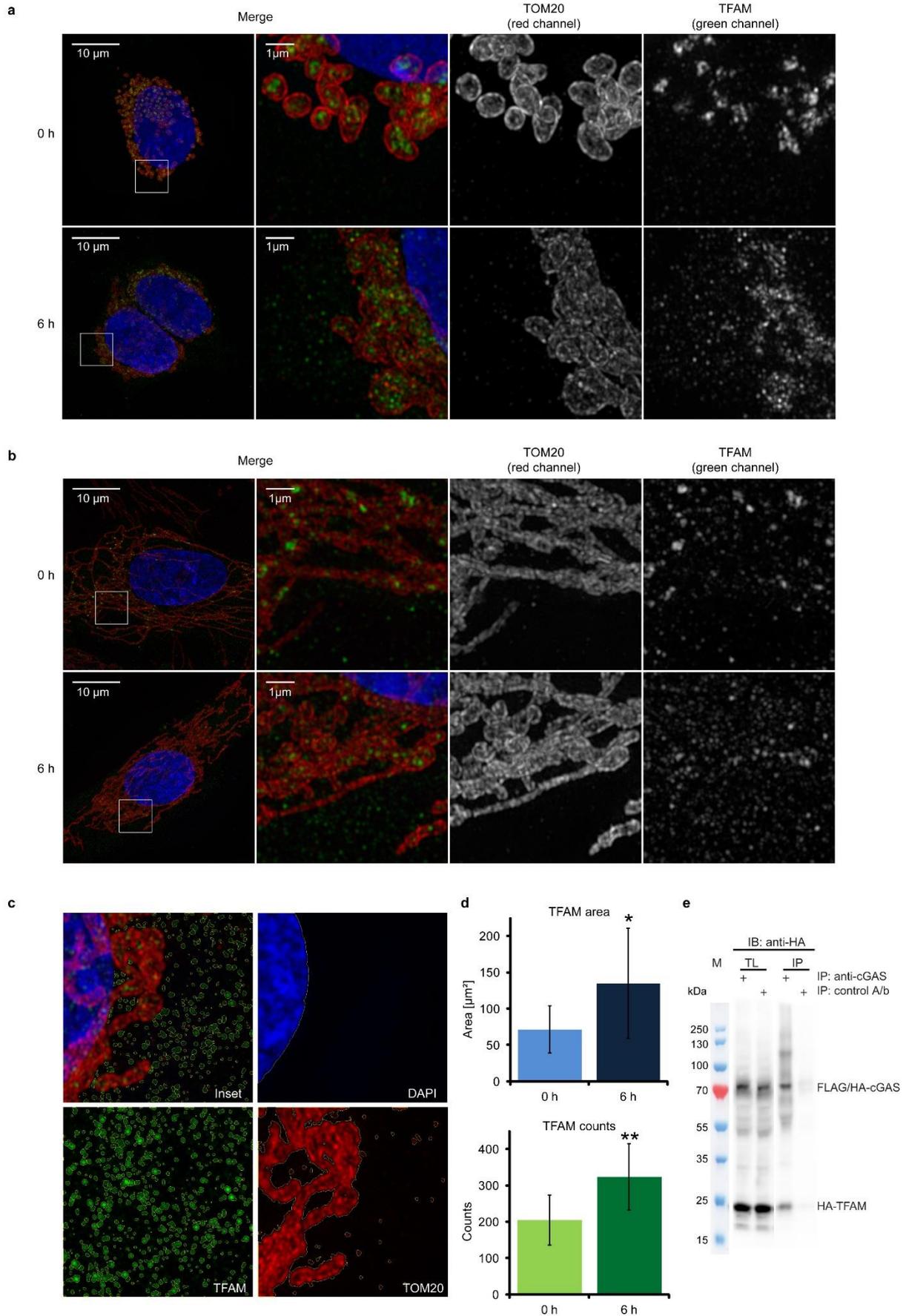
Extended Data Figure 9. TFAM relocalizes from mitochondria into cytosol during mitochondrial stress. a, **b**, TFAM (green) localization upon mitochondrial stress induction. Cells were incubated with caspase and Bcl-2 inhibitors for indicated time and stained for super-resolution fluorescence microscopy (3D SIM). Mitochondria were visualized by TOM20 (mitochondrial import receptor subunit, red) staining, cell nuclei by DAPI (blue). **a**, TFAM cytosolic localization under mitochondrial stress conditions in HEK293T cells. **b**, TFAM decondensation in mitochondria and its leakage into cytoplasm in human primary fibroblasts in response to mitochondrial stress. **c**, Example of statistical analysis in human primary fibroblasts (**b**). Cells were incubated with caspase and Bcl-2 inhibitors for 6 h and stained for super-resolution fluorescence microscopy (3D SIM). Mitochondria were visualized by TOM20 (mitochondrial import receptor subunit) staining, cell nuclei by DAPI. Areas stained for TFAM (green), TOM20 (red), and nucleus (blue) are defined with yellow line. Cytosolic TFAM spots are depicted on the top left view. **d**, Comparison of cytosolic TFAM signal in control cells (0 h) and under mitochondrial stress conditions (6 h) (**b, c**). Plotted are mean values of cytosolic TFAM signal calculated as area (top) or number of spots (bottom) ± s.d. Three cells of each type and three segments per cell were used ($n = 9$). Single asterisk indicates a statistically significant difference ($P < 0.05$, $P = 0.0267$), double asterisk indicates a statistically significant difference ($P < 0.01$, $P = 0.0050$), compared with control. The TFAM signal in the cytosol of control cells (0 h) represents background signal. **e**, Co-immunopurification of cGAS and TFAM. Cell lysates with overexpressed Flag/HA-cGAS and HA-TFAM were incubated with anti-hcGAS or with a control antibody (A/b) and the proteins in total lysates (TL) and elution fractions (IP) were visualized by immunoblotting (IB).
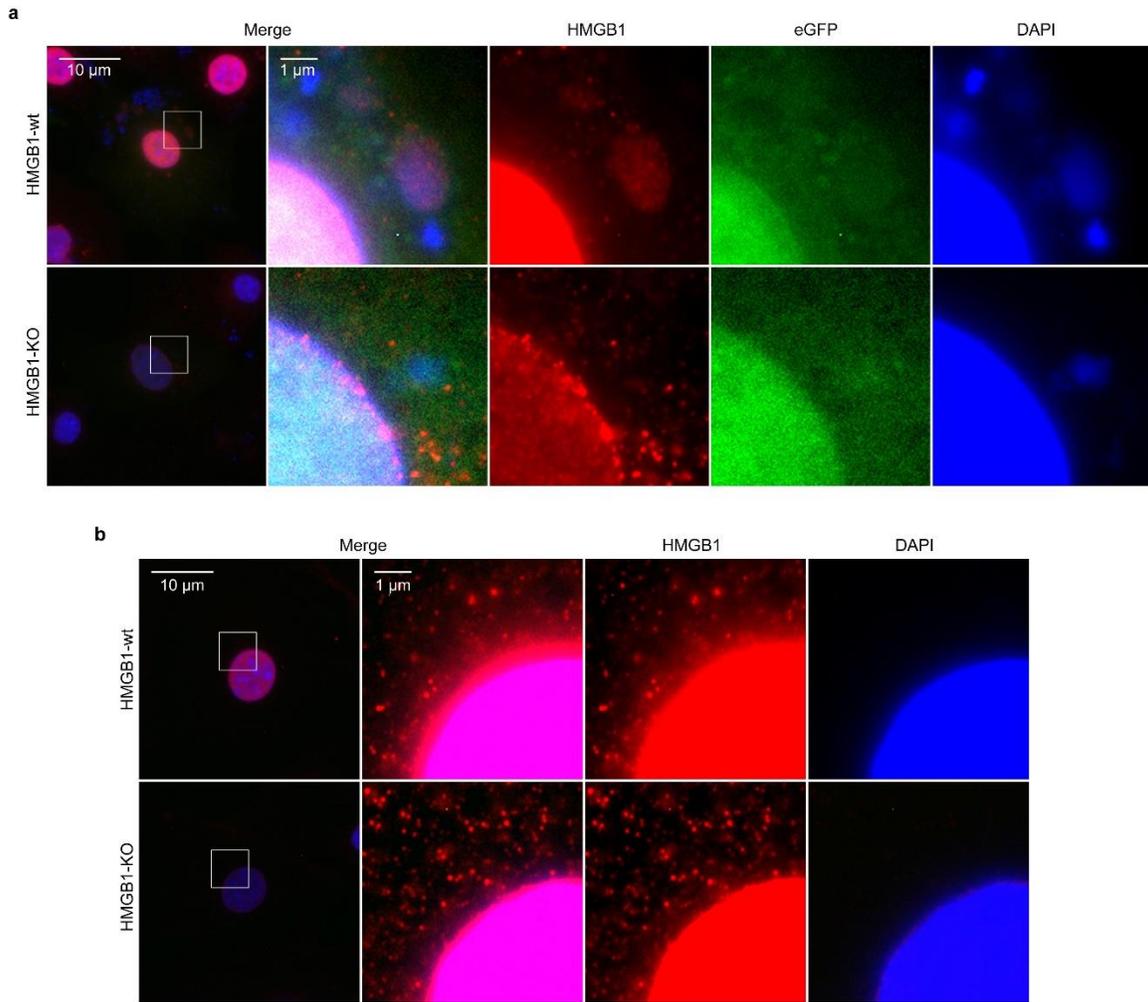
Extended Data Figure 10. HMGB1 but not eGFP co-localizes with cytosolic DNA. a, **b**, HMGB1 (red), eGFP (green), and DNA (blue) localization in WT or HMGB1-KO MEFs. DNA was stained by DAPI (blue). Images were obtained by wide-field fluorescence microscopy (pDV). Fluorescent signal within enlarged images is enhanced until saturation of the nuclear signal to better visualize cytoplasmic structures. **a**, Cells were transfected with eGFP expression construct. HMGB1 co-localizes with DAPI-staining in WT but not HMGB1-KO cells. eGFP does not co-localize with DNA or HMGB1. **b**, In non-transfected cells, HMGB1 and DAPI stainings are present only in the nucleus. The HMGB1 staining in the cytosol of non-transfected and HMGB1-KO cells represents background signals that become visible with the artificial signal amplification beyond saturation, as shown in the zoom-in images.

2. Sequence-specific activation of the DNA sensor cGAS by Y-form DNA structures as found in primary HIV-1 cDNA

Summary

In this publication reverse-transcribed HIV-1 negative-strand ssDNA is reported as the pathogen pattern recognized by major cytosolic DNA sensor cGAS and specific short dsDNA constructs with single-stranded guanine-overhangs are defined as the minimal cGAS recognition motives. Apart from cytosolic DNA emerging from dsDNA viruses like HSV-1 cGAS senses a range of retroviruses including HIV-1, murine leukemia virus and simian immunodeficiency depending on their reverse transcription step and independently from their integration. However, the exact nature of retroviral cGAS ligands remained unknown. This work discovers stem-loop structures within negative-strand ssDNA transcribed from HIV-1 genome during the first stages of infection to stimulate cGAS activity and type I IFN production in human macrophages. Strikingly, previously known to recognize dsDNA sequence-unspecifically, cGAS activation by HIV-1 or endogenous retroelement ssDNA was found to strongly depend on unpaired guanine bases flanking stem-loop regions. Mutations of these bases abolished IFN production in infected peripheral blood mononuclear cells. Based on this observation, synthetic short dsDNA constructs with G-overhangs on both sides – Y-shaped DNAs (YSDs) - were designed and shown to potently activate cGAS *in vitro* and in human cells. Given that cGAS was previously shown to require at least 45 bp to recognize blunt-end dsDNA in cell lines, G-overhangs allow to overcome this dsDNA length limitation. This study shows that 20 bp dsDNA with flanking guanines is as potent cGAS activator, as plasmid or genomic DNA and that G-overhangs enhance cGAS-stimulating capacity of golden-standard 45 bp immunostimulatory DNA. Novel cGAS ligands discovered in this work were shown to activate cGAS independently of G-quadruplex formation and YSD aggregation and thus represent a new class of short highly-immunostimulatory motives recognized by cGAS in a sequence-dependent manner.

Author contribution

The author of this thesis participated in *in vitro* characterization of the cGAS complexes with various short guanine overhangs-containing dsDNA constructs – Y-shaped DNAs (YSDs). She assisted the authors of this publication with interpretation of the *in vitro* and cell-based assays and actively participated in discussion of possible mechanisms of cGAS specific activation by YSDs.

Sequence-specific activation of the DNA sensor cGAS by Y-form DNA structures as found in primary HIV-1 cDNA

Anna-Maria Herzner[1,11], **Cristina Amparo Hagmann**[1,10,11], **Marion Goldeck**[1], **Steven Wolter**[1], **Kirsten Kübler**[2], **Sabine Wittmann**[3], **Thomas Gramberg**[3], **Liudmila Andreeva**[4], **Karl-Peter Hopfner**[4], **Christina Mertens**[1], **Thomas Zillinger**[1,5], **Tengchuan Jin**[6], **Tsan Sam Xiao**[7], **Eva Bartok**[1], **Christoph Coch**[1], **Damian Ackermann**[8,10], **Veit Hornung**[9], **Janos Ludwig**[1], **Winfried Barchet**[1,5], **Gunther Hartmann**[1,11], and **Martin Schlee**[1,11]

[1]Institute of Clinical Chemistry and Clinical Pharmacology, University Hospital Bonn, Bonn, Germany

[2]Department of Obstetrics and Gynecology, Center for Integrated Oncology, University of Bonn, Bonn, Germany

[3]Institute of Clinical and Molecular Virology, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany

[4]Department Biochemistry, Gene Center, Ludwig-Maximilians University, Munich, Germany

[5]German Center of Infectious Disease, Cologne-Bonn, Germany

[6]School of Life Sciences, University of Science and Technology of China, Hefei, China

[7]Department of Pathology, Case Western Reserve University, Cleveland, Ohio, USA

[8]LIMES Institute, Chemical Biology, University of Bonn, Bonn, Germany

[9]Institute of Molecular Medicine, University Hospital, University of Bonn, Bonn, Germany

Abstract

Cytosolic DNA that emerges during infection with a retrovirus or DNA virus triggers antiviral type I interferon responses. So far, only double-stranded DNA (dsDNA) over 40 base pairs (bp) in length has been considered immunostimulatory. Here we found that unpaired DNA nucleotides

flanking short base-paired DNA stretches, as in stem-loop structures of single-stranded DNA (ssDNA) derived from human immunodeficiency virus type 1 (HIV-1), activated the type I interferon–inducing DNA sensor cGAS in a sequence-dependent manner. DNA structures containing unpaired guanosines flanking short (12- to 20-bp) dsDNA (Y-form DNA) were highly stimulatory and specifically enhanced the enzymatic activity of cGAS. Furthermore, we found that primary HIV-1 reverse transcripts represented the predominant viral cytosolic DNA species during early infection of macrophages and that these ssDNAs were highly immunostimulatory. Collectively, our study identifies unpaired guanosines in Y-form DNA as a highly active, minimal cGAS recognition motif that enables detection of HIV-1 ssDNA.

Sensing of nucleic acids is crucial to antiviral defense. Unlike pathogen-associated molecular patterns (PAMPs) of bacterial origin that are foreign to the host, nucleic acids are vital to both host and pathogen alike. Thus, receptors that are part of the innate immune system recognize foreign genetic material through its unusual localization or structural features or modifications.

In the endolysosome of some cells of the immune system, Toll-like receptor 9 (TLR9) 'preferentially' detects DNA containing CpG dinucleotides[1–3]. In the cytosol, recognition of DNA triggers the secretion of both interferon-α (IFN-α) and IFN-β (collectively called 'IFN-α/β' here) and proteolytically activated interleukin 1β (IL-1β). Sensing of DNA by the inflammasome-forming receptor AIM2 is considered essential for the activation of IL-1β[4–6]. In contrast, several cytosolic DNA receptors that induce IFN-α/β have been proposed[7–15], although it is now broadly accepted that the IFN-α/β-inducing mitochondrial adaptor STING is downstream of this process[16,17].

Two candidate receptors upstream of STING, IFI16 and cGAS ('cyclic GMP-AMP (cGAMP) synthetase'), have been reported[18,19]. Involvement of IFI16 in the induction of IFN-α/β during infection with herpes simplex virus, human cytomegalovirus, human immunodeficiency virus (HIV) or *Listeria monocytogenes* has been reported[20]. However, no genetic proof confirming those findings has been provided so far. In contrast, cGAS-deficient mice and cells demonstrate clear deficits in their immune response to cytosolic DNA. Moreover, direct interaction of DNA with cGAS promotes synthesis of the second messenger cGAMP, which activates STING[19,21–29]. Furthermore, cGAS is reported to be essential for the immunodetection of DNA viruses[19,30,31] and retroviruses[27,32,33].

Several studies have defined cytosolic DNA-recognition motifs[11,34,35]. Double-stranded DNA (dsDNA) with any sequence longer than 24 base pairs (bp) is known to induce IFN-α/β in mouse cells, and a 45-bp dsDNA sequence (interferon-stimulatory DNA (ISD)) has been established as the 'gold standard' for the induction of IFN-α/β[11]. In human monocytes or the human monocytic cell line THP-1, length-dependent induction of IFN-α has a lower bound of 40–50 bp, with much less secretion of IFN-α in response to these short sequences[7,18]. Thus, it is assumed that recognition of DNA in the cytosol depends on duplex character and length but not sequence. However, it has been reported that lentivirus single-stranded DNA (ssDNA) stem-loop structures comprising far fewer than 40 bp can also induce IFN-α/β, although induction of IFN-α/β has been observed to depend on base-paired stretches of DNA within the stem-loop structures[36].

In this study, we delineated the recognition of a 181-nucleotide early HIV type 1 (HIV-1) reverse transcript ('strong-stop (−)-strand DNA' (sstDNA)) by the immune system and found that an isolated stem-loop–structured sequence induced cGAS-dependent activation of the immune system. Such recognition of the stem-loop structure depended on the presence of 3′ and 5′ stem-flanking sequences containing unpaired guanosines. We also found that increasing the guanosine content enhanced the induction of IFN-α/β. The addition of unpaired guanosines to otherwise inactive, blunt, 20-nucleotide DNA duplexes rendered these immunoactive at a level comparable to that of plasmid or genomic dsDNA. Strikingly, additional unpaired guanosine flanks even enhanced the activity of the prototypic blunt, 45-nucleotide ISD[11]. Furthermore, our data demonstrated the importance of these immunostimulatory Y-form DNA structures for the sensing of HIV-1 early reverse transcripts by the immune system in primary human macrophages as a model of infection with macrophage-tropic HIV-1. Collectively, our study documents a minimal immunostimulatory DNA motif that induces cGAS activity in a structure- and sequence-dependent manner and thereby enables the recognition of partially mismatched stem-loop structures as found in ssDNA of HIV-1.

RESULTS

Detection of unpaired guanosines in HIV cDNA stem loops

HIV-1 is detected via the cGAS-STING pathway[27,32,33]. It has been reported that ssDNA is the predominant cytosolic DNA species during the first 4 h of HIV infection, while dsDNA is found later in the nuclear or perinuclear fractions[37]. Replication of HIV-1 is initiated by reverse transcription of the first ~181 nucleotides of the HIV-1 RNA genome, primed by lysyl-tRNA (tRNA[Lys3]), which generates sstDNA (Fig. 1a). A published study has reported the cytosolic recognition of sstDNA-derived sequences as being dependent on the presence of the base-paired stretches[36]. Although they are double stranded, these base-paired structures still seemed too short (<40 bp) to elicit IFN-α/β in human monocytes[7] (**Supplementary Fig. 1a**). Thus, we attempted to elucidate if features other than base pairing could enable activation of a cellular response. We modeled the complete HIV-1 sstDNA complementary to bases 1–181 of the (+) strand using the mfold server (HIV-1 strain HXB2; NCBI accession code, AF033819) and identified three stem-loop structures (SL1–SL3; Fig. 1a, top). Then, we transfected chloroquine-treated human peripheral blood mononuclear cells (PBMCs) with DNA corresponding to each stem-loop structure isolated. In this experimental setting, IFN-α/β is derived from monocytes, while TLR-dependent secretion of IFN-α (for example, activation of TLR9 by the CpG ODN 2216) is blocked by chloroquine[7] (**Supplementary Fig. 1b**).

We found that among the stem-loop DNA structures, SL2 was immunostimulatory (Fig. 1b), despite containing only 21 bp in a stem sequence maximally 24–26 nucleotides in length. Notably, homologous sequences of another HIV-1 strain, as well as of a simian immunodeficiency virus strain, also demonstrated comparable immunostimulatory activity ('immunoactivity') (**Supplementary Fig. 1c**). Furthermore, the combination of SL2 and SL3 (116 nucleotides) had greater immunoactivity than SL2 alone (Fig. 1c), indicative of a synergistic immunostimulatory effect.

To determine the origin of the immunoactivity, we systematically altered the structures of the stem-loop structure to modify melting temperature, mismatches and bulges, and unpaired stretches flanking the stem. While lowering the melting temperature of the stem via the introduction of A:T base pairs resulted in less secretion of IFN-α, increasing the melting temperature via the introduction of G:C base pairs resulted in a slight but insignificant increase in the secretion of IFN-α (Fig. 1d). The removal of mismatches enhanced the secretion of IFN-α (Fig. 1e), which excluded the possibility of mismatches as a recognition motif. However, the most prominent change resulted from the removal of the unpaired stem-loop stretches, which abrogated the IFN-α signal and was not compensated by removal of the mismatches (Fig. 1f). This observation linked the adjacent, unpaired sequences to stem-loop–associated immunoactivity. Next we assessed whether such recognition was sequence dependent. The removal of guanosines from the unpaired stretches of the hairpin rendered it inactive, while increasing the guanosine content increased its immunoactivity (Fig. 1g). Moreover, the immunoactivation induced by the combined stem-loop structures SL2 plus SL3 was also largely dependent on the presence of unpaired guanosines (Fig. 1h). Native PAGE and melting-curve analysis confirmed that mutations that removed guanosines did not substantially change folding or complex formation (**Supplementary Fig. 1d**). As observed for wild-type stems, the mismatch–bulge-deficient stem-loop structure induced far less secretion of IFN-α when guanosines within the loop and 3′ and 5′ flanking stretches were substituted (Fig. 1i). Notably, the combination of additional guanosines in the unpaired regions and a perfect duplex in the stem led to an IFN-α/β response comparable to that observed for standard stimuli such as genomic DNA (Fig. 1j). Furthermore, isolated primary human monocytes and monocyte-derived macrophages, which represent the main targets of macrophage-tropic HIV-1 infection, recognized short hairpin structures, and again, IFN-α-induction depended on unpaired guanosines (**Supplementary Fig. 1e,f**. In conclusion, these data indicated that the unpaired regions of stem-loop structures within ssDNA substantially contributed to detection of the ssDNA and that guanosines were essential for this recognition mechanism.

Unpaired guanosines in Y-form DNA are a robust yet flexible PAMP

Next we investigated whether a closed loop of SL2 was essential for immunostimulation or if Y-form DNA junctions (the transition from dsDNA to ssDNA) were sufficient for immunoactivation. Opening the loop led to a slight yet insignificant reduction in the immunoactivity of the SL2 variants without mismatches (Fig. 2a), which showed that the unpaired stretches rather than the closed loop were necessary for recognition of the short stem-loop structures. Finally, simple, short DNA duplexes (20–21 bp) with various stem sequences (HIV derived, random non-palindromic, and palindromic) flanked by unpaired guanosine trimer (G_3) ends exhibited the same IFN-α/β-inducing activity as did the guanosine-enriched SL2 hairpin or genomic DNA with removal of mismatches (Fig. 2b). That finding prompted us to further study the flexibility of this recognition motif, which we called 'G_n-ended Y-form short DNA' (G_n-YSD, where 'n' indicates the number of unpaired guanosines at each 3′, and 5′, end).

In line with those observations, substitution of the terminal G_3 with adenosine (A_3-YSD), thymidine (T_3-YSD) or cytidine (C_3-YSD) almost completely abrogated the induction of

IFN-α (Fig. 2c). Furthermore, the stimulatory effect of the G_3 overhangs was neutralized by hybridization of G_3-YSD with the fully complementary C_3-YSD, but not by hybridization with the T_3-YSD control, since the latter left free G_3 ends available for recognition (Fig. 2c). Further experiments revealed that two G_3 overhangs at opposite ends of the DNA duplex (regardless of whether it was at the 3′ end or 5′ end) were sufficient for robust induction of IFN-α (Fig. 2d). Nevertheless, the sequence requirements for recognition of the unpaired overhangs demonstrated considerable flexibility; one guanosine within one of five unpaired bases at each 5′ and 3′ end of a DNA duplex was sufficient to initiate an IFN-α/β response (**Supplementary Fig. 2a**). While a decrease in the length of the core duplexes of G_3-YSD to 12 bp elicited a diminished but still substantial IFN-α response, an increase in its length did not enhance IFN-α induction (Fig. 2e). Notably, G_3 ends also further enhanced the already potent immunostimulatory activity of the prototypic blunt 45-nucleotide ISD[11] by approximately fourfold (**Supplementary Fig. 2b**). Mismatches or bulges of up to 4 bp in the dsDNA core sequence of G_3-YSD were tolerated, although this correlated with diminished immunoactivity (Fig. 2f).

We also observed specific recognition of G_3-YSD in purified human monocytes as well as in monocyte-derived macrophages, even though both cell types responded weakly to ISD (**Supplementary Fig. 2c,d**). Finally, G end–dependent recognition of Y-form DNA was not restricted to human cells, as immortalized mouse macrophages also showed a similar response to YSD (**Supplementary Fig. 2d**), although these cells also responded to 30-nucleotide blunt-ended dsDNA, as reported[11]. Together, these results demonstrated that unpaired guanosine extensions potently enhanced the immunoactivity of short DNA duplexes.

G-ended Y-form DNA is active as a monomeric duplex

To address whether variations in DNA stability and transfection efficiency might be responsible for the differences observed in the stimulatory activity of the short DNA duplexes[38], we transfected THP-1 cells with 5′-labeled stimulatory DNA (G_3-YSD; ISD) or non-stimulatory DNA (C_3-YSD; blunt-ended 26-nucleotide DNA). At 4 h and 8 h after transfection, we assessed their presence in cytosolic lysates by denaturing polyacrylamide-gel electrophoresis (PAGE) of equal lysate proportions and subsequent detection of label fluorescence. Here, G_3-YSD exhibited fluorescence intensity in THP-1 cytosolic cell lysates similar to or somewhat less than that of the non-stimulatory C_3-YSD or blunt-ended 26-nucleotide DNA (**Supplementary Fig. 3a**). Furthermore, *in vitro* digestion of the same unlabeled DNA stimuli with recombinant DNase TREX1, DNase I or DNase II[38] did not substantially reduce degradation (**Supplementary Fig. 3b,c**). Thus, we concluded that direct ligand-receptor interaction, not stability or translocation effects, was relevant for the induction of IFN-α/β by guanosine-ended Y-form DNA.

One characteristic feature of guanosine-rich DNA sequences is their ability to assemble into guanosine quadruplex ('G quadruplex') complexes ('G tetrad') (Fig. 3a). One possible explanation for the enhanced recognition of guanosine-ended YSD was the potential formation of long chains by such intermolecular G-quadruplex interactions. By native PAGE, high-molecular-weight bands were visible for G_4-YSD or G_5-YSD (Fig. 3c), which

Nat Immunol. Author manuscript; available in PMC 2015 December 03.

-99-

indicated G quadruplex–mediated oligomerization (sequences and structures, Fig. 3b), while G$_2$-YSD or G$_3$-YSD migrated as a single, low-molecular-weight band (Fig. 3c). Since strong electric fields may interrupt weak interactions during electrophoresis, we investigated the molecular interactions in solution. Anti-parallel G quadruplexes are arranged chirally and are optically active at 295 nm (ref. 39). Their formation can be detected by circular dichroism spectroscopy, which measures the difference in the absorption of left-circularly polarized light and that of right-circularly polarized light (ellipticity) of the analyte at 30–70 °C. G$_4$-YSD and G$_5$-YSD exhibited a substantial positive ellipticity at 295 nm that was lost at high temperatures (70 °C) (Fig. 3d), which indicated G-quadruplex formation at physiological temperatures. In contrast, G$_2$-YSD or G$_3$-YSD, like the negative controls A$_3$-YSD and a blunt-ended 30-nucleotide DNA, did not show considerable optical activity at 295 nm (Fig. 3d), which indicated that G$_2$-YSD and G$_3$-YSD were present as monomeric duplexes. Notably, the quadruplex-forming YSDs did not induce larger amounts of IFN-α than the prototype G$_3$-YSDs did (Fig. 3e). To exclude the intracellular formation of G quadruplexes, we made use of 26-nucleotide G$_2$-YSD, which was as active as 20-nucleotide G$_3$-YSD (Fig. 3e). Here, we replaced the terminal guanosines with 7-deaza-guanosine or inosine. Substitution of nitrogen N7 of the guanine by a C-H group (7-deaza-guanosine) and the missing amine group at carbon C3 of inosine disables formation of the hydrogen bonds essential for the G-quadruplex arrangement[40]. Substitution with 7-deaza-guanosine as well as substitution with inosine did not substantially alter the immunostimulatory activity (Fig. 3f). We concluded that intermolecular G-quadruplex interactions were not involved in the recognition of G-ended YSD but that the monomeric duplex was detected and induced IFN-α/β.

G-ended Y-form DNA activates cGAS via direct interaction

Of all human cell lines we tested, only the human monocytic cell line THP-1 was responsive to G$_3$-YSD (data not shown), in line with what has been described for dsDNA with a random base composition[41]. The RNA helicase RIG-I has been linked to the detection of long AT-rich DNA[7,8]. To exclude the possibility of involvement of signaling via RIG-I and the signaling adaptor MAVS in the recognition of YSD in this cell line, we performed RNA-mediated interference (RNAi) with small interfering (siRNA) targeting MAVS or STING. Knockdown of STING substantially repressed the induction of IFN-α/β by G$_3$-YSD and plasmid DNA, but knockdown of MAVS did not; instead, knockdown of MAVS repressed the response to the RIG-I ligand 3P-dsRNA (triphosphorylated double-stranded RNA) but not to any of the DNA ligands (Fig. 4a and **Supplementary Fig. 4a**). These results demonstrated that G$_3$-YSD indeed induced secretion of IFN-α/β via a STING-dependent pathway.

To identify the receptor responsible for the recognition of G$_3$-YSD, we expressed Flag-tagged candidate receptors in HEK293T human embryonic kidney cells and performed bead-coupled co-precipitation experiments with immobilized DNA ligands (Fig. 4b–d). IFI16, ZBP1, DDX41 and cGAS have been described as IFN-α/β-inducing cytosolic receptors for dsDNA in fibroblasts, monocytes, macrophages or conventional dendritic cells[18,19,29,42,43]. Among those candidates, IFI16 and cGAS bound to both G$_3$-YSD and C$_3$-YSD, while ZBP1, DDX41 and the negative control RIG-I did not detectably precipitate

together with YSD (Fig. 4c,d). Both IFI16 and cGAS bound to long (79-nucleotide) blunt-ended DNA, whereas short (26-nucleotide) blunt-ended DNA did not precipitate these candidates (Fig. 4d). Only Ku80, a nonspecific DNA-binding control, did not show substantial 'preference' for any DNA structure or length (Fig. 4d). Of note, IFI16 showed a slight (threefold) 'preference' for the immunostimulatory YSD (G_3-YSD over C_3-YSD; Fig. 4c,d and **Supplementary Fig. 4b**).

We performed RNAi experiments with siRNA in THP-1 cells to determine the contribution of IFI16 and cGAS to the secretion of IFN-α/β after stimulation with dsDNA. RNAi of cGAS almost completely abrogated the induction of IFN-α/β by dsDNA species, including G_3-YSD (Fig. 4e and **Supplementary Fig. 4c**). In contrast, RNAi of IFI16 did not substantially suppress the induction of IFN-α/β by dsDNA (**Supplementary Fig. 4d**). Moreover, RNAi of MAVS again inhibited only RIG-I-dependent, 3P-dsRNA-induced secretion of IFN-α/β, not the G_3-YSD-dependent response (Fig. 4e and **Supplementary Fig. 4c**). These results demonstrated that G_3-YSD-induced secretion of IFN-α/β depended on the cGAS-STING-axis.

To investigate whether guanosines were directly detected by cGAS, we performed *in vitro*–activation assays of truncated recombinant cGAS (amino acids 155–522). In the presence of a ligand, cGAS catalyzes the conversion of ATP and GTP to the non-canonical cyclic dinucleotide cGAMP. We used untagged G_3-YSD, C_3-YSD and 26-nucleotide blunt DNA, as well as 45-nucleotide ISD as a positive control[11]: As expected, in the presence of ISD, ATP and GTP were converted to cGAMP, as visualized by the corresponding ultraviolet irradiation–absorption maxima in the HPLC chromatogram (Fig. 4f). While G_3-YSD activated cGAS to the same extent as ISD did, C_3-YSD and the blunt-ended 26-nucleotide DNA resulted in less conversion (Fig. 4f). Thus, we concluded that G-rich overhangs of short duplexes specifically and directly enhanced the activation of cGAS, which led to STING-dependent secretion of IFN-α/β.

The interferon response correlates with viral ssDNA content

During the life cycle of lentiviruses, reverse transcription of (−) ssDNA is followed by replication of second strand ((+)-strand) DNA, which leads to the generation of dsDNA[37]. A published study has shown detection of HIV-1 ssDNA predominantly in the cytosol, with dsDNA in the nuclear and perinuclear fraction[37]. Thus, we hypothesized that recognition of ssDNA by cGAS might be especially important during early infection. To investigate the kinetics of ssDNA formation, we developed a strand-specific quantitative PCR protocol to quantitatively detect (−)-strand DNA and (+)-strand DNA in infected cells (**Supplementary Fig. 5a**). To determine if dsDNA or ssDNA is crucial for cGAS stimulation, we generated replication-deficient HIV-1-derived lentiviral particles containing a mutant reverse transcriptase, RT(N265D) (**Supplementary Fig. 5b**), reported to be impaired in its use of DNA as template during replication[44]. Thus, we expected this mutant to produce little or no dsDNA, represented by the presence of (+)-strand DNA. We infected THP-1 cells for 4 h with these HIV-1 particles during inhibition of SAMHD1 (which controls reverse transcription) by the nonstructural protein Vpx, delivered by virus-like-particles, and analyzed cytosol-enriched fractions of these cells. To confirm that the induction of

interferon-stimulated genes by either wild-type particles or RT(N265D) particles was cGAS dependent, we infected wild-type THP-1 cells as well as cGAS-deficient THP-1 cells[45]. As reported[37], in both wild-type and cGAS-deficient cells, we identified (−)-strand DNA as the predominant species in the cytosol (Fig. 5a). Compared with strand synthesis induced by wild-type RT particles, for the RT(N265D) particles, (+)strand synthesis was impaired but (−)-strand synthesis was not (Fig. 5a). Despite the diminished induction of (+)-strand synthesis by RT(N265D), induction of the interferon-stimulated gene *IFIT2* by RT(N265D) particles was not diminished compared with its induction by wild-type particles but instead was slightly increased, in the responsive wild-type THP-1 cells (Fig. 5b). In addition, the induction of *IFIT2* by both strains, as well as by the transfection of genomic DNA, but not by the transfection of 3P-dsRNA, was lost in cGAS-deficient cells (Fig. 5b,c), which demonstrated cGAS-mediated induction of an interferon-stimulated gene.

To verify that effect in primary cells, we also infected monocyte-derived macrophages, since these are particularly physiologically relevant for HIV infection. Again, we identified (−)ssDNA as the predominant DNA species in the cytosol-enriched fraction, with an even greater portion of (−)ssDNA than of (+)DNA (Fig. 5d); this confirmed both the presence and excess of ssDNA. While macrophages infected with RT(N265D) particles had slightly more (−)ssDNA than did cells infected with wild-type particles (Fig. 5d), the abundance of (+)-strand DNA was reduced in the former cells, which led to a significantly lower ratio of (+) strand to (−) strand (Fig. 5e). Notably, the induction of both *IFNB1* and *IFIT2* was greater in cells infected with the mutant RT(N265D) particles than in cells infected with wild-type particles (Fig. 5f,g), which demonstrated a correlation with the presence of the (−) strand rather than the (+) strand for the induction of IFN-β and interferon-stimulated genes and linked ssDNA to the recognition of early HIV-1 infection. Together these data offered evidence that (−)-strand ssDNA was the predominant DNA species detected by cGAS, at least during early stages of HIV-1 infection, and that the impairment of (+)-strand synthesis did not hinder the activation of cGAS. Furthermore, these data emphasized the importance of the detection of ssDNA by cGAS *in vivo*.

Long ssDNA is highly immunostimulatory

Since the probability of forming cGAS-stimulating YSD structures increases with ssDNA length, we hypothesized that the induction of IFN-α/β might correlate with length of the HIV (−)ssDNA strand. Thus, we established a protocol to enrich very pure ssDNA from PCR products comprising the first 180 or 381 nucleotides of the (−) strand emerging during infection with HIV-1 strain HXB2. Here, we combined digestion of the (+) strand with 5′ phosphate–dependent λ-exonuclease with purification via biotin-mediated immobilization and digestion with a dsDNA-specific restriction enzyme (**Supplementary Fig. 6a**). Using this protocol, we produced very pure ssDNA (100-fold excess of (−) strand over (+) strand), with each product appearing as one prominent band by PAGE (**Supplementary Fig. 6b**). Stimulation of monocyte-derived macrophages with these ssDNA species revealed a correlation between length and immunoactivity that culminated in the induction of IFN-α by the 381-nucleotide species that was comparable to that induced by G_3-YSD and was even greater than that induced by human genomic DNA (Fig. 6a). This indicated that long ssDNAs emerging during HIV-1 infection could be as immunoactive as comparable dsDNA

sequences. In conclusion, these data supported the hypothesis that the detection of HIV-1 depends on the recognition of ssDNA.

cGAS senses unpaired G in human retroelement sequences

We hypothesized that analogous to the detection of lentivirus, the recognition of ssDNA or YSD might be involved in the sensing of endogenous ssDNA derived from endogenous retroelements. These can accumulate and activate cGAS when the function of the cytosolic DNase TREX1 is impaired[46,47], which leads to interferon-driven autoimmune diseases such as Aicardi-Goutieres syndrome. We assessed the immunstimulatory capacity of two long hairpin structures derived from endogenous retroelements (HERV-E and ERV3.1). Indeed, the structures with more unpaired guanosines (ERV3.1) elicited a mild IFN-α response in human PBMCs (Fig. 6b). This activity was abrogated upon mutation of guanosine in the unpaired stretches (Fig. 6b). Thus, our findings indicated that recognition of a G-YSD motif was also a possible mechanism for enabling the detection of single-stranded endogenous retroelements.

DISCUSSION

Length-dependent but not sequence-dependent recognition of cytosolic base-paired DNA has been reported[7,11,19,29]. By investigating the recognition of HIV-1 ssDNA stem-loop structures by the immune system, we have identified a hitherto-unknown cGAS recognition motif consisting of unpaired guanosines in Y-form DNA junctions (G-YSD). This motif rendered short (<20-bp), otherwise inert dsDNA highly immunostimulatory and was of particular importance in ligands compromised by limited duplex length, mismatches or bulges, such as the short stem-loop structures in single-stranded HIV-1 primary reverse transcripts. The predominance of cytosolic (−)-strand DNA and not (+)-strand DNA and its close correlation with the observed immunostimulation indicated that ssDNA might represent the main stimulus during early HIV-1 infection in macrophages, which are an important target of HIV.

Unpaired guanosines in YSD promoted high cGAMP-synthase activity of purified cGAS. Since we were able to exclude the possibility of effects of G-quadruplex self-assembly, stability or transfection, guanosines most probably activated cGAS itself. Although sequence-independent activation of cGAS by various DNA types (plasmids, PCR products and genomic DNA) would indicate internal binding of DNA duplexes, crystal-structure studies are less conclusive at present[28,29]. Here, two cGAS molecules formed a tight 2:2 conformation with two DNA duplexes, which made internal binding to long DNA sterically unfavorable, since strand assembly would lead to duplex collision. Thus, binding of cGAS to DNA ends seems more plausible. Alternatively, a specific guanosine-cGAS (base-protein) interaction might be a plausible mechanism. Notably, although the *in vitro* cGAMP-synthesis assay with cGAS truncated at the amino terminus generally correlated with activation in cell-based assays, the sequence-specific effect was much more pronounced in the cellular 'read-out'. Therefore, we cannot exclude the possibility of a contribution by the cGAS amino terminus or a co-receptor (for example, PQBP1)[48] to the recognition of G-YSD.

Lentiviral genomic (+)-strand RNA is rich in adenosine with a low proportion of cytidine, especially in unpaired regions[49–51], corresponding to a low guanosine content in (−)ssDNA. Several mechanisms for selective pressure toward the elimination of cytidine in HIV-1 RNA have been proposed, including mutational bias of reverse transcriptase, avoidance of transcriptional silencing by methylation of CpG motifs, and detection by cytidine deaminases of the APOBEC family or TLR9 (ref. 50). Our results might also link the recognition of ssDNA by cGAS to this elimination process. Notably, cGAS is the main sensor of HIV-1 in myeloid cells, and a rapid and sustained IFN-α/β response in dendritic cells has been linked to an effective HIV-1-specific CD8[+] T cell response[52]. Notably, the proportion of cytidine in the conserved, highly-structured HIV-1 long terminal repeats appears to be normal[50]. Here, RNA structure–dependent functions might limit mutagenesis and thereby preserve these cGAS recognition sites.

Apart from the detection of retroviruses, the YSD motif–dependent recognition of ssDNA by cGAS might be responsible for the elevated levels of IFN-α/β observed in Aicardi-Goutieres syndrome and other auto-inflammatory disorders with impaired function of TREX1. TREX1 degrades endogenous cytosolic DNA derived mainly from single-stranded endogenous retroelements, including the long terminal repeat segments of endogenous retroviruses[46]. TREX1 deficiency thus leads to the accumulation of cytosolic DNA and the induction of a STING-cGAS–dependent IFN-α/β response[47,53]. We assessed two ssDNA stem-loop structures derived from endogenous retroviral long terminal repeats and found that one induced IFN-α/β via the detection of unpaired guanosines. Thus, endogenous ssDNA might also be sensed via the YSD-recognition motif.

In conclusion, cGAS can sense very short dsDNA sequences containing guanosines in Y-form DNA (YSD recognition), and this PAMP enables the detection of ssDNA in the reverse transcripts of exogenous or endogenous retroviruses. Thus, we propose two distinct mechanisms for the cGAS-mediated recognition of DNA: While a base-paired stem of 12 bp or more is a necessary prerequisite, the additional presence of either a stem length of at least 40 bp (ISD recognition) or guanosines in Y-form junctions (YSD recognition) is sufficient for the activation of cGAS. Nevertheless, it remains to be determined whether the recognition of ISD and YSD results in similar or distinct modes of cGAS activation.

ONLINE METHODS

Ethics statement

The studies of human PBMCs were approved by the local ethics committee (Ethikkommission der Medizinischen Fakultät Bonn) according to guidelines of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use and Good Clinical Practice. Written informed consent was provided by voluntary blood donors.

Stimulatory nucleic acids

Synthetic DNA was obtained from Metabion, IDT, Invitrogen or EUROGENTEC (deazaguanosine-DNA). pDNA (pBluescript) was isolated from transformed *Escherichia coli* K12 with a PureLink HiPure Plasmid Filter Maxiprep Kit (Life Technologies).

Poly(dAdT) was obtained from Sigma-Aldrich. 3P-dsRNA (*in vitro* transcript 4 (IVT4)) was prepared as described[9]. DNA was hybridized in 25 mM Tris-HCl, pH 7.4 and 25 mM NaCl. DNA was heated to 72 °C for 5 min and was then cooled to 18 °C at a cooling rate of 1 °C/ min. DNA with a melting temperature of >70 °C was heated to 95 °C and then was cooled to 18 °C at a cooling rate of 1 °C/min.

Cell culture and stimulation

Cell lines were cultured in RPMI-1640 medium (THP-1 cells) or DMEM (HEK293T cells (immortalized mouse macrophages[54]); Gibco), supplemented with 10% FCS and penicillin/ streptomycin (standard medium; Gibco). Human PBMCs were isolated from buffy coats derived from human blood from healthy, voluntary donors by Ficoll-Hypaque density-gradient centrifugation (Biochrom). Monocytes were generated from PBMCs by magnetic cell sorting of $CD14^+$ cells (Miltenyi Biotec). Human macrophages were differentiated *in vitro* from $CD14^+$ monocytes in the presence of recombinant cytokine GM-CSF (50 U/ml) in standard RPMI-1640 medium. The generation of $CD14^{hi}CD80^{hi}CD163^-$ M1-polarized macrophages was confirmed by flow cytometry after 5 d (BD Biosciences).

For stimulation experiments, 4×10^5 PBMCs, 2×10^5 monocytes 1.5×10^5 macrophages or 6×10^4 THP-1 cells per well were seeded into 96-well plates in standard RPMI-1640 medium. For prevention of TLR9 activation, PBMCs were incubated in 2.5 μg/ml chloroqine 30 min before stimulation. Cells were stimulated for 20 h (PBMCs, monocytes or THP-1 cells) or 36 h (monocyte-derived macrophages). Nucleic acid stimuli were used at a final concentration of 0.8 μg/ml (PBMCs, monocytes or monocyte-derived macrophages) or 1.3 μg/ml (THP-1 cells; immortalized mouse macrophages). RNA and DNA stimuli were used in complex with Lipofectamine 2000 (Invitrogen). Secretion of IFN-α was measured with ELISA kits supplied by eBioscience. Human IFN-α/β activity was quantified by incubation of the IFN-α/β-sensing reporter cell line HEK-Blue IFN-α/β (Invivogen) with supernatants of stimulated cells. Upregulation of *Ifnb1* mRNA in immortalized mouse macrophages was detected by quantitative PCR 6 h after stimulation. The induction of *IFNB1* or *IFIT2* mRNA in infection experiments was measured 4 h after infection or transfection (described below in the subsection entitled "Infection of monocyte-derived macrophages and THP-1 cells").

PAGE of nucleic acids

Samples were separated by electrophoresis through native polyacrylamide gels (15%) in KCl-supplemented TBE buffer (45 mM Tris-borate, pH 8.0, 1 mM EDTA and 20 mM KCl) at 12.5 V/cm and 4 °C. DNA was annealed in a buffer of 100 mM KCl, 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA; 1 μg DNA was loaded per lane and, after electrophoresis, DNA was stained with methylene blue. For denaturing PAGE, nucleic acids or proteinase K–digested lysate (described below in the subsection entitled "Detection of cytosolic 5′ IRD700–labeled DNA") were denatured for 10 min at 65 °C (long ssDNA) or for 5 min at 90 °C (short DNA duplexes) in an excess of formamide supplied with Orange G and then were 'snap-cooled' on ice. Samples were loaded onto a TBE-buffered 6% polyacrylamide gel (long ssDNA) or 15% polyacrylamide gel (short DNA duplexes) containing 48% (wt/ vol) urea and the gel was run at 12.5 V/cm and 18–20 °C. DNA was detected by infrared

fluorescence of the 5′-linked dye IRD-700 (described below) or by staining with GelRed or GelGreen (Biotium).

Circular dichroism spectroscopy

Circular-dichroism spectroscopy was recorded on a Jobin–Yvon Dichrograph model CD6 in a 1-cm quartz cell. For measurement of circular-dichroism, DNA was resolved in NaCl buffer (150 mM NaCl and 10 mM TRIS-HCl, pH 8) and was subsequently heated to 70 °C.

Detection of cytosolic 5′ IRD-700–labeled DNA

THP-1 cells seeded into 24-well plates (3×10^5 cells per well) were transfected with 1 µg hybridized DNA labeled at the 5′ end with IRD-700 (Metabion), followed by incubation for 4 h or 8 h. The cells were washed with PBS and then were lysed for 5 min on ice with 40 µl NP40 buffer (25 mM Tris-HCl, pH 7.4, 1 mM MgCl2, 5 mM KCl, 0.5% NP-40 and 10 mM EDTA). Nuclei were precipitated by centrifugation at $300g$ (5 min at 4 °C), then supernatants were cleared by another centrifugation step at $5,000g$ (5 min at 4 °C). Lysates (5 µl each) were mixed with 5 µl proteinase K mix (1% (wt/vol) SDS, 20 mM EDTA, 60 mM Tris-HCl, pH 6.8, and 0.2 mg/ml proteinase K), followed by digestion for 20 min at 50 °C, then by heat inactivation for 10 min at 90 °C. Samples were separated by electrophoresis through denaturing gels, and IRD-700 fluorescence was detected with an Odyssey Imaging System (Li-COR). Loading controls were prepared as follows: hybridized DNA (2.5 ng) was mixed with 5 µl proteinase K mix and then was incubated, denatured with formamide and separated by electrophoresis as described above.

DNA interaction assays

FLAG-tagged proteins were expressed by transient transfection into HEK293T cells (seeded in six-well plates) via $CaPO_4$ complexes. After 24 h, the cells were mechanically detached and then were resuspended in ice-cold PBS, precipitated and washed. Cells were lysed in DNA lysis buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 10% (vol/vol) glycerol, 0.3% (vol/vol) Triton X-100, 1× Phosphatase-Inhibitor-Cocktail and 1× complete protease inhibitor (Roche); 100 µl per six wells), then nuclei were precipitated for 30 min at $20,000g$ and 4 °C. Hybridized DNA tagged with biotin at the 3′ end (5 µg ssDNA, or 10 µg dsDNA) was incubated for 30 min at 18–20 °C with 30 µl supernatant. Neutravidin beads (10 µl per reaction; Thermo Scientific) were washed twice with DNA lysis buffer and were loaded onto polyethylene filter columns (0.8 ml; Thermo Scientific). DNA-lysate mixes were added, followed by incubation for further 2 h. After that, the columns were opened at the bottom and the supernatant was collected in a 1.5 ml reaction tube. The beads were washed twice with DNA lysis buffer and once with DNA lysis buffer with a higher NaCl concentration (300 mM). Bound proteins were eluted for 30 min at 40 °C with Lämmli buffer (40 µl per reaction). A volume of 15 µl eluate or 3 µl cell lysate was assessed by standard immunoblot analysis with the monoclonal antibody ANTI-FLAG M2-Peroxidase (Sigma). The interaction assay analyzing the binding of the HIN domain of IFI16 to DNA in solution was performed as described[55].

cGAS *in vitro* activity assay

For *in vitro* reactions of cGAS in the presence of varying dsDNA stimuli, 2 μM recombinant human cGAS (truncated version: residues 155–522) was mixed with 88 ng/μl hybridized DNA, 0.1 mM ATP and 0.1 mM GTP in buffer A (100 mM NaCl, 40 mM Tris, pH 7.5, and 10 mM MgCl$_2$). After 90 min of incubation at 37 °C, the reaction mixture was analyzed by reverse-phase HPLC. Samples were prepared in 300 mM triethylammonium acetate, then were applied to a 4.1- × 250-mm PRP-1 column (Hamilton) and separated for 18 min at a flow rate of 1 ml/min in an isocratic gradient of 100 mM ammonium acetate.

RNAi

THP-1 cells were transfected by electroporation with siRNA at 72 h (Fig. 3a and **Supplementary Fig. 2b**) or 48 h (Fig. 3e) before stimulation. The siRNAs with 3′ dTdT modification (Biomers) targeted the following sequences: STING.1, 5′-GGCAUCAAGGAUCGGGUUU-3′; STING.2, 5′-GGUCAUAUUACAUCGGAUA-3′; MAVS, 5′-UUAAAGGAGUUUAUCG AUGUA-3′; cGAS.2, 5′-GGAAGAAAUUAACGACAUU-3′; cGAS.3, 5′-GAAGA AACAUGGCGGCUAU-3′; cGAS.4, 5′-AGAGAAAUGUUGCAGGAAA-3′; IFI16.2, 5′-UCAGAAGACCACAAUCUAC-3′; IFI16.3, 5′-GGUGCUGAACGCA ACAGAAUCAUUU-3′; and controls, 5′-CAUAAGGCAAUGAAGAGAUAG-3′ (luciferase; Fig. 3a) or 5′-GCACCCCAAUAACGAGCUU-3′ (no target; Fig. 3e and **Supplementary Fig. 2b**). mRNA expression at the time point of stimulation was analyzed by quantitative PCR.

Quantitative PCR (RNA detection)

Cells were lysed in 350 μl RLT buffer (Qiagen), then 350 μl 70% ethanol was added and the mixture was loaded onto a Zymo-SpinTM IIIC column. The column was washed sequentially with 350 μl buffer RW1 (Qiagen) and 350 μl Zymo RNA Wash buffer (Zymo). The columns were dried by centrifugation and RNA was eluted with desalted RNase-free water. For the detection of mouse *Ifnb1* mRNA, 1 h of on-column digestion (at 18–20 °C) with DNase I (Zymo) and an additional wash step (500 μl Zymo RNA Wash buffer) were included after the application of Zymo RNA Wash buffer. cDNA was synthesized with a RevertAid RT Kit according to the manufacturer's protocol (Thermo Scientific), with random hexamers. Quantitative PCR was performed with EvaGreen QPCR-Mix II (ROX) (Biobudget) in an ABI7900HT cycler (primer sequences, **Supplementary Table 2**).

Lentiviral particles

Lentiviral particles containing wild-type reverse transcriptase or the mutant RT(N265D) reverse transcriptase were produced by cotransfection of HEK293T cells (by the calcium phosphate method) with expression plasmids encoding vesicular stomatitis virus glycoprotein (pVSV-G), HIV-1 transactivator of transcription (Tat) (pcTat), or regulator of the expression of virion proteins (Rev) protein of Rous sarcoma virus (pRev). HIV-1 group-associated antigen and polymerase (Gag-Pol) sequence encoding wild-type reverse transcriptase (pMDL) or the mutant RT(N265D) reverse transcriptase (pMDL-N256D), and a lentiviral reporter construct expressing green fluorescent protein under the control of a

cytomegalovirus promoter (pHR-CMV-GFP W Sin18) at a ratio of 1.5:1:1:3.5:3.5. For all transfection, the culture medium was replaced after 6 h. The supernatants were harvested 48 h after transfection, then were filtered (0.4 μm) and then purified by ultracentrifugation (90 min at 175,000g) with a sucrose cushion (20% sucrose, 10 mM Tris, pH 7.5, 1 mM EDTA and 100 mM NaCl), separated into aliquots and frozen at −80 °C. Reporter viruses were 'titrated' on HEK293T cells, and green fluorescent protein–positive cells were quantified at 72 h after infection by flow cytometry. In addition, particles containing wild-type or RT(N265D) reverse transcriptase were also normalized for the amount of p24 capsid protein, measured by ELISA (Innogenetics). The mutation encoding the N265D reverse transcriptase mutant[44] was introduced by cloning of a fusion PCR product into the *Age*I and *Mfe*I sites of the Gag-Pol gene of pMDL.

Infection of monocyte-derived macrophages and THP-1 cells

For infection, THP-1 cells (1×10^5 cells per well) or monocyte-derived macrophages (1.5×10^5 cells per well) were seeded into 96-well plates. Cells were preincubated for 2 h with Vpx-delivering virus-like particles and then were infected with lentiviral particles containing wild-type reverse transcriptase at a multiplicity of infection of 1 (THP-1 cells) or 0.67 (monocyte-derived macrophages; 1×10^5 infectious units per well). Infection with lentiviral particles generated with the N265D mutant reverse transcriptase was adjusted according to the concentration of p24 (details in figure legends and Results). At 4 h after infection, cells were washed twice with PBS and then were gently lysed for 5 min on ice in Nonidet-P40 buffer (0.1 M NaCl, 10 mM Tris, pH 8, 2 mM EDTA, 1% β-mercaptoethanol and 0.5% Nonidet-P40). Nuclei were precipitated by centrifugation at 300g for 5 min at 4 °C, and the supernatant cleared by further centrifugation at 5,000g for 5 min at 4 °C. Supernatants were diluted in RLT buffer (Qiagen) and nucleic acids were isolated by a standard RNA-purification protocol (described above in the subsection entitled "Quantitative PCR (RNA detection)"). This protocol did not markedly affect DNA yield compared with results obtained by other protocols (data not shown).

Strand-specific detection of HIV reverse transcripts

Specific detection of HIV (+) and (−) strands included two steps: first, introduction of a strand-specific linker by reaction with DNA polymerase I without exonuclease activity (Klenow exo), and second, linker-specific quantitative PCR. For the strand-specific quantitative PCR (first step), linker-introducing primers (3′ Linker-LTR.1 for (+)-strand detection; 5′ Linker-LTR.2 for (−)-strand detection; **Supplementary Table 2**) were used for first-strand synthesis. Here, cytosolic nucleic acid extracts, mixed with the respective linker primer (4 μM), in 1× FastDigest Buffer (Thermo Scientific), were heated to 95 °C for 5 min, then were cooled to 37 °C. Next, dNTPs (final concentration, 0.1 mM each) and Klenow exo–(1 U/10 μl; Thermo Scientific) were added in 1× FastDigest Buffer for the initiation of first-strand synthesis. After incubation for 15 min at 37 °C, the reaction was stopped by incubation for 10 min at 95 °C.

For the second step, the reaction mix was diluted 1:2 in water and used for quantitative PCR as described for classic quantitative PCR, except that a two-step PCR-protocol was used (50 s at 72 °C and 15 s at 95 °C for each cycle). Primers were 3′-Linker+5′LTR.3 for (+)-strand

detection and 5′Linker+3′LTR.4 for (−)-strand detection (**Supplementary Table 2**). In each experiment, serial dilution of a plasmid (Sigma SHC003) was used as a standard in the Klenow reaction and quantitative PCR. Copy numbers were normalized to the number of mitochondrial genome copies, determined by classic three-step quantitative PCR.

Preparation of ssDNA from PCR products

PCR products were amplified from gBlocks long synthetic DNA (IDT) by Dreamtaq polymerase (Thermo Scientific), with (−)-strand primers comprising a 5′ biotin tag, 5 phosphothioate linkages at the 5′ end and a *Apa*I restriction site, followed by a sequence complementary to the HIV-1 RNA genomic sequence immediately 5′ of the original primer-binding site (identical to the 3′ end of the genomic RNA). The (+)-strand primers had a 5′ phosphate and were derived from the (+)strand sequence 180 bp or 381 bp from the 3′ end of the RNA genome. Primer sequences are in **Supplementary Table 2**. PCR products were purified with an AnalytikJena PCRpure kit and were digested for 30 min with Lambda Exonuclease (Thermo Scientific), followed by heat inactivation for 10 min at 95 °C. DNA was diluted in PBS and then was coupled to NeutrAvidin beads overnight at 4 °C. Beads were spun down (3800*g* and 4 °C for 30 s), then were resuspended in 1× FastDigest buffer containing FastDigest *Apa*I, followed by incubation for 15 min at 37 °C. PBS was added in excess, and the beads were incubated again for 2 h at 4 °C. The beads were washed with PBS twice, then were resuspended in 1× FastDigest Buffer containing *Apa*I anti-sense DNA oligonucleotide and were incubated for 30 min at 18–20 °C. Then, FastDigest *Apa*I was added again and the beads were incubated for 30 min at 37 °C. The ssDNA was again purified with an AnalytikJena PCRpure kit. Single-strandedness was controlled by strand-specific quantitative PCR (a ratio of (−) strand to (+) strand >100) and quality was assessed by denaturing gel electrophoresis (6%).

Statistical analysis

If not stated otherwise, statistical analysis was performed by repeated-measures two-sided, one-way ANOVA with data matched according to donor (PBMCs) or the same experiment (cell lines) with GraphPad Prism 6. If the *P* value calculated by ANOVA was considered significant (<0.05), individual comparisons were corrected for multiple comparisons by Tukey's post-hoc test (if not stated otherwise).

Acknowledgments

References

1. Krieg AM, et al. CpG motifs in bacterial DNA trigger direct B-cell activation. Nature. 1995; 374:546–549. [PubMed: 7700380]

2. Hemmi H, et al. A Toll-like receptor recognizes bacterial DNA. Nature. 2000; 408:740–745. [PubMed: 11130078]

3. Hartmann G, et al. Rational design of new CpG oligonucleotides that combine B cell activation with high IFN-alpha induction in plasmacytoid dendritic cells. Eur J Immunol. 2003; 33:1633–1641. [PubMed: 12778481]

4. Hornung V, et al. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. Nature. 2009; 458:514–518. [PubMed: 19158675]

5. Fernandes-Alnemri T, Yu JW, Datta P, Wu J, Alnemri ES. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. Nature. 2009; 458:509–513. [PubMed: 19158676]

6. Bürckstümmer T, et al. An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. Nat Immunol. 2009; 10:266–272. [PubMed: 19158679]

7. Ablasser A, et al. RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. Nat Immunol. 2009; 10:1065–1072. [PubMed: 19609254]

8. Chiu YH, Macmillan JB, Chen ZJ. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. Cell. 2009; 138:576–591. [PubMed: 19631370]

9. Schlee M, et al. Recognition of 5′triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. Immunity. 2009; 31:25–34. [PubMed: 19576794]

10. Cavlar T, Ablasser A, Hornung V. Induction of type I IFNs by intracellular DNA-sensing pathways. Immunol Cell Biol. 2012; 90:474–482. [PubMed: 22450802]

11. Stetson DB, Medzhitov R. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. Immunity. 2006; 24:93–103. [PubMed: 16413926]

12. Yang P, et al. The cytosolic nucleic acid sensor LRRFIP1 mediates the production of type I interferon via a beta-catenin-dependent pathway. Nat Immunol. 2010; 11:487–494. [PubMed: 20453844]

13. Bagashev A, et al. Leucine-rich repeat (in Flightless I) interacting protein-1 regulates a rapid type I interferon response. J Interferon Cytokine Res. 2010; 30:843–852. [PubMed: 20586614]

14. Zhang X, et al. Cutting edge: Ku70 is a novel cytosolic DNA sensor that induces type III rather than type I IFN. J Immunol. 2011; 186:4541–4545. [PubMed: 21398614]

15. Kim T, et al. Aspartate-glutamate-alanine-histidine box motif (DEAH)/RNA helicase A helicases sense microbial DNA in human plasmacytoid dendritic cells. Proc Natl Acad Sci USA. 2010; 107:15181–15186. [PubMed: 20696886]

16. Ishikawa H, Ma Z, Barber GN. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. Nature. 2009; 461:788–792. [PubMed: 19776740]

17. Zhong B, et al. The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. Immunity. 2008; 29:538–550. [PubMed: 18818105]

18. Unterholzner L, et al. IFI16 is an innate immune sensor for intracellular DNA. Nat Immunol. 2010; 11:997–1004. [PubMed: 20890285]

19. Sun L, Wu J, Du F, Chen X, Chen ZJ. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. Science. 2013; 339:786–791. [PubMed: 23258413]

20. Jakobsen MR, Paludan SR. IFI16: At the interphase between innate DNA sensing and genome regulation. Cytokine Growth Factor Rev. 2014; 25:649–655. [PubMed: 25027602]

21. Wu J, et al. Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. Science. 2013; 339:826–830. [PubMed: 23258412]

22. Gao P, et al. Cyclic [G(2′,5′)pA(3′,5′)p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. Cell. 2013; 153:1094–1107. [PubMed: 23647843]

23. Ablasser A, et al. cGAS produces a 2–5′′-linked cyclic dinucleotide second messenger that activates STING. Nature. 2013; 498:380–384. [PubMed: 23722158]

24. Civril F, et al. Structural mechanism of cytosolic DNA sensing by cGAS. Nature. 2013; 498:332–337. [PubMed: 23722159]

25. Kato K, et al. Structural and functional analyses of DNA-sensing and immune activation by human cGAS. PLoS ONE. 2013; 8:e76983. [PubMed: 24116191]

Nat Immunol. Author manuscript; available in PMC 2015 December 03.

-110-

26. Kranzusch PJ, Lee AS, Berger JM, Doudna JA. Structure of human cGAS reveals a conserved family of second-messenger enzymes in innate immunity. Cell Rep. 2013; 3:1362–1368. [PubMed: 23707061]

27. Gao D, et al. Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. Science. 2013; 341:903–906. [PubMed: 23929945]

28. Zhang X, et al. The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. Cell Rep. 2014; 6:421–430. [PubMed: 24462292]

29. Li X, et al. Cyclic GMP-AMP synthase is activated by double-stranded DNA-induced oligomerization. Immunity. 2013; 39:1019–1031. [PubMed: 24332030]

30. Li XD, et al. Pivotal roles of cGAS-cGAMP signaling in antiviral defense and immune adjuvant effects. Science. 2013; 341:1390–1394. [PubMed: 23989956]

31. Lam E, Stein S, Falck-Pedersen E. Adenovirus detection by the cGAS/STING/TBK1 DNA sensing cascade. J Virol. 2014; 88:974–981. [PubMed: 24198409]

32. Lahaye X, et al. The capsids of HIV-1 and HIV-2 determine immune detection of the viral cDNA by the innate sensor cGAS in dendritic cells. Immunity. 2013; 39:1132–1142. [PubMed: 24269171]

33. Rasaiyaah J, et al. HIV-1 evades innate immune recognition through specific cofactor recruitment. Nature. 2013; 503:402–405. [PubMed: 24196705]

34. Suzuki K, et al. Activation of target-tissue immune-recognition molecules by double-stranded polynucleotides. Proc Natl Acad Sci USA. 1999; 96:2285–2290. [PubMed: 10051633]

35. Ishii KJ, et al. A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. Nat Immunol. 2006; 7:40–48. [PubMed: 16286919]

36. Jakobsen MR, et al. From the Cover: IFI16 senses DNA forms of the lentiviral replication cycle and controls HIV-1 replication. Proc Natl Acad Sci USA. 2013; 110:E4571–E4580. [PubMed: 24154727]

37. Galvis AE, Fisher HE, Nitta T, Fan H, Camerini D. Impairment of HIV-1 cDNA Synthesis by DBR1 Knockdown. J Virol. 2014; 88:7054–7069. [PubMed: 24672043]

38. Gehrke N, et al. Oxidative damage of DNA confers resistance to cytosolic nuclease TREX1 degradation and potentiates STING-dependent immune sensing. Immunity. 2013; 39:482–495. [PubMed: 23993650]

39. Huppert JL. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. Chem Soc Rev. 2008; 37:1375–1384. [PubMed: 18568163]

40. Murchie AI, Lilley DM. Retinoblastoma susceptibility genes contain 5′ sequences with a high propensity to form guanine-tetrad structures. Nucleic Acids Res. 1992; 20:49–53. [PubMed: 1738603]

41. Hagmann CA, et al. RIG-I detects triphosphorylated RNA of Listeria monocytogenes during infection in non-immune cells. PLoS ONE. 2013; 8:e62872. [PubMed: 23653683]

42. Takaoka A, et al. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. Nature. 2007; 448:501–505. [PubMed: 17618271]

43. Zhang Z, et al. The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. Nat Immunol. 2011; 12:959–965. [PubMed: 21892174]

44. Fisher TS, Darden T, Prasad VR. Mutations proximal to the minor groove-binding track of human immunodeficiency virus type 1 reverse transcriptase differentially affect utilization of RNA versus DNA as template. J Virol. 2003; 77:5837–5845. [PubMed: 12719577]

45. Mankan AK, et al. Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. EMBO J. 2014; 33:2937–2946. [PubMed: 25425575]

46. Stetson DB, Ko JS, Heidmann T, Medzhitov R. Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell. 2008; 134:587–598. [PubMed: 18724932]

47. Ablasser A, et al. TREX1 deficiency triggers cell-autonomous immunity in a cGAS-dependent manner. J Immunol. 2014; 192:5993–5997. [PubMed: 24813208]

48. Yoh SM, et al. PQBP1 Is a Proximal Sensor of the cGAS-Dependent Innate Response to HIV-1. Cell. 2015; 161:1293–1305. [PubMed: 26046437]

49. Watts JM, et al. Architecture and secondary structure of an entire HIV-1 RNA genome. Nature. 2009; 460:711–716. [PubMed: 19661910]

50. van der Kuyl AC, Berkhout B. The biased nucleotide composition of the HIV genome: a constant factor in a highly variable virus. Retrovirology. 2012; 9:92. [PubMed: 23131071]

51. van Hemert FJ, van der Kuyl AC, Berkhout B. The A-nucleotide preference of HIV-1 in the context of its structured RNA genome. RNA Biol. 2013; 10:211–215. [PubMed: 23235488]

52. Martin-Gayo E, et al. Potent cell-intrinsic immune responses in dendritic cells facilitate HIV-1-specific T cell immunity in HIV-1 elite controllers. PLoS Pathog. 2015; 11:e1004930. [PubMed: 26067651]

53. Gall A, et al. Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease. Immunity. 2012; 36:120–131. [PubMed: 22284419]

54. Bauernfeind F, et al. Cutting edge: reactive oxygen species inhibitors block priming, but not activation, of the NLRP3 inflammasome. J Immunol. 2011; 187:613–617. [PubMed: 21677136]

55. Jin T, et al. Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor. Immunity. 2012; 36:561–571. [PubMed: 22483801]
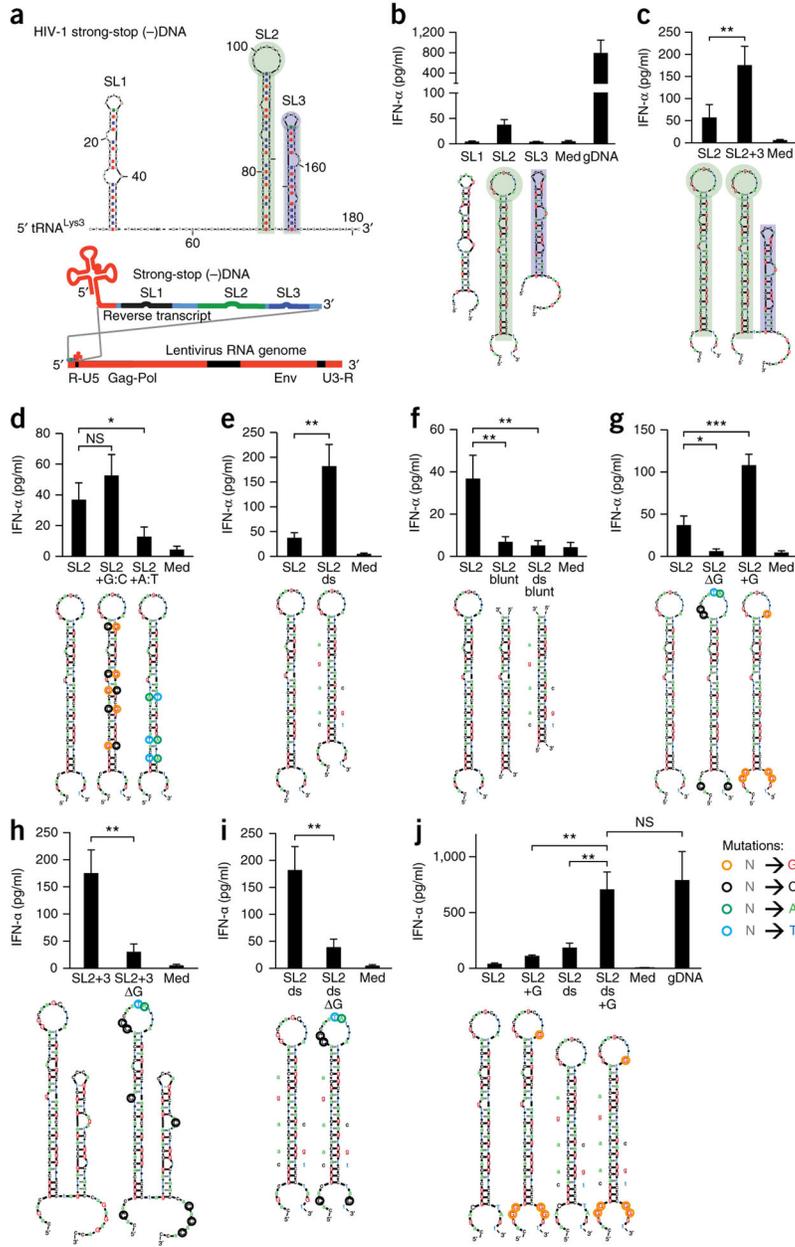
Figure 1.

Immunostimulation by stem-loop-forming structures in HIV-1 sstDNA depends on unpaired guanosines flanking the stem. (**a**) mFOLD model of HIV-1 sstDNA, the primary lysyl-tRNA (tRNA^{Lys3})-primed reverse transcript of the 5′ untranslated region (R-U5). Gag-Pol and Env, regions encoding HIV-1 group-associated antigen (Gag), polymerase (Pol) and envelope (Env) proteins; U3-R, 3′ untranslated region and repeat. (**b**–**j**) Enzyme-linked immunosorbent assay (ELISA) of IFN-α in supernatants of chloroquine-treated human PBMCs 20 h after treatment with medium alone (Med) or transfection of genomic DNA (gDNA) or wild-type stem-loop structures SL1, SL2 or SL3 (**b**), or of wild-type SL2 (**c**–**j**) or a combination of SL2 and SL3 (SL2+3) (**c**), SL2 with introduced C:G base pairs (SL2

+G:C) or A:T base pairs (SL2 +A:T) (**d**), SL2 with mismatches and bulges removed from the stem (SL2 ds) (**e**), SL2 (SL2 blunt) or the SL2 mutant in **e** (SL2 ds blunt) with single-stranded parts removed (**f**), SL2 with substitution of guanosines (SL2 G) or conversion of various nucleotides (color loops) to guanosine (SL2+G) in the loop as well as the 3′ and 5′ ends (**g**), the combined SL2 plus SL3 in **c** with (SL2+3 G) or without (SL2+3) substitution of unpaired guanosines (**h**), the SL2 mutant in **e** with (SL2 ds G) or without (SL2 ds) substitution of unpaired guanosines (**i**), or SL2 with conversion of various nucleotides (color loops) to guanosine (SL2+G), removal of mismatches and bases (SL2 ds) or a combination of those changes (SL2 ds+G) (**j**). Below plots (**b**–**j**), mFOLD models of secondary structures (sequences, **Supplementary Table 1**): colors indicate mutations (key; 'N' is any base.); letters adjacent to stems indicate bases removed. NS, not significant ($P > 0.05$); *P 0.05; **P 0.01 and ***P 0.001 (one-way analysis of variance (ANOVA) followed by Fisher's least-significant difference LSD test.). Data are pooled from two (**b**,**d**,**e**,**g**–**j**) or four (**c**,**f**) experiments with two biological replicates in each (mean and s.e.m. of $n = 4$ donors (**b**,**d**,**e**,**g**–**j**; same experiments) or $n = 8$ donors (**c**,**f**; same experiments)).

-114-
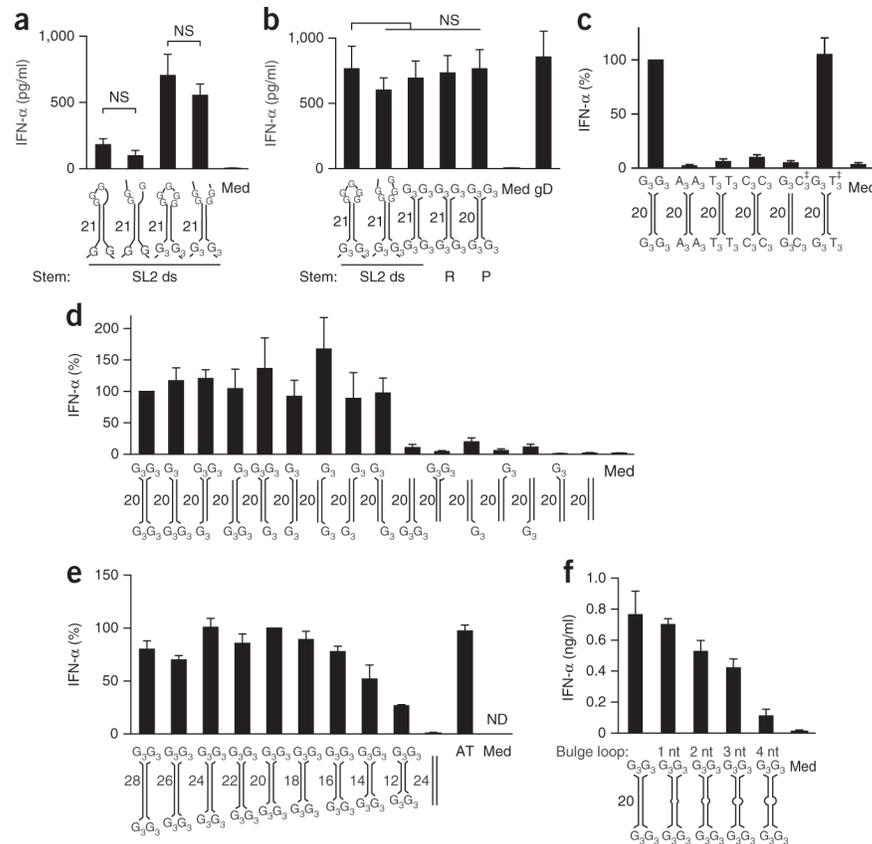
Figure 2.
Guanosine extensions render short DNA duplexes highly immunostimulatory. (**a**,**b**) ELISA of IFN-α in the supernatant of chloroquine-treated human PBMCs 20 h after treatment with medium alone or transfection of genomic DNA (as in Fig. 1) or the SL2 variant with mismatches and bulges removed from the stem (SL2 ds) and with a closed loop (first and third bars) or open loop (second and fourth bars) comprising the unmodified unpaired regions (first and second bars) or with conversion of various nucleotides to guanosine (as in Fig. 1j; third and fourth bars) (**a**) or with short duplexes comprising variants of single-stranded regions (first bar as in the third bar in **a**; second bar as in the fourth bar in **a**; third bar with G_3 at each end) as well as different stem elements (third, fourth and fifth bars), randomized stem sequence (R) or palindromic stem sequence (P) (**b**). (**c**) ELISA of IFN-α in chloroquine-treated human PBMCs as in **a**,**b**, transfected with palindromic G_3-YSD, A_3-YSD, T_3-YSD or C_3-YSD, or with G_3-YSD hybridized to C_3-YSD or T_3-YSD, the latter two at a twofold excess (‡); results are presented relative to those of cells transfected with G_3-YSD, set as 100%. (**d**) ELISA of IFN-α in chloroquine-treated human PBMCs as in **a**,**b**, transfected with G_3-YSD or YSD with various numbers and positions of unpaired G trimers flanking 20-nucleotide dsDNA (presented as in **c**). (**e**) ELISA of IFN-α in chloroquine-treated human PBMCs as in **a**,**b**, transfected with G_3-YSD of various duplex lengths (12–28 bp) or with poly(dAdT) (AT); results are presented relative to those of cells transfected with G_3-YSD with a 20-bp duplex, set as 100%. (**f**) ELISA of IFN-α in chloroquine-treated human PBMCs as in **a**,**b**, transfected with G_3-YSD of various bulge-loop sizes in the base-

paired region. Below plots, models of secondary structures (sequences, **Supplementary Table 1**); numbers adjacent to stems indicate stem length. ND, not detectable. NS, not significant (one-way-ANOVA followed by Fisher's least-significant difference test). Data are pooled from two (**a–c,e,f**) or three (**d**) experiments with one or two biological replicates in each (mean and s.e.m. of $n = 4$ donors (**a–c,f**), $n = 5$ donors (**d**) or $n = 3$ donors (**e**)).
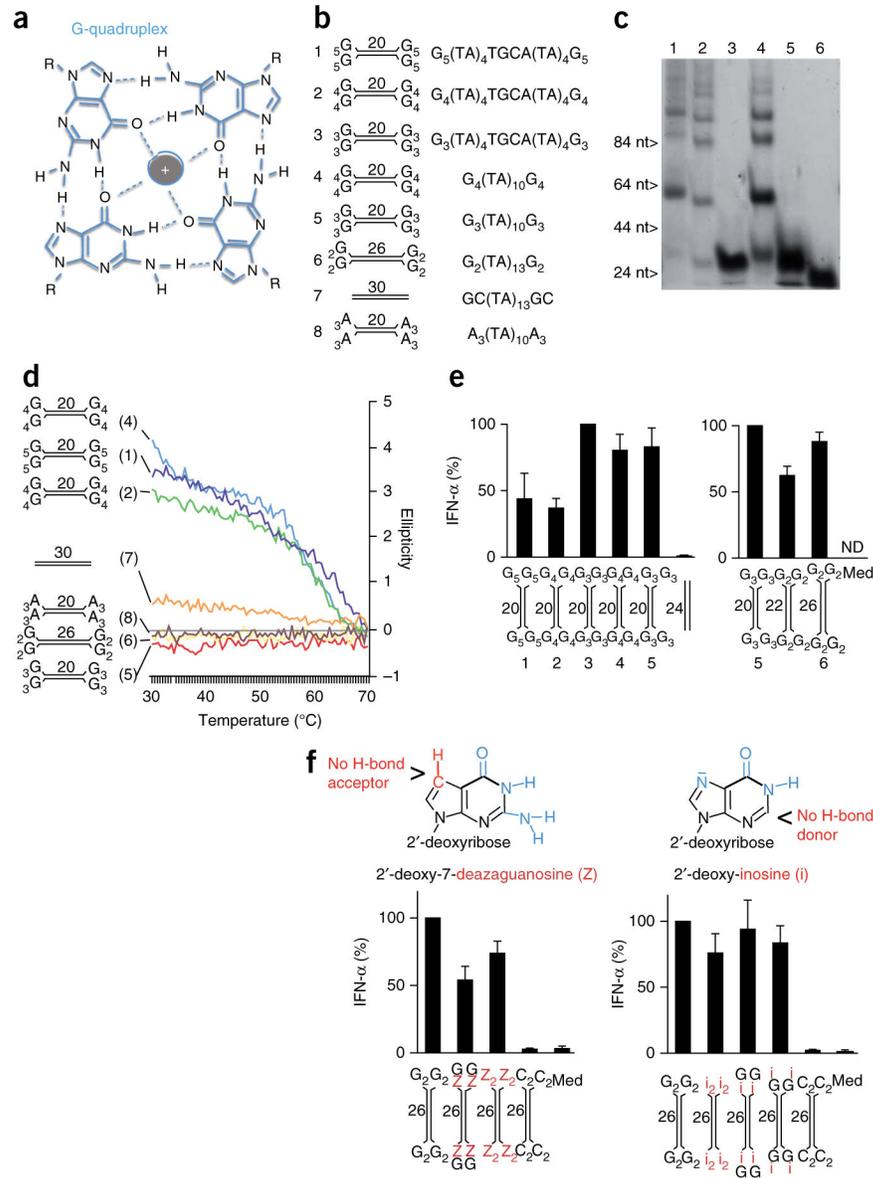
Figure 3.

Induction of IFN-α/β by Y-form DNA is independent of the G quadruplex. (**a**) G-quadruplex assembly: four guanosines form a planar structure, mediated by hydrogen bonds and stabilized by a central cation (gray). R, deoxyribose backbone. (**b**) Structures and sequences of YSD or short blunt DNA used in **c**–**e**. (**c**) Native PAGE of G_n-YSD variants (numbers above lanes correspond to models in **b**). (**d**) Circular-dichroism melting curve at 295 nm, with the ellipticity of YSD variants or blunt-end DNA (numbers in parentheses and diagrams at left match those in **b**) measured at 30–70 °C. (**e**) ELISA of IFN-α in supernatants of chloroquine-treated human PBMCs 20 h after transfection of G_n-YSD variants (numbers and diagrams below plots match those in **b**); results are presented relative to those of cells transfected with G_3-YSD variant 3 (left) or variant 5 (right), set as 100%. (**f**) ELISA of IFN-α (bottom) as in **e**, after transfection of 26-nucleotide G_2-YSD with stepwise

substitution of terminal guanosines (G) with deazaguanosine (Z; left) or inosine (i; right); top, chemical structure of 2′-deoxy-7-deazaguanosine and 2′-deoxy-inosine. Numbers adjacent to stems (**b**,**d**–**f**) indicate stem length (sequences of DNA structures, **Supplementary Table 1**). Data are representative of two experiments (**c**,**d**) or are pooled from two (**e**, **f**, right) or three (**f**, right) experiments with one or two biological replicates in each (**e**,**f**; mean and s.e.m. of $n = 3$ donors (**e**), $n = 6$ donors (**f**, left) or $n = 4$ donors (**f**, right)).
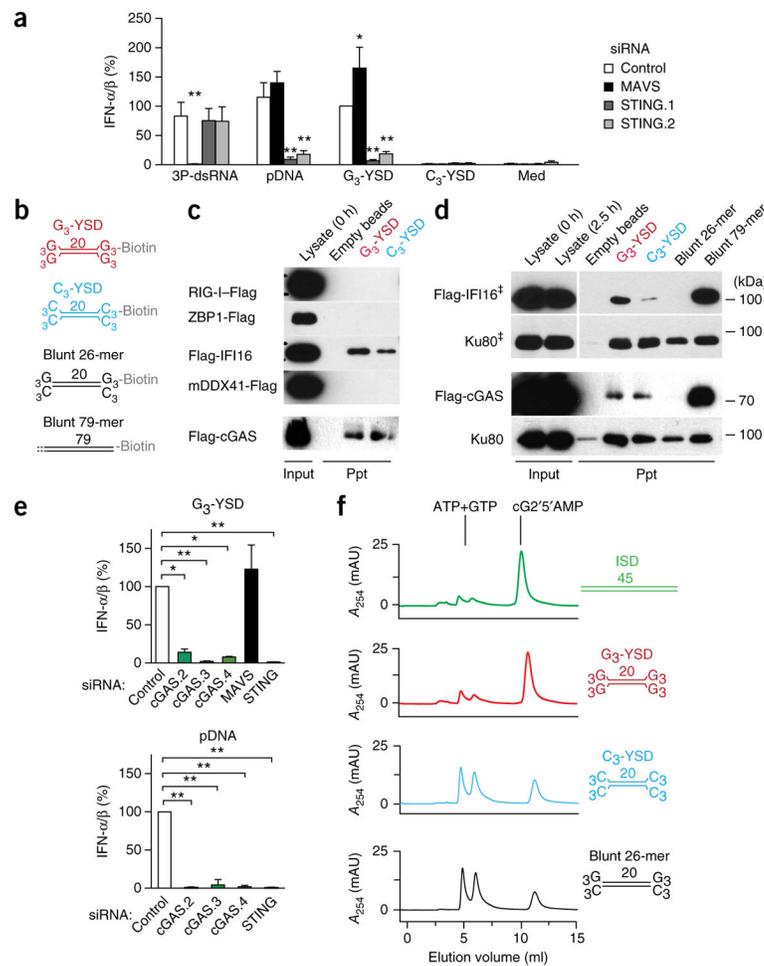
Figure 4.
G-YSD activates the cGAS-STING axis. (**a**) IFN-α/β activity in supernatants of THP-1 cells treated with control siRNA or siRNA targeting MAVS or STING (either of two target sequences (STING.1 or STING.2)) (key) and then, 72 h later, stimulated for 20 h with the RIG-I ligand 3P-dsRNA, plasmid DNA (pDNA), G_3-YSD or C_3-YSD; results are presented relative to those of THP-1 cells treated with control siRNA and stimulated with G_3-YSD, set as 100%. (**b**) Secondary structures of biotinylated DNA duplexes used in interaction assays in **c,d**: G_3-YSD or C_3-YSD, or short, 26-nucleotide (26-mer) or long, 79-nucleotide (79-mer) blunt dsDNA. (**c**) Immunoblot analysis of Flag-tagged receptor candidates (left margin) with lysate alone (0.1 volume of the lysate used in the precipitation assays; Input) at the beginning (0 h) of the assay (far left lane) or after precipitation (Ppt) with empty beads (middle left lane) or with G_3-YSD or C_3-YSD (above lanes; as in **b**). (**d**) Immunoblot analysis of Flag-tagged IFI16 or cGAS or of the nonspecific DNA-binding control Ku80 with lysate alone (Input) at the beginning (0 h) or end (2.5 h) of the experiment, or after precipitation with empty beads, G_3-YSD, C_3-YSD (as in **c**), or with short (26-nucleotide) or long (79-nucleotide) blunt dsDNA (as in **b**) (above lanes). ‡, lysate and precipitate detection are from the same blot with one empty lane removed. (**e**) IFN-α/β in supernatants of THP-1 cells treated with control siRNA or siRNA targeting cGAS (one of three target sequences

-119-

(cGAS.2, cGAS.3 or cGAS.4)), MAVS or STING (horizontal axes) and then, 48 h later, stimulated for 20 h with G_3-YSD (top) or plasmid DNA (bottom); results are presented relative to those of THP-1 cells treated with control siRNA, set as 100%. (**f**) HPLC detecting the conversion of ATP and GTP to cG2′5′AMP (cyclic [G(2′,5′)pA(3′,5′)p]) by cGAS in the presence of ISD, G_3-YSD, C_3-YSD or the 26-nucleotide blunt DNA (without biotin) in **b**, presented (in milli-absorption units (mAU)) as absorption at 254 nm (A_{254}). Sequences of DNA structures, **Supplementary Table 1**. *P 0.01 and **P 0.001, presented only for results significantly different from control (two-way ANOVA followed by Bonferroni's post-hoc test (**a**) or one-way ANOVA followed by Tukey's post-hoc test (**e**)). Data are pooled from three (**a,e**, bottom) or four (**e**, top) independent experiments with biological replicates (mean and s.e.m.) or are representative of three experiments (**c,d,f**).
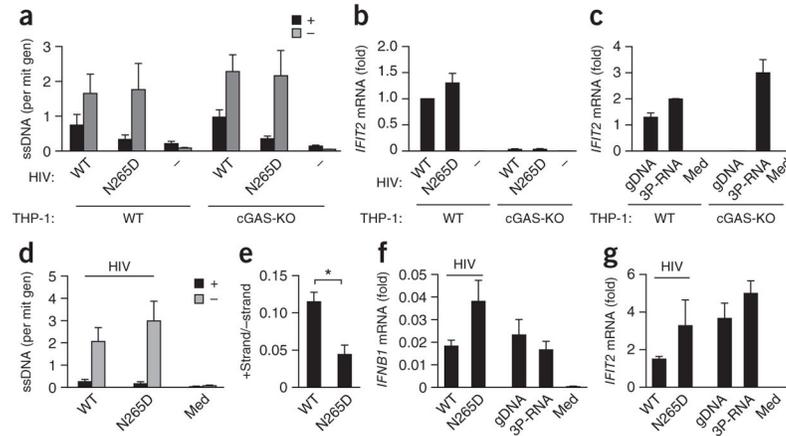
Figure 5.
The interferon response induced by HIV-1 infection correlates with the presence of (−)-strand DNA but not with the presence of (+)-strand DNA. (**a**) Strand-specific quantitative detection of the HIV-1 (+) strand and (−) strand (key) in cytosol-enriched nucleic acid preparations of wild-type (WT) or cGAS-deficient (cGAS-KO) THP-1 cells 4 h after infection with HIV-1 particles containing wild-type (HIV: WT) or mutant (HIV: N265D) reverse transcriptase or without infection (HIV: −); results are presented as ssDNA copies per mitochondrial genome (mit gen). (**b,c**) *IFIT2* mRNA in THP-1 cells 4 h after treatment as in **a** (**b**) or after transfection of genomic DNA or 3P-dsRNA or treatment with medium alone (Med) (**c**); results are presented as copies of *IFIT2* mRNA per copy of control *GAPDH* mRNA (**c**), or that value relative to the results of wild-type THP-1 cells infected with wild-type particles (**b**). (**d**) Strand-specific quantitative detection of the HIV-1 (+) and (−) strand (key) in cytosol-enriched nucleic acid preparations of monocyte-derived macrophages 4 h after no infection (Med) or infection with HIV-1 particles as in **a** (presented as in **a**). (**e**) Ratio of (+) strand to (−) strand in **d**. (**f,g**) *IFNB1* mRNA (**f**) and *IFIT2* mRNA (**g**) in monocyte-derived macrophages 4 h after stimulation with HIV-1 particles as in **a** or transfection of genomic DNA or 3P-dsRNA; results are presented as copies of *IFNB1* mRNA (**f**) or *IFIT2* mRNA (**g**) per copy of control *GAPDH* mRNA. *P 0.01 (ratio-paired *t*-test). Data are from three (**a,b**) i or two (**c**) independent experiments (mean and s.e.m.) or are pooled from two experiments with two biological replicates in each (**d–g**; mean and s.e.m. of $n = 4$ donors).
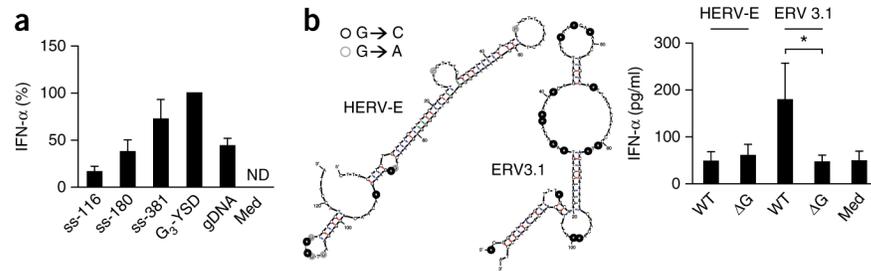
Figure 6.
Long ssDNA comprising the 5′-terminal HIV-1 (−)-strand sequence is highly immunostimulatory, and the recognition of endogenous retroelement-derived ssDNA depends on unpaired guanosines. (**a**) ELISA of IFN-α in supernatants of monocyte-derived macrophages 36 h after treatment with medium alone (Med) or transfection of genomic DNA, G_3-YSD or ssDNA species of various lengths (116 nucleotides (ss-116; this is SL2 plus SL3 as in Fig. 1c), 180 nucleotides (ss-180) or 381 nucleotides (ss-381)); results are presented relative to those of cells transfected with G_3-YSD, set as 100%. (**b**) IFN-α in supernatants of chloroquine-treated PBMCs 24 h after treatment with medium alone (Med) or transfection of wild-type (WT) or mutant (ΔG) HERV-E or ERV3.1 (right), and mFOLD-derived models of the secondary structures of HERV-E or ERV3.1 (left); circles indicate mutation of guanosine (black, G to C; gray: G to A). *P 0.05 (one-way ANOVA followed by Tukey's post-hoc test). Data are pooled from two experiments with one or two biological replicates in each (mean and s.e.m. of $n = 3$ donors (**a**) or $n = 4$ donors (**b**)).

3. Cytosolic RNA:DNA hybrids activate the cGAS–STING axis

Summary

This publication for the first time describes and characterizes RNA:DNA hybrids as cGAS activators. cGAS was discovered as a central sensor for cytosolic DNA driving type I IFN production. Since autoinflammatory disorders like Aicardi–Goutières syndrome (AGS) or systemic lupus erythematosus (SLE) are characterized by elevated type I IFN production in the serum and are associated with mutations in nucleic acids metabolizing enzymes SAMHD1, TREX1, RNase H2 and ADAR1, an involvement of cytosolic nucleic acids sensors in their development was hypothesized. cGAS was shown to be responsible for IFN signature in TREX1-deficient mice recognizing endogenous DNA molecules. Though TREX1 deficiency was reported to result in accumulation of reverse transcription products emerging from retroelements, the exact nature of these DNA species remains unclear. In this work poly(rA):poly(dT) RNA:DNA hybrids were found to stimulate type I IFN response in mouse and human cells, as measured by IFN-β mRNA synthesis, cytokine IP-10 production and with luciferase-reporter assays. Knock-out studies in THP-1 cells revealed such IFN response to RNA:DNA hybrids to be strongly dependent on intact cGAS-STING pathway. In vitro studies confirmed a direct activation of cGAS by RNA:DNA hybrids of different lengths. Intriguingly, RNA:DNA hybrids were found to activate cGAS less efficiently in comparison to dsDNA of the same length both in vitro and in the cell-based assays. Though a modelling of cGAS complexes with dsDNA and RNA:DNA hybrids revealed a strong similarity of cGAS binding modes, RNA:DNA decreased efficiency in cGAS activation might arise from insufficient or unsteady conformational change introduced to cGAS, in contrast to dsDNA binding. However, given a self-enhancing positive feedback loop, even small amounts of cGAMP produced by cGAS upon binding of RNA:DNA hybrids might be sufficient to induce a powerful innate immune response. Thus, this works introduces RNA:DNA hybrids as potent cGAS activators and offers a link between cGAS activation and retroelement and retroviral replication intermediates, as well as autoinflammation.

Author contribution

The author of this thesis expressed and purified recombinant human cGAS Mab-21 construct in a large-scale.

Article

Cytosolic RNA:DNA hybrids activate the cGAS–STING axis

Arun K Mankan[1], Tobias Schmidt[1,†], Dhruv Chauhan[1,†], Marion Goldeck[2], Klara Höning[1], Moritz Gaidt[1], Andrew V Kubarenko[1,‡], Liudmila Andreeva[3], Karl-Peter Hopfner[3] & Veit Hornung[1,*]

Abstract

Intracellular recognition of non-self and also self-nucleic acids can result in the initiation of potent pro-inflammatory and antiviral cytokine responses. Most recently, cGAS was shown to be critical for the recognition of cytoplasmic dsDNA. Binding of dsDNA to cGAS results in the synthesis of cGAMP(2′–5′), which then binds to the endoplasmic reticulum resident protein STING. This initiates a signaling cascade that triggers the induction of an antiviral immune response. While most studies on intracellular nucleic acids have focused on dsRNA or dsDNA, it has remained unexplored whether cytosolic RNA:DNA hybrids are also sensed by the innate immune system. Studying synthetic RNA:DNA hybrids, we indeed observed a strong type I interferon response upon cytosolic delivery of this class of molecule. Studies in THP-1 knockout cells revealed that the recognition of RNA:DNA hybrids is completely attributable to the cGAS–STING pathway. Moreover, *in vitro* studies showed that recombinant cGAS produced cGAMP upon RNA:DNA hybrid recognition. Altogether, our results introduce RNA:DNA hybrids as a novel class of intracellular PAMP molecules and describe an alternative cGAS ligand next to dsDNA.

Introduction

The innate immune system has evolved to defend the host against invading pathogens. An important prerequisite for this task is the specific and reliable detection of different microbial pathogens as non-self. This is achieved by a conserved set of germ-line-encoded pathogen recognition receptors (PRRs) that have evolved to detect so-called pathogen-associated molecular patterns (PAMPs) as foreign. PRRs include, amongst others, Toll-like receptors (TLRs), Nod-like receptors (NLRs) and RIG-I-like receptors (RLRs) (Medzhitov, 2007).

Detection of virus-derived nucleic acids plays a central role in the initiation of antiviral immunity. Nucleic acid recognition by PRRs results in the secretion of type I interferon (IFN) cytokines and IFN-stimulated genes (ISGs), which function to impede viral replication. For example, 5′-triphosphorylated RNA is detected by the cytoplasmic RNA helicase RIG-I, whereas long double-stranded RNA is sensed by its related family member MDA5. Both RIG-I and MDA5 signal via their shared signaling adapter molecule MAVS, located at the mitochondrion (Goubau *et al*, 2013). In certain cell types, RNA can also be detected by TLR7 and TLR8, located in the endolysosomal compartment. While these receptors display a certain preference for non-self RNA, they are in principle also responsive to endogenous RNA molecules. However, under normal circumstances, their localization in the endolysosomal compartment shields them from endogenous RNA molecules, thereby precluding erroneous activation by self-molecules. A similar scenario holds true for TLR9, which is also located in the endolysosome, detecting DNA of certain sequence composition (CpG motifs) (Barbalat *et al*, 2011). More recently, a cytosolic nucleotidyltransferase named cGAS (cyclic GMP-AMP synthase) was identified as the key sensor required for DNA recognition in the cytoplasm (Sun *et al*, 2013; Wu *et al*, 2013). Upon binding to dsDNA, cGAS catalyzes the formation of the cyclic dinucleotide (CDN) molecule cGAMP using ATP and GTP as substrates. Unlike previously known prokaryotic CDNs, the cGAS-derived CDN contains an unusual 2′–5′ phosphodiester linkage between GMP and AMP, with its second phosphodiester linkage being 3′–5′: >Gp(2′–5′)Ap(3′–5′)>(cGAMP(2′–5′)) (Ablasser *et al*, 2013a; Diner *et al*, 2013; Gao *et al*, 2013b; Wu *et al*, 2013; Zhang *et al*, 2013). Upon formation, cGAMP(2′–5′) binds to the endoplasmic reticulum resident protein STING, which results in the activation of the TBK1/IRF3 pathway and as such the production of type I interferons. Interestingly, cGAMP(2′–5′) is not only bound to exert its antiviral activity within the cell it has been produced in, but it can also be passed on to bystander cells via gap junctions. This, in

1 Institute of Molecular Medicine, University Hospital, University of Bonn, Bonn, Germany
2 Institute of Clinical Chemistry and Clinical Pharmacology, University Hospital, University of Bonn, Bonn, Germany
3 Department of Biochemistry and Gene Center, Ludwig-Maximilians-University, Munich, Germany
*Corresponding author. Tel: +49 228 287 51203; Fax: +49 228 287 51201; E-mail: veit.hornung@uni-bonn.de
†These authors contributed equally
‡Present address: Institute of Clinical Chemistry and Clinical Pharmacology, University Hospital, University of Bonn, Bonn, Germany

-124-

turn, can then initiate antiviral immune responses in cells that have not been in contact with a virus yet (Ablasser *et al*, 2013c).

As mentioned above, certain PRRs not only recognize pathogenic components but can also sense self-molecules, once mislocalized in the respective PRR compartment (Ablasser *et al*, 2013b). For example, defects in cytoplasmic nucleases can result in the accumulation of endogenous nucleic acids in the cytosol and thereby lead to the activation of innate sensing pathways. Aicardi–Goutières syndrome (AGS) is a rare but generally fatal childhood inflammatory condition with neurological dysfunction, that is associated with increased production of type I interferons (Crow, 2011; Lee-Kirsch *et al*, 2014). This disease predominantly results from loss-of-function mutations in nucleic acid metabolizing enzymes such as SAMHD1, TREX1, Ribonuclease H2 (RNASE H2A, RNASE H2B, RNASE H2C) or ADAR1 (Gall *et al*, 2012; Lee-Kirsch *et al*, 2014). TREX1 deficiency in mice has been shown to result in the accumulation of endogenous DNA molecules, which can serve as ligands for the cGAS–STING axis and as such initiate antiviral immunity (Gall *et al*, 2012; Ablasser *et al*, 2014). However, the exact nature of these ligands is currently not known. On the other hand, ablation of the *Rnaseh2b* gene in mice results in the incorporation of ribonucleotides in the genomic DNA; however, a mechanistic link to PRR-mediated proinflammatory gene expression has not been established (Reijns *et al*, 2012).

Intrigued by the possibility of cytosolic RNA:DNA hybrids serving as PAMP molecules on their own, we set forth to investigate the immunostimulatory capacity of this class of molecules in cells competent for both dsRNA and dsDNA sensing pathways.
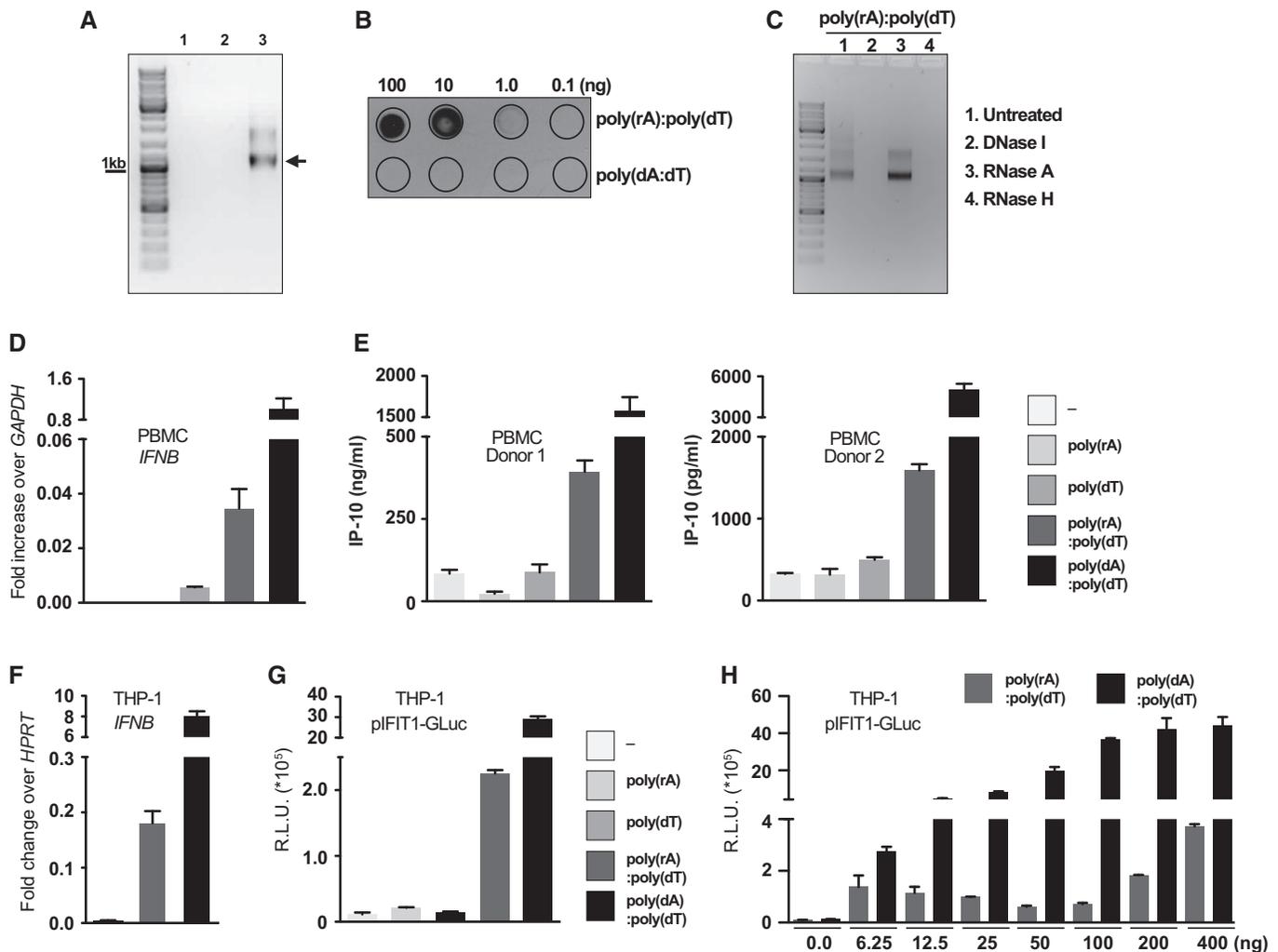
Results

Intracellular RNA:DNA hybrids induce antiviral gene expression

Using *in vitro* transcribed RNA molecules (e.g. T7 RNA polymerase) as templates for reverse transcriptase, RNA:DNA hybrids of different lengths can easily be produced enzymatically. However, reverse transcriptase also possesses DNA-dependent DNA polymerase activity and as such can generate dsDNA from the newly synthesized ssDNA (Hsieh *et al*, 1993). Therefore, to faithfully exclude dsDNA contamination in assessing the immunostimulatory capacity of DNA:RNA hybrids, we made use of enzymatically generated homopolymers of poly(rA) and poly(dT). Annealing of poly(rA) and poly(dT) resulted in the formation of a poly(rA):poly(dT) RNA:DNA hybrid with a predominant band of around 1,300 bp (Fig 1A). To confirm that this product indeed represented a RNA:DNA hybrid, we performed dot-blot assays wherein membrane-immobilized nucleic acids were probed with a monoclonal antibody (S9.6) specific for RNA:DNA hybrids (Boguslawski *et al*, 1986; Hu *et al*, 2006). We observed a specific detection of the hybrid by the antibody, correlating with the amount of nucleic acids immobilized (Fig 1B). Digestion of the hybrid with RNase A which digests dsRNA, DNase I which digests dsDNA or ssDNA and RNase H which digests RNA part of RNA:DNA hybrids provided further confirmation regarding the purity of the hybrid. Thus, while the RNA:DNA hybrid was completely digested by RNase H and DNase I enzymes, it was resistant, as expected, to RNase A activity (Fig 1C). Control experiments confirmed the specificity of these enzymes on dsDNA or dsRNA, respectively (A.K. Mankan, unpublished

observations). In order to assess the biological activity of RNA:DNA hybrids, we then transfected the poly(rA):poly(dT) hybrids or the single components poly(rA) or poly(dT) into murine bone marrow-derived macrophages. Transfection of the hybrids resulted in a robust induction of type I IFNs and pro-inflammatory genes (*Ifna*, *Ifnb*, *Il6* and *Ip10*) (Supplementary Fig S1A). Of note, a significant activity of poly(dT), by itself, was also observed for some readouts studied (see discussion below). To assess the relevance of hybrids in the human system, we next transfected hybrids or the individual polynucleotide components into peripheral blood mononuclear cells. RT–PCR confirmed enhanced expression of *IFNB* gene in PBMCs transfected with hybrids (Fig 1D). Synthetic poly(dA):poly(dT), which was transfected as a dsDNA control, showed higher activity with regard to *IFNB* induction. We also observed a significant secretion of IP-10 in response to the RNA:DNA hybrids, while only a minimal IP-10 induction was seen in cells transfected with the single polynucleotides (Fig 1E). We next tested the activity of RNA:DNA hybrids in differentiated human THP-1 cells. In line with the PBMC data, RNA:DNA hybrids induced robust transcription of *IFNB* in THP-1 cells, again with dsDNA being more active (Fig 1F). In order to have a sensitive readout for antiviral gene expression, we equipped THP-1 cells with Gaussia luciferase (GLuc) under the promoter of *IFIT1*, a well-characterized ISG that is also directly transcribed upon PRR stimulation (Supplementary Fig S2A). Studying pIFIT1-GLuc THP-1 in response to these different stimuli paralleled the data obtained by measuring *IFNB* (Fig 1G). However, as observed in the previous set of experiments, the transactivation of the *IFIT1* promoter in response to RNA:DNA hybrid was never as strong as upon dsDNA transfection, also when compared over a broad range of different ligand concentrations (Fig 1H). Interestingly, assessing the production of the antiviral chemokine IP-10 in response to these stimuli revealed a similar plateau for both hybrids as well as dsDNA at higher ligand concentrations, most likely due to the fact that this chemokine is induced by both PRR-dependent as well as type I IFN-dependent mechanisms (Supplementary Fig S1B). Altogether, these data indicated that intracellular RNA:DNA hybrids are indeed sensed by the innate immune system, leading to strong antiviral immune responses.

RNA:DNA hybrid detection is independent of MAVS but requires the cGAS–STING axis

We next focused on the identification of the intracellular receptor that was essential for mediating the RNA:DNA hybrid-induced innate immune response. As RNA:DNA hybrids contain a single polyribonucleotide molecule, we first assessed the involvement of the RLR system as sensors of RNA-DNA hybrids. MAVS constitutes the critical signaling adapter downstream of both RIG-I and MDA-5 and as such, MAVS-deficient cells are devoid of both RIG-I and MDA5 signaling. We made use of the CRISPR/Cas9 gene editing system to knock out MAVS in THP-1 cells. As such, we targeted a critical coding exon to disrupt the reading frame and therefore the expression of MAVS (Supplementary Fig S2B). Transfection of differentiated wild-type (WT) THP-1 cells with RNA:DNA hybrids resulted in increased *IFNB*, and a similar response was also observed in *MAVS* knockout cells (Fig 2A). As expected, MAVS-deficient THP-1 cells failed to respond to 5′-triphosphate RNA, whereas dsDNA-mediated activation of THP-1 cells was not affected.

-125-

Figure 1. Synthetic annealed poly(rA):poly(dT) hybrid induces type I interferon response.

A   Synthetic poly(rA) (lane 1) and poly(dT) (lane 2) oligos were annealed, and the samples were run on a 1% agarose gel. Arrows indicate the annealed poly(rA):poly(dT) (lane 3) RNA:DNA hybrid.

B   Different amounts of synthetic annealed poly(rA):poly(dT) hybrid was spotted on a nitrocellulose membrane and probed with anti-hybrid S9.6 antibody.

C   Poly(rA):poly(dT) was either undigested (1) or digested with DNase I (2), RNase A (3) or RNase H (4), and the digested samples were analyzed on a 1% agarose gel.

D   Human peripheral blood mononuclear cells (PBMCs) were transfected with or without poly(rA), poly(dT) or poly(rA):poly(dT) hybrid or poly(dA):poly(dT) dsDNA for 6 h. RT–PCR was performed to check for the expression of *IFNB*. Figure shows values from two biological replicates. Data plot shows mean with SEM.

E   Human PBMCs from two different donors were transfected with or without poly(rA), poly(dT) or poly(rA):poly(dT) for 16 h, and the amount of secreted IP-10 was measured by ELISA. Each experiment was performed in duplicate. Data plot shows mean with SEM.

F   Differentiated human monocyte THP-1 cells were transfected with or without poly(rA):poly(dT) hybrid, poly(dA):poly(dT) dsDNA and RT–PCR was performed to check for the expression of *IFNB*. Each experiment was performed in duplicate. Data are representative of one of two independent experiments. Data plot shows mean with SEM.

G   Differentiated human monocyte THP-1 cells with Gaussia luciferase knockin under *IFIT1* promoter were transfected with or without poly(rA):poly(dT) hybrid, poly(dA): poly(dT) dsDNA and luciferase assay was performed. Data shown are from two biological replicates. Data plot shows mean with SEM.

H   Differentiated human monocyte THP-1 cells with Gaussia luciferase knockin under *IFIT1* promoter were transfected with different amounts of poly(rA):poly(dT) hybrid or poly(dA):poly(dT) dsDNA and luciferase assay was performed. Data shown are from two biological replicates. Data plot shows mean with SEM.

To study the involvement of the cGAS–STING axis in RNA:DNA hybrid recognition, we next generated THP-1 cells deficient in either STING or cGAS (Supplementary Fig S2C and D). Interestingly, we did not observe any induction of *IFNB* in response to transfection of RNA:DNA hybrids in differentiated *STING* or *cGAS* KO THP-1 cells (Fig 2B–D). As anticipated, dsDNA-mediated *IFNB* induction was also completely abrogated in *cGAS/STING* KO cells, whereas 5′-triphosphate RNA detection was still intact. Similar results were

also observed for IL-6 (Supplementary Fig S3). These data suggested that, next to dsDNA, cGAS also serves as the receptor for cytosolic RNA:DNA hybrids.

RNA:DNA hybrids bind with cGAS and produce cGAMP

To confirm the direct recognition of RNA:DNA hybrids by cGAS, we took advantage of the fact that cGAS-mediated activation can be
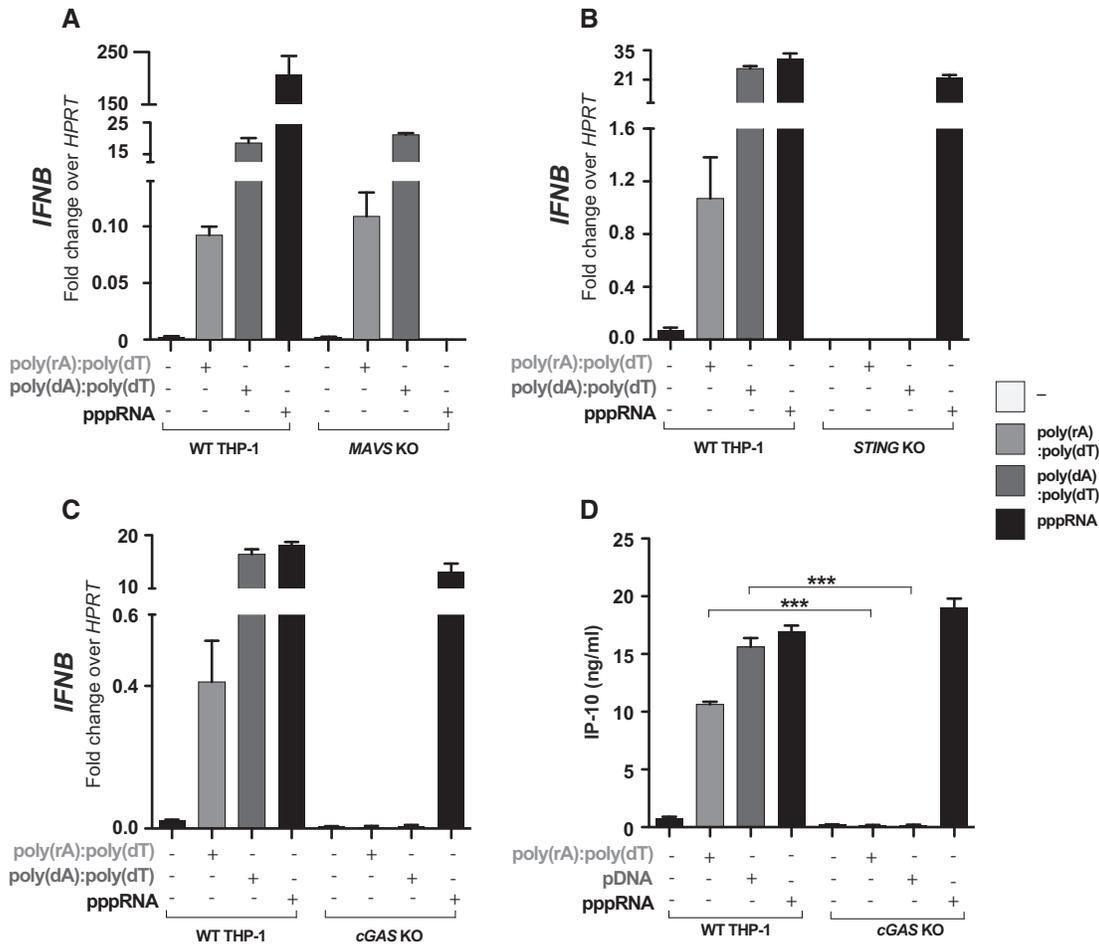
-126-

Figure 2. Poly(rA):poly(dT) RNA:DNA hybrids-induced type I interferon response is cGAS dependent.

A–C   WT THP-1, *MAVS* KO THP-1, *STING* KO THP-1 and *cGAS* KO THP-1 cells were transfected with or without poly(rA):poly(dT) (hybrid), poly(dA):poly(dT) dsDNA or pppRNA for 6 h. RT–PCR was performed to check for the expression of *IFNB*. Data shown are from experimental duplicates representative of two independent experiments. Data plot shows mean with SEM.

D   WT THP-1 cells or *cGAS* KO THP-1 cells were transfected with or without poly(rA):poly(dT) (hybrid), dsDNA or pppRNA for 16 h, and amount of IP-10 secreted was detected by ELISA. Data shown are from experimental triplicate, representative of three independent experiments. Data plot shows mean with SEM.

Data information: ***$P < 0.0001$, one-way ANOVA, Tukey's multiple comparison test.

studied *in vitro*. To this effect, we incubated the single polynucleotides (poly(rA) and poly(dT)) or the synthetic RNA:DNA hybrids (poly(rA):poly(dT)) with recombinant cGAS in the presence of ATP and GTP. Moreover, we included synthetic dsRNA (poly(rA):poly(rU)) and synthetic dsDNA (poly(dA):poly(dT)) as controls. The samples were incubated for 45, 90 or 180 min. Analysis of the RNA:DNA/cGAS sample via HPLC revealed a peak correlating with the expected peak for cGAMP(2′–5′) (Fig 3A). However, no such product was observed upon the incubation of cGAS with dsRNA or any of the single polynucleotide preparations (Fig 3C and Supplementary Fig S4A). Compared to dsDNA, RNA:DNA hybrids induced less cGAMP, also when studying cGAMP production over time (Supplementary Fig S4B). Several groups have recently used shorter length synthetic ssRNA and ssDNA to generate hybrids and to study the immune cell activation by these shorter hybrids. To confirm whether such short hybrids can also activate cGAS–cGAMP pathway, we used 60-bp synthetic poly(rA) ssRNA and poly(dT) ssDNA

oligos and generated RNA:DNA hybrids by annealing them (Fig 4A). To verify whether these shorter oligos were functionally active, we transfected them into pIFIT1-GLuc THP-1 cells and observed a robust secretion of luciferase only in response to hybrids and dsDNA (Fig 4B). As above, incubation of these shorter hybrids with recombinant cGAS clearly resulted in the formation of cGAMP (Fig 4C).

To understand the structural basis for the interaction of cGAS with RNA:DNA hybrids, we generated protein–nucleic acid interaction models *in silico*. Modeling of the complexes of cGAS with either RNA:DNA hybrid, dsDNA or dsRNA revealed that RNA:DNA hybrid could bind in the cleft of cGAS (regardless of the orientation of RNA and DNA strands) exactly in the same mode as dsDNA. At the same time, the dsRNA helix could not be accommodated in this cleft (Fig 5A–D). Structural alignment of dsDNA, RNA:DNA hybrid and dsRNA helixes showed that dsDNA and RNA:DNA hybrid molecules have a similar conformation of their double-stranded helix with similar shapes of their minor and major groves. While there was an
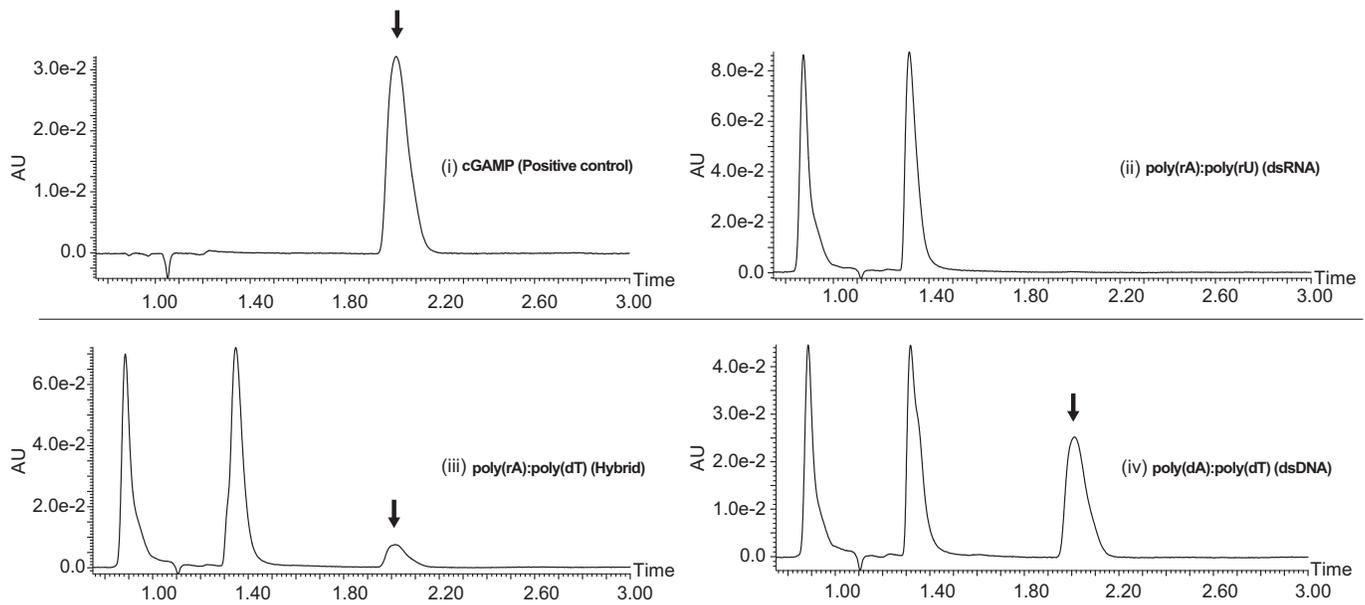
Figure 3.   Recombinant cGAS protein synthesizes cGAMP(2′−5′) in the presence of poly(rA):poly(dT) hybrids.
RP-HPLC chromatograms showing the presence of cGAMP(2′−5′) after incubation of recombinant cGAS protein for 90 min in the presence of ATP and GTP with different ligands. (i) Synthetic 2′−5′ cGAMP as a positive control, (ii) poly(dA):poly(rU) dsRNA, (iii) poly(rA):poly(dT) hybrid and (iv) poly(dA):poly(dT) dsDNA.

obvious overlap between the orientation of the bases of RNA:DNA hybrids and dsDNA, the shape of a dsRNA helix differs from both the dsDNA and hybrid helices with the dsRNA helix being wider (Fig 5E and F). It is likely that this prevents its proper binding into the cGAS cleft and as such the conformational switch of cGAS that is required for its nucleotidyltransferase activity. As RNA:DNA hybrids can bind to cGAS in both orientations of the RNA chain, its binding to cGAS should lead to the same switch-like conformational changes of the activation loop, which are induced by dsDNA binding (Zhang *et al*, 2014) (Supplementary Fig S5). Altogether, these data proved that RNA:DNA hybrids can serve as ligands initiating cGAS activity.

Discussion

The presence of foreign nucleic acids or aberrant formation or translocation of self-nucleic acids triggers an immune response that is dependent on the interaction of these nucleic acids with different cytoplasmic PRRs. While the receptors and the subsequent pathways activated by intracellular dsDNA and dsRNA have been clearly ascertained, it was not established whether cytosolic RNA:DNA hybrids can also initiate immune responses. Here, we report that cytoplasmic RNA:DNA hybrids can indeed induce potent antiviral immune responses. Furthermore, we can ascribe the immunostimulatory activity of this ligand to the direct activation of the recently established dsDNA sensor cGAS.

To address the immunostimulatory activity RNA:DNA hybrids in the first place, it was important to obtain hybrids devoid of dsDNA contamination. A critical confounding factor present when generating RNA:DNA hybrids using reverse transcriptase is the synthesis of dsDNA from the first strand cDNA (Hsieh *et al*, 1993). To check for

this possibility, we digested our *in vitro* transcribed hybrid samples with RNase H; however, we still observed some residual activation suggesting the presence of dsDNA. Consequently, we decided to use enzymatically generated long polynucleotides that could be annealed to form RNA:DNA hybrids. Of course, there is no biological correlate for this artificial duplex molecule; however, we consider it a sound model ligand, given the fact that ligand receptor interactions of double-stranded nucleic acids usually occur independent of their nucleobase composition, but are rather dependent on their tertiary structure.

Our *in vitro* studies using recombinant cGAS in conjunction with different nucleic acids show that the amount of cGAMP produced in response to the hybrids is not within the same range as observed with dsDNA. We hypothesize that this could be due to suboptimal activation of cGAS resulting from an inefficient conformational switch that is required for its enzyme activity. However, in a situation in which multiple signals are integrated by the cell (e.g. IP-10 production), low levels of cGAMP production might be sufficient to reach activity levels that are comparable to dsDNA stimulation. Interestingly, in some cases, we also observed an enhanced IFN response to single-stranded poly(dT) molecule while no such response was observed for poly(rA). However, at the same time, our *in vitro* data clearly demonstrate that poly(dT) by itself cannot activate cGAS. As such, we consider it likely that this weak, but consistent cGAS activation by poly(dT) *in vivo* is due to the intracellular formation of a cGAS stimulatory ligand (Supplementary Fig S6). While this phenomenon might not directly involve poly(dT) as a ligand, we speculate that poly(dT) can form DNA:RNA hybrids with endogenous poly(rA) in the context of mRNA tails.

Aberrant activation of the immune system in response to self-nucleic acids forms the basis for several chronic inflammatory
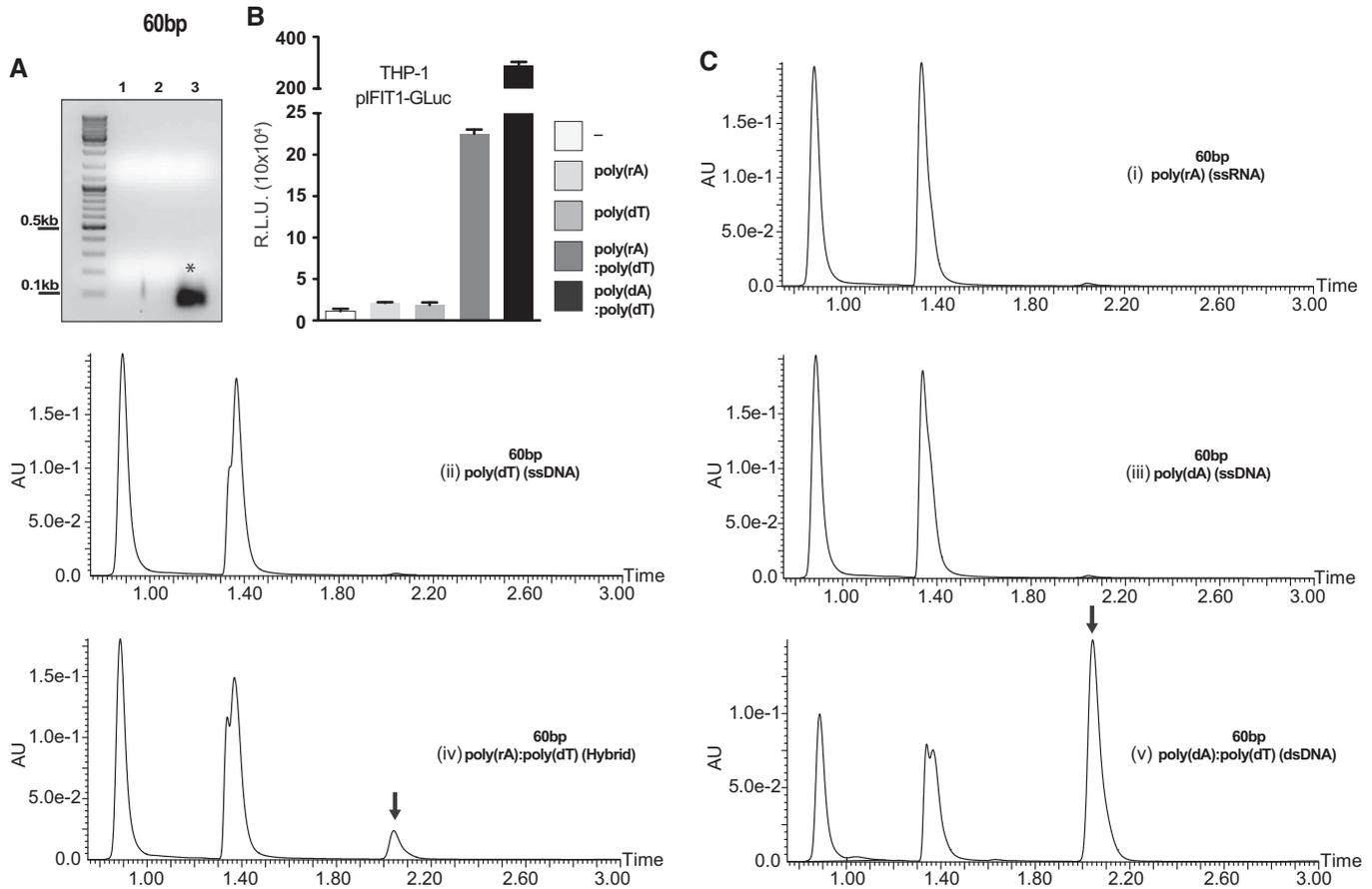
-128-

Figure 4.   Recombinant cGAS protein synthesizes cGAMP(2′–5′) in the presence of 60-bp synthetic poly(rA):poly(dT) hybrids.

A   60-bp synthetic poly(rA) (lane 1) and 60-bp poly(dT) (lane 2) oligos were annealed, and the samples were run on a 1% agarose gel. Asterisk indicates annealed poly(rA):poly(dT) RNA:DNA hybrid (lane 3).

B   Differentiated human monocyte THP-1 cells with Gaussia luciferase knockin under *IFIT1* promoter were transfected with or without 60-bp poly(rA), 60-bp poly(dT) or 60-bp poly(rA):poly(dT) (hybrid) and 60-bp poly(dA):poly(dT) (dsDNA) and luciferase assay was performed. Data shown are from two biological replicates. Data plot shows mean with SEM.

C   RP-HPLC chromatograms showing the presence of cGAMP(2′–5′) after incubation of recombinant cGAS protein for 180 min in the presence of ATP and GTP with: (i) 60-bp poly(rA) (ssRNA), (ii) 60-bp poly(dT) (ssDNA), (iii) 60-bp poly(dA) (ssDNA), (iv) 60-bp poly(rA):poly(dT) (hybrid), or (v) 60-bp poly(dA):poly(dT) (dsDNA). Arrow indicates the expected peak for cGAMP (2′–5′), the product of cGAS catalyzed reaction.

diseases. Aicardi–Goutières syndrome (AGS), for example, is a rare but generally fatal condition associated with increased type I interferon production in the serum of respective patients. This condition can result from mutations in genes encoding for the nucleic acid metabolizing enzymes SAMHD1, TREX1, Ribonuclease H2 (RNASE H2A, RNASE H2B, RNASE H2C) or ADAR1 (La Piana *et al*, 2014; Lee-Kirsch *et al*, 2014). Even though it is quite conceivable that AGS-related mutations lead to the activation of cytosolic nucleic acid sensing pathways, this has only been documented for TREX1 deficiency. Here, it has been shown *in vivo* that the fatal consequences of TREX1 deficiency can be rescued by the deletion of STING, and *in vitro* studies have revealed that this is fully attributable to the activation of cGAS (Gall *et al*, 2012; Ablasser *et al*, 2014). However, the exact nature of the ligand that activates cGAS in this context still remains unclear. Interestingly, TREX1 has been shown to metabolize the reverse-transcribed DNA, and the absence of TREX1 resulted in the accumulation of DNA from endogenous retro elements (Stetson, 2012). On the other hand, the ribonuclease H2 enzyme complex is involved in the removal of the Okazaki fragments as well as aberrantly introduced ribonucleotides during the replication of the genome (Reijns *et al*, 2012). As such, the hypomorphic mutations in RNase H2 proteins, observed in AGS, could result in less efficient removal of the RNA strand or the incorporated ribonucleotides, thereby culminating in the accumulation of RNA:DNA hybrids. Unfortunately, the *Rnaseh2b* knockout mouse model does not display the expected phenotype of type I IFN-driven immune activation, which might be due to the fact that this condition does not lead to a cell-autonomous activation of nucleic acid sensing PRRs. A secondary activation of immune cells in the context of RNase H2 deficiency might, however, not be evident due to early embryonic lethality of these mice (Hiller *et al*, 2012; Reijns *et al*, 2012). As such, it would be worth revisiting this model with regard to identifying the culprit stimulatory ligand using conditionally ablated *Rnaseh2b*.

-129-
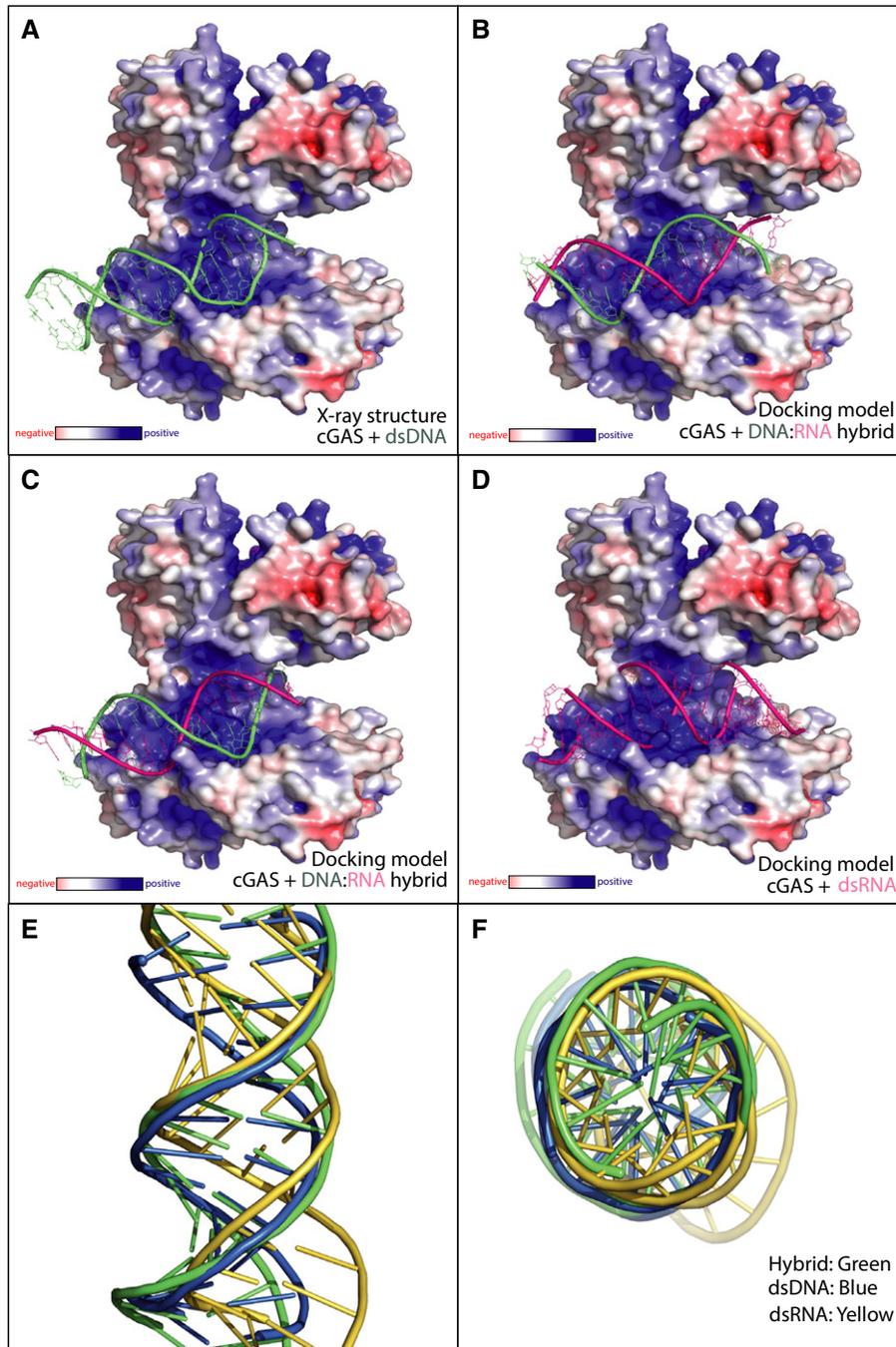
Figure 5. Modeling of dsDNA, RNA:DNA hybrids and dsRNA interaction with cGAS.

A–D HADDOCK models of different double-stranded nucleic acids in the crystal structure of pig cGAS are shown (the cGAS structure is based on the cGAS:dsDNA PDB ID 4KB6). The DNA strand is shown in green, whereas the RNA strand is depicted in pink. (A) dsDNA bound to cGAS (PDB ID 4KB6). (B, C) Best docking results of an RNA:DNA hybrid in two different orientations (PDB ID 4KB6 chain A) into cGAS are depicted. Hybrid molecules have been located in the cleft formed by the Zn-thumb and Arg150 of cGAS. (D) The best docking solution of dsRNA into the dsDNA binding region of cGAS.

E, F Cartoon and surface representation of superposition of dsDNA (blue), DNA:RNA hybrid (green) and dsRNA (yellow) molecules based on published structures.

Retroviruses could provide another source of intracellular RNA:DNA hybrids, which are generated by reverse transcriptase in the course of their replicative cycle (Telesnitsky & Goff, 1997). In this context, it was recently shown that viral-derived RNA:DNA hybrids accumulate in the cytoplasm and endosomal compartments of retrovirus (MMLV)-infected fibroblasts (Rigby *et al*, 2014). In the same study, the authors used synthetic 45- and 60-bp RNA:DNA hybrid molecules as stimuli and observed that while the response to the 45-bp hybrid molecule was completely TLR9 dependent, it was only partially dependent in the case of 60-bp hybrids. As such, to

-130-

account for the partial response still observed in the 60-bp hybrid-stimulated cells, the authors alluded to the existence of other cytoplasmic hybrid sensors. Given the restricted expression of TLR9 in the human system, we speculate that RNA:DNA hybrid detection by cGAS could constitute the predominant sensing mechanism in the human system. Indeed, the fact that retroviral infection has been shown to trigger proinflammatory gene expression in cells other than pDCs indicates the existence of additional cytoplasmic receptors for retroviral infection (Luban, 2012). In this regard, under permissive conditions, cGAS has been identified as the key intracellular sensor essential for mounting an antiviral response in the course of HIV infection in various cell types, including fibroblasts, myeloid dendritic cells and macrophages (Gao *et al*, 2013a). Here, it was demonstrated that inhibition of viral reverse transcriptase but not of viral integrase resulted in significant inhibition of IFN-β production. Unfortunately, this set of experiments cannot distinguish between the presence of RNA:DNA hybrids versus dsDNA as there is currently no technical means to differentially influence the formation of such nucleic acids during *in vivo* infections. Consequently, further studies will be required to identify the cGAS-stimulatory nucleic acid in the context of retroviral infection.

Interestingly, the presence of RNA:DNA hybrids has also been described in the life cycle of a number of DNA viruses, including herpesviridae (Prichard *et al*, 1998; Rennekamp & Lieberman, 2011) and hepadnaviruses (Summers & Mason, 1982; Miller *et al*, 1984). Moreover, mitochondrial nucleic acids, which have previously been described as DAMP molecules (Oka *et al*, 2012), harbor stable RNA:DNA hybrid structures that arise in the course of mtDNA replication (Yasukawa *et al*, 2006). Again, additional studies will be needed to decipher the possible contribution of RNA:DNA hybrids to the immunostimulatory capacities of these infectious or sterile inflammatory conditions.

In conclusion, we demonstrate that the intracellular presence of RNA:DNA hybrids elicits cGAS-dependent antiviral immune responses. Future studies will be required to identify the biological context of this sensing modality, yet as outlined above, we hypothesize that this mechanism could be of relevance in the context of aberrant recognition of endogenous nucleic acids or viral infections.

Materials and Methods

Generation of RNA:DNA hybrids

Equal amounts of synthetic poly(dT) (Sigma #P6905) and poly(rA) (Sigma # P9403) were mixed in 5× annealing buffer (50 mM Tris–HCl pH 7,6, 250 mM NaCl, 5 mM EDTA), for example, 40 μl poly(rA) (300 ng/μl) + 40 μl poly(dT) (300 ng/μl) in 20 μl 5× annealing buffer. The samples were incubated in a PCR block starting at 95°C with decreasing gradient of −1°C per 100 seconds until 20°C. 60-bp poly(rA) was purchased from Biomers. 60-bp poly(dA) and 60-bp poly(dT) were purchased from Integrated DNA Technologies.

Dot-blot for hybrids

Nitrocellulose membrane was marked and spotted with the RNA:DNA hybrids or poly(dA:dT) (dsDNA) as a negative control. The membrane was allowed to air dry for 2 h and then blocked in 0.5%

milk in PBST buffer. The membrane was then incubated with the anti-hybrid S9.6 antibody overnight at 4°C, washed and probed with anti-mouse secondary antibody.

Cell culture

Murine BMDMs were isolated and cultured as per standard protocols. THP-1 cells were maintained in RPMI medium. The cells were differentiated for 3 h with PMA (330 ng/ml), the medium was removed, and the cells were washed with 1× PBS three times. The cells were detached in PBS, and 7×10^4 cells were plated in 96-well flat-bottomed plate. Isolation of PBMC from buffy coat was performed using standard protocol.

Stimulation with RNA:DNA hybrids

7×10^4 cells were incubated overnight in flat-bottomed 96-well plate. A transfection mix with Lipofectamine and 1 μg/ml of poly(rA), poly(dT), poly(rA):poly(dT) was prepared as per standard protocol. The cells were incubated with the nucleic acids overnight. The supernatants were collected and subjected to further assays. For RT–PCR, the cells were stimulated with the nucleic acids for 6 h.

RNase H, RNase A and DNase I digestion

1 μg of hybrid was digested with 1 μl of the RNase H, RNase A or DNase I (Fermentas) enzyme in a final reaction volume of 10 μl. The sample was incubated at 37°C for 30 min. The enzyme was heat-inactivated by incubating the samples at 65°C for 10 min, and the efficiency of digestion was checked by running the samples in 0.8% agarose gel.

RNA analysis

RNA isolation was performed using the QIAcube system (Qiagen), and equal amounts of RNA were used for cDNA synthesis. Real-time PCR analysis with Eva Green PCR Master Mix (Biobudget) was performed on Real-time PCR system (Roche). The expression of target genes was normalized to HPRT or GAPDH expression and plotted as arbitrary units on a linear scale. Primer sequences are available on request.

ELISA

Cell culture supernatants were assayed for human IP-10 and human IL-6 (BD Biosciences) according to the manufacturer's instructions.

STING, *MAVS* and *cGAS* KO THP-1 cells

THP-1 knockout cells were generated using the CRISPR/Cas9 system. A plasmid containing mCherry Cas9 and a U6 promoter-driven gRNA against STING, MAVS or cGAS was electroporated into THP-1 cells under following conditions: 5 μg plasmid were mixed with 250 μl cell suspension (10×10^6/ml in OptiMEM) and electroporated at 950 μF and 250 V. Cells were FACS sorted for mCherry expression 24 h after electroporation. Positive cells were

diluted under limiting conditions and plated in 96-well plates to obtain single cell clones. The genotype of THP-1 clones was analyzed by deep sequencing (Illumina, MiSeq) (Schmid-Burgk *et al*, 2014).

Generation of pIFIT1-GLuc THP-1 reporter cell line

To study induction of interferon-stimulated genes at the natural gene expression level using a reporter system, we generated a IFIT1 reporter cell line. Using the CRISPR/Cas9 system, the Gaussia luciferase gene was knocked into the *IFIT1* gene locus in THP-1 cells. For this purpose, a reporter insert cassette, harboring flanking 50-bp homology arms, a 2A peptide and the luciferase gene, was synthesized by a two-step PCR. THP-1 cells were co-electroporated with a CRISPR targeting the *IFIT1* locus and the Gaussia luciferase insert cassette. As the CRISPR construct contained a mCherry gene, electroporated cells were sorted for mCherry expression by FACS 24 h after electroporation. For isolation of reporter cell clones, cells were seeded under limiting dilution conditions. Single cell clones were tested for knockin of the luciferase gene by stimulation with synthetic triphosphate RNA (pppRNA) and subsequent measurement of luciferase induction. Positive reporter clones were validated by sequencing of the insert locus.

In vitro assay for cGAS activity

For *in vitro* reactions of cGAS in the presence of varying nucleic acids, 2 mM recombinant cGAS was mixed with 0.2 μg/μl poly(rA): poly(dT), poly(dA):poly(dT), poly(rA):poly(rU) or the corresponding single-stranded controls and 0.1 M ATP and 0.1 mM GTP in buffer A (100 mM NaCl, 40 mM Tris pH 7.5, 10 mM MgCl$_2$). After 45, 90 or 180 min of incubation at 37°C, the reaction mixture was analyzed by RP-HPLC. Samples were prepared in 0.3 M triethyl-ammonium acetate, applied to a Waters XBridge C18 column (4.6 × 50 mm, 2.5 μm particle size) and separated in an isocratic gradient of 100 mM ammonium acetate for 5 min at a flow rate of 1 ml/min. Chemically synthesized cGAMP(2′–5′) (Ablasser *et al*, 2013a) served as a positive control.

Modeling of cGAS dsDNA, DNA:RNA hybrid and dsRNA complexes

Complexes of cGAS with three short double-stranded nucleic acids (NA) were modeled by docking of corresponding NA (dsDNA (PDB ID 4KB6) (Civril *et al*, 2013) or (PDB ID 3V6T) (Deng *et al*, 2012), DNA:RNA hybrid (PDB ID 4GG4) (Yin *et al*, 2012) and dsRNA (PDB ID 3KS8) (Kimberlin *et al*, 2010) into the structure of pig cGAS (PDB ID 4KB6 chain A) (Civril *et al*, 2013)). Docking was done on the HADDOCK server (de Vries *et al*, 2010). HADDOCK permits position restrained docking of NA into proteins. Sets of residues for docking in cGAS and dsDNA that should interact with NA have been taken from the original crystal structure of the cGAS–dsDNA complex (PDB ID 4KB6) (Civril *et al*, 2013). Visualization of the results was performed using PyMol.

Statistics

Statistical tests were performed using GraphPad Prism program. For column statistics, one-way ANOVA with Bonferroni's or Tukey's multiple comparison test was performed. $P \leq 0.05$ was considered significant.

Author contributions

AKM and VH planned the experiments. AKM, TS, DC and MG performed the experiments. KH and MG generated the pIFIT1-GLuc THP-1 cells. LA purified recombinant cGAS. K-PH and AVK performed the *in silico* structural analysis. AKM and VH drafted the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Ablasser A, Goldeck M, Cavlar T, Deimling T, Witte G, Rohl I, Hopfner KP, Ludwig J, Hornung V (2013a) cGAS produces a 2′–5′-linked cyclic dinucleotide second messenger that activates STING. *Nature* 498: 380–384

Ablasser A, Hertrich C, Wassermann R, Hornung V (2013b) Nucleic acid driven sterile inflammation. *Clin Immunol* 147: 207–215

Ablasser A, Schmid-Burgk JL, Hemmerling I, Horvath GL, Schmidt T, Latz E, Hornung V (2013c) Cell intrinsic immunity spreads to bystander cells via the intercellular transfer of cGAMP. *Nature* 503: 530–534

Ablasser A, Hemmerling I, Schmid-Burgk JL, Behrendt R, Roers A, Hornung V (2014) TREX1 deficiency triggers cell-autonomous immunity in a cGAS-dependent manner. *J Immunol* 192: 5993–5997

Barbalat R, Ewald SE, Mouchess ML, Barton GM (2011) Nucleic acid recognition by the innate immune system. *Annu Rev Immunol* 29: 185–214

Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, Patterson WL, Carrico RJ (1986) Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J Immunol Methods* 89: 123–130

Civril F, Deimling T, de Oliveira Mann CC, Ablasser A, Moldt M, Witte G, Hornung V, Hopfner KP (2013) Structural mechanism of cytosolic DNA sensing by cGAS. *Nature* 498: 332–337

Crow YJ (2011) Type I interferonopathies: a novel set of inborn errors of immunity. *Ann N Y Acad Sci* 1238: 91–98

de Vries SJ, van Dijk M, Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* 5: 883–897

Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, Shi Y, Yan N (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* 335: 720–723

Diner EJ, Burdette DL, Wilson SC, Monroe KM, Kellenberger CA, Hyodo M, Hayakawa Y, Hammond MC, Vance RE (2013) The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. *Cell Rep* 3: 1355–1361

Gall A, Treuting P, Elkon KB, Loo YM, Gale M Jr, Barber GN, Stetson DB (2012) Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease. *Immunity* 36: 120−131

Gao D, Wu J, Wu YT, Du F, Aroh C, Yan N, Sun L, Chen ZJ (2013a) Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science* 341: 903−906

Gao P, Ascano M, Wu Y, Barchet W, Gaffney BL, Zillinger T, Serganov AA, Liu Y, Jones RA, Hartmann G, Tuschl T, Patel DJ (2013b) Cyclic [G(2′,5′)pA(3′,5′)p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. *Cell* 153: 1094−1107

Goubau D, Deddouche S, Reis e Sousa C (2013) Cytosolic sensing of viruses. *Immunity* 38: 855−869

Hiller B, Achleitner M, Glage S, Naumann R, Behrendt R, Roers A (2012) Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome integrity. *J Exp Med* 209: 1419−1426

Hsieh JC, Zinnen S, Modrich P (1993) Kinetic mechanism of the DNA-dependent DNA polymerase activity of human immunodeficiency virus reverse transcriptase. *J Biol Chem* 268: 24607−24613

Hu Z, Zhang A, Storz G, Gottesman S, Leppla SH (2006) An antibody-based microarray assay for small RNA detection. *Nucleic Acids Res* 34: e52

Kimberlin CR, Bornholdt ZA, Li S, Woods VL Jr, MacRae IJ, Saphire EO (2010) Ebolavirus VP35 uses a bimodal strategy to bind dsRNA for innate immune suppression. *Proc Natl Acad Sci USA* 107: 314−319

La Piana R, Uggetti C, Olivieri I, Tonduti D, Balottin U, Fazzi E, Orcesi S (2014) Bilateral striatal necrosis in two subjects with Aicardi-Goutieres syndrome due to mutations in ADAR1 (AGS6). *Am J Med Genet A* 164: 815−819

Lee-Kirsch MA, Wolf C, Gunther C (2014) Aicardi-Goutieres syndrome: a model disease for systemic autoimmunity. *Clin Exp Immunol* 175: 17−24

Luban J (2012) Innate immune sensing of HIV-1 by dendritic cells. *Cell Host Microbe* 12: 408−418

Medzhitov R (2007) Recognition of microorganisms and activation of the immune response. *Nature* 449: 819−826

Miller RH, Tran CT, Robinson WS (1984) Hepatitis B virus particles of plasma and liver contain viral DNA-RNA hybrid molecules. *Virology* 139: 53−63

Oka T, Hikoso S, Yamaguchi O, Taneike M, Takeda T, Tamai T, Oyabu J, Murakawa T, Nakayama H, Nishida K, Akira S, Yamamoto A, Komuro I, Otsu K (2012) Mitochondrial DNA that escapes from autophagy causes inflammation and heart failure. *Nature* 485: 251−255

Prichard MN, Jairath S, Penfold ME, St Jeor S, Bohlman MC, Pari GS (1998) Identification of persistent RNA-DNA hybrid structures within the origin of replication of human cytomegalovirus. *J Virol* 72: 6997−7004

Reijns MA, Rabe B, Rigby RE, Mill P, Astell KR, Lettice LA, Boyle S, Leitch A, Keighren M, Kilanowski F, Devenney PS, Sexton D, Grimes G, Holt IJ, Hill RE, Taylor MS, Lawson KA, Dorin JR, Jackson AP (2012) Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* 149: 1008−1022

Rennekamp AJ, Lieberman PM (2011) Initiation of Epstein-Barr virus lytic replication requires transcription and the formation of a stable RNA-DNA hybrid molecule at OriLyt. *J Virol* 85: 2837−2850

Rigby RE, Webb LM, Mackenzie KJ, Li Y, Leitch A, Reijns MA, Lundie RJ, Revuelta A, Davidson DJ, Diebold S, Modis Y, Macdonald AS, Jackson AP (2014) RNA:DNA hybrids are a novel molecular pattern sensed by TLR9. *EMBO J* 33: 542−558

Schmid-Burgk JL, Schmidt T, Gaidt MM, Pelka K, Latz E, Ebert TS, Hornung V (2014) OutKnocker: a web tool for rapid and simple genotyping of designer nuclease edited cell lines. *Genome Res* 24: 1719−1723

Stetson DB (2012) Endogenous retroelements and autoimmune disease. *Curr Opin Immunol* 24: 692−697

Summers J, Mason WS (1982) Replication of the genome of a hepatitis B−like virus by reverse transcription of an RNA intermediate. *Cell* 29: 403−415

Sun L, Wu J, Du F, Chen X, Chen ZJ (2013) Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* 339: 786−791

Telesnitsky A, Goff SP (1997) Reverse transcriptase and the generation of retroviral DNA. In *Retroviruses*, Coffin JM, Hughes SH, Varmus HE (eds). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Wu J, Sun L, Chen X, Du F, Shi H, Chen C, Chen ZJ (2013) Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science* 339: 826−830

Yasukawa T, Reyes A, Cluett TJ, Yang MY, Bowmaker M, Jacobs HT, Holt IJ (2006) Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J* 25: 5358−5371

Yin P, Deng D, Yan C, Pan X, Xi JJ, Yan N, Shi Y (2012) Specific DNA-RNA hybrid recognition by TAL effectors. *Cell Rep* 2: 707−713

Zhang X, Shi H, Wu J, Sun L, Chen C, Chen ZJ (2013) Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol Cell* 51: 226−235

Zhang X, Wu J, Du F, Xu H, Sun L, Chen Z, Brautigam CA, Chen ZJ (2014) The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. *Cell Rep* 6: 421−430

-133-

Supplementary Information

Supplementary Figure 1: Synthetic annealed poly(rA):poly(dT) hybrid induces type I interferon response.

a) Bone marrow derived macrophages (BMDM) were transfected with or without poly(rA), poly(dT) or poly(rA):poly(dT) for 6 hrs. RT-PCR was performed to check for the expression of *Ifn-α, Ifn-β, Il-6* and *Ip-10.* **b)** Differentiated Human monocyte THP-1 cells were transfected with different amounts of poly(rA):poly(dT) (hybrid) or poly(dA):poly(dT) dsDNA and IP-10 in supernatants was measured by ELISA. In all figures error bars represent SEM.

Supplementary Figure 2: Scheme representing the coding region of IFIT1, MAVS, STING and cGAS targeted by the CRISPR/Cas9 system

Depicted is a schematic representation of the genomic loci of human IFIT1, MAVS, STING and cGAS genes. Small black and large grey squares represent non-coding and coding exons, respectively. The red arrow highlights the location of the CRISPR target site. The sequences under the genomic loci are a magnification of the targeted region. The top sequences are the reference sequences, whereby the red letters highlight the protospacer (CRISPR binding site) and orange letters the PAM (protospacer adjacent motif). For MAVS, STING and cGAS the sequences below the reference sequence show the mutations that were detected by deep sequencing.

Supplementary Figure 3: cGAS-dependent secretion of IL-6

WT THP-1 cells or cGAS KO THP-1 cells were transfected with or without poly(rA):poly(dT) (hybrid), pDNA or pppRNA for 16hrs and amount of IL-6 secreted was detected by ELISA.

Supplementary Figure 4: Single stranded homopolymeric polynucleotides do not activate cGAS

a) RP-HPLC chromatograms showing the synthesis of cGAMP(2′-5′) after incubation of recombinant cGAS protein for 90 minutes in the presence of

1

ATP and GTP with different ligands (i) with poly(rA), (ii) with poly(dT), (iii) poly(dA) and (iv) poly(rU). **b)** Synthetic dsDNA, dsRNA and RNA:DNA hybrids were incubated for 45, 90 or 180 minutes with recombinant cGAS and the synthesis of 2' 5' cGAMP checked by RP-HPLC. In order to normalize for the total amount of sample loaded the % of AUC was calculated for each time point and each sample.

Supplementary Figure 5: cGAS "activation loop" upon dsDNA and RNA:DNA hybrid binding

Structure of cGAS bound to dsDNA (**a**) and RNA:DNA hybrid in two modes (**b**) and (**c**). The DNA strand is shown in green, whereas the RNA strand is depicted in light pink. The "activation loop" of cGAS is highlighted in blue.

Supplementary Figure 6: Induction of Type I IFNs by poly(dT) is cGAS dependent. WT THP-1 cells or cGAS KO THP-1 cells were transfected with or without poly(rA), poly(dT) and poly(rA):poly(dT) (hybrid) for 16hrs and amount of IP-10 secreted was detected by ELISA.

Supplementary Figure 1

A

BMDM

Ifn-α

Ifn-β

Il-6

Ip-10

Fold change over *Hprt*

- –
- poly(rA)
- poly(dT)
- poly(rA):poly(dT)
- pDNA

B

THP-1

IP-10 (ng/ml)

0.0 6.25 12.5 25 50 100 200 400 (ng)

- poly(rA):poly(dT)
- poly(dA):poly(dT)

Supplementary Figure 2

A human IFIT1 genomic locus

CAGAGATCGCATACCCAGCGCTGGATTCAGGGTTTTCAGGGATAGCGGACCTACCGAAATTTGGAAGCAAAGAAAATGGCTTGTGATTTTTTGTGGCTAA
CTGAAAACCCTGAATCCAGCGCTGGGTATGCGATCTCTGCCTATCGCCTGGATGGCTTTAAATTAGCCACAAAAAATCACAAGCCATTTTCTTTGCTTCC

homology arm   2A   Gaussia luciferase gene   homology arm

B human MAVS genomic locus
Exon3

TGGGTGGAG**TACTTCATTGCGGCACTGA**GGGGCTGTGAGCTAGTTGATCTCGCGGACGAAGTGGCCTCT   REF
TGGGTGGAGTACTTCATTGCGGCACTG--------TGAGCTAGTTGATCTCGCGGACGAAGTGGCCTCT   -Del

C human STING genomic locus
Exon3

ATGCCCCACTCCAGCCTGCA**TCCATCCATCCCGTGTCCCA**GGGGTCACGGGGCCCAGAAGGCAGCCTTGG   REF
ATGCCCCACTCCAGCCTGCATCCATCCATCCCGTGG-----GGGTCACGGGGCCCAGAAGGCAGCCTTGG   -5nt
ATGCCCCACTCCAGCCTGCATCCATCCATCCCGTGTCCA-GGGGTCACGGGGCCCAGAAGGCAGCCTTGG   -1nt

D human cGAS genomic locus
Exon1

CGACTTCCTGGCGGGGGCC**GAACTTTCCCGCCTTAGGCA**GGGCGGCCTCGGGGGCAGCCGGAGACTCGGTG   REF
CGACTTCCTGGCGGGGGCCGAACTTTCCCGCCTT-GGCAGGGCGGCCTCGGGGGCAGCCGGAGACTCGGTG   -1nt
CGACTTCCTGGCGGGGGCCGAACTTTCCCGCCTTAG-CAGGGCGGCCTCGGGGGCAGCCGGAGACTCGGTG   -1nt

Supplementary Figure 3

A

B

A — x-ray structure
cGAS + dsDNA

B — docking model
cGAS + DNA:RNA hybrid

C — docking model
cGAS + DNA:RNA hybrid

Supplementary Figure 6

4. A fluorescent cyclic dinucleotide and its use in methods of identifying substances having an ability to modulate the cGAS/STING pathway (Patent)

Hopfner K.-P., <u>Andreeva L.</u>, Drexler D.J. (Filed 24 July 2017) "A fluorescent cyclic dinucleotide and its use in methods of identifying substances having an ability to modulate the cGAS/STING pathway." European Patent Application EP 17182689.4.

Summary

This work introduces a novel fluorescence-based high-throughput method for measuring cGAS activity and STING binding using a commercially available fluorescent ATP analogue 2-aminopurine-riboside-5'-triphosphate. A growing knowledge of cGAS-STING pathway involvement in autoinflammatory diseases and cancer and successful trials on cGAMP as immune adjuvant for vaccination and anti-cancer therapy make cGAS-STING pathway a promising target for drug development. This makes a method development for easy and fast cGAS-STING activity assay an important direction of research. Previously reported methods for cGAS activity measurement, though having their advantages, are either low-throughput and require expensive and not easily accessible equipment, or do not precisely discriminate specific 2'3'-cGAMP from other CDNs. Among others, cGAS activity assays based on radiolabeled ATP and GTP analogues or cGAMP quantitative measurements using anion-exchange chromatography, nuclear magnetic resonance spectroscopy (NMR) or mass spectrometry (MS) are either not safe, or time-consuming and require several purification steps. Fluorescence-based assay based on fluorescence quenching of cyclic dinucleotide (CDN) 3'3'-cG(d2AP)MP – a fluorescent analogue of 3'3'-cGAMP – by its dimerization with other CDNs was proven to be a high-throughput non-toxic method for 3'3'-cGAMP and c-di-GMP detection, but failed to recognize cGAS product 2'3'-cGAMP. Other high-throughput methods utilizing a fluorescently labeled riboswitch or an antibody against 2'3'-cGAMP were either cross-reactive with 2'3'- and 3'3'-cGAMP isoforms, or did not allow a continuous measurement. A suitable assay for STING activity *in vitro* is an even harder to find, since STING is a receptor and not an enzyme. In this work we describe a novel assay for a reproducible time-resolved cGAS activity measurement and a method for STING binding studies using fluorescent CDN analogue fGAMP. This technique allows simultaneous measurements of cGAS and STING activity using a conventional fluorescence plate reader, is non-toxic and not limited by reaction conditions. Our assay is well-tailored for identifying cGAS and STING agonists and inhibitors in a high-throughput manner and is perfectly suitable for clinical applications.

Author contribution

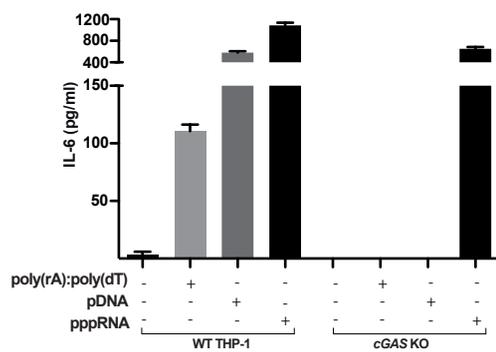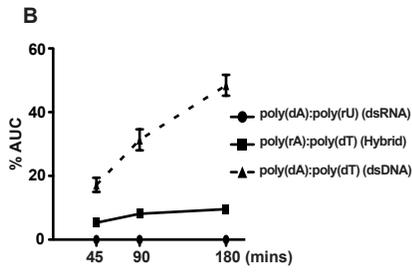The author of this thesis together with Prof. Dr. K.-P. Hopfner and D.J. Drexler invented a novel high-throughput fluorescence-based method for studying cGAS activity and STING binding. She established and optimized its application for cGAS activity measurements with several cGAS constructs in a range of different conditions. She confirmed its reliability comparing the novel assay with the previously established radiolabeled cGAS activity measurement technique. She also evaluated limitations of the novel method and a possibility of its application to study cGAS activity-modulating factors. Moreover, together with D.J. Drexler she tested another application of this method for characterization of STING ligands' affinity and proved its effectiveness for studying various STING-binding molecules.

Discussion

1. Methods for cGAS activity measurement

cGAS has a great importance in innate immune recognition of various viruses and intracellular bacteria, plays an emerging role in recognition of self-DNA and as involved in autoimmune diseases and cancer development. This makes a search for precise and high-throughput method for cGAS activity measurement of huge interest. The first discoveries of cGAS were made using incorporation of radioactively labeled ATP or GTP into the products of cGAS reaction and separation of these products by thin layer chromatography [235, 239, 240]. This method allows sensitive and direct measurement of cGAMP synthesis rate using small reaction volumes, however, its precision is limited by decreasing volume of the reaction between the time points, asynchronous termination of the reaction and variation in exposure time between experiments, making it difficult to compare cGAS activities obtained in different experiments. In different variations of this method reaction can be either stopped by direct plotting on the cellulose plate or by preceding ultrafiltration of the sample that might generate radioactive aerosols and additionally limits the precision of the assay by uncontrollable loss of the reaction material at the transferring step. Moreover, radioactivity assays are limited by the number of reactions that can be performed simultaneously and time-resolved kinetics performed using this method is rather time-consuming.

Another *in vitro* method developed to study cGAS activity was introduced by P. Li's group and is based on anion exchange chromatography of cGAS reaction mixture on MonoQ column (GE Healthcare) [314]. In contrast to radiolabeled assays, in which the resolution and interpretation of side products of reaction are difficult to achieve, anion exchange chromatography allows a good separation of all the nucleotides and quantitative analysis of reaction products using UV absorption. Another advantage of the assay includes its application without additional labeling of the nucleotides and its nontoxicity. However, with this method the samples can not be measured simultaneously and variable loss of material by previous ultrafiltration and transfer to the column may occur. Similar to radiolabeled assays, anion exchange chromatography of cGAS products is a time-consuming method that is difficult to use for extensive screening. Moreover, the variation of reaction condition is limited by its influence on nucleotide binding to the column. Some reaction conditions can result in a weak interaction of reaction products with the column and an unsteady separation of the products thus limiting the method application for comparison of different cGAS reaction conditions.

Two more methods of cGAS activity measurement without a need for labeling were suggested by J. Hall and colleagues [312, 594]. The first of them is based on fluorescence polarization of Cy5-labeled cGAMP in combination with an antibody [594]. Non-fluorescent cGAMP synthesized by cGAS was shown to compete for the antibody binding with its fluorescent analogue leading to a decrease in polarization signal. Though the antibody did not show a cross reactivity to ATP, GTP, c-di-AMP and c-di-GMP, the studies of its affinity to 2'3'- and 3'3'-cGAMP would be important. Given that the proposed antibody can be purchased commercially, this method is advantageous for a large-scale analysis of cGAS activity *in vitro*, but was not proved for measuring cGAMP concentrations in cell extracts. Moreover, it is not well-suited for continuous time-resolved measurements of cGAS activity, because the reaction must be stopped before the addition of Cy5-labeled cGAMP and the antibody. However, the development of an anti-

cGAMP antibody may be of interest for other applications like STING inhibition and used for medical research.

The second method utilizes surface plasmon resonance at Biacore with serpentine flow connecting cells with immobilized cGAS-DNA complex and STING [312]. In this set-up injection of ATP and GTP resulted in production of cGAMP that could be detected by resonance in STING-containing cell. This method of an indirect measurement of cGAS activity was successfully used to determine K_M values of the reaction, as well as to study substrate inhibition and the mechanism of cGAS catalysis, since it allows to overcome a competition for cGAS active site between the substrates and linear or side products of the reaction by applying a constant flow that counteracts the accumulation of the reaction intermediates and side products. However, though a high precision of this method allows to determine a range of kinetic parameters of the reaction, its flexibility in using different cGAS constructs and DNA ligands is limited, since each of the constructs must be separately coupled for the channel. Moreover, a coupling of cGAS-DNA complex may influence complex arrangement and stability and interfere with cGAS dimerization and oligomerization. Furthermore, this method does not allow a direct measurement of cGAS activity and is limited by specific STING-mediated detection of cGAMP leaving analysis of other reaction products for complementary methods like NMR or MS. The latter techniques are also broadly used to study cGAS activity and have an advantage of quantitative and structural analysis of the reaction mixtures. For example, NMR and MS were used for the first discovery of cGAS product composition [238-240, 250]. Nevertheless, these specialized techniques require high sample purity and are commonly used after purification of the products by HPLC. This increases the number of steps and complexity of cGAMP measurement decreasing the precision of cGAMP quantitative measurement. Moreover, these techniques are time consuming and complex and are not always easily accessible. Unlike the *in vitro* methods for cGAS activity measurement described above, NMR and MS are broadly used to analyze cGAS activity in cells after purifying cGAMP from the cell lysates.

Several fluorescence-based methods were suggested for a high-throughput cGAS activity measurement. Based on the structure of natural riboswitches and a finding that GEMM-I riboswitch from *Geobacter* recognizes 3'3'-cGAMP and can be fluorescently modified, M.C. Hammond's group engineered a fluorescent biosensor capable of detecting 2'3'-cGAMP [595-598]. In this method binding of 2'3'-cGAMP to a fluorescently labeled riboswitch results in its conformational change and increase of fluorescence thus allowing an easy large-scale monitoring of cGAMP production by fluorescent readouts in a fluorescence plate reader [597]. Moreover, it can be applied for measuring cGAMP levels in the cell extracts without need for luciferase reporter assays, which would introduce high amounts of plasmid DNA into the cell and thus trigger a potent cGAS activation interfering with a studied response. Disadvantages of such method are cross-reactivity of a biosensor with 3'3'-cGAMP and c-di-GMP, sensibility of the biosensor to DNA-intercalating agents inhibiting cGAS activity and its instability against RNase contaminations.

An alternative fluorescent method of CDN detection was introduced by B.T. Roembke, *et al.* and is based on an elegant observation that in the presence of Mn^{2+} cations fluorescent analog of 3'3'-cGAMP - 3'3'-cG(d2AP)MP – forms heterodimers with other non-fluorescent cyclic dinucleotides. Such dimerization was found to lead to quenching of fluorescence [599]. Fluorescence of 3'3'-cG(d2AP)MP emerges from incorporation of the fluorescent nucleic acid base analog 2-aminopurine instead of adenine in bacterial
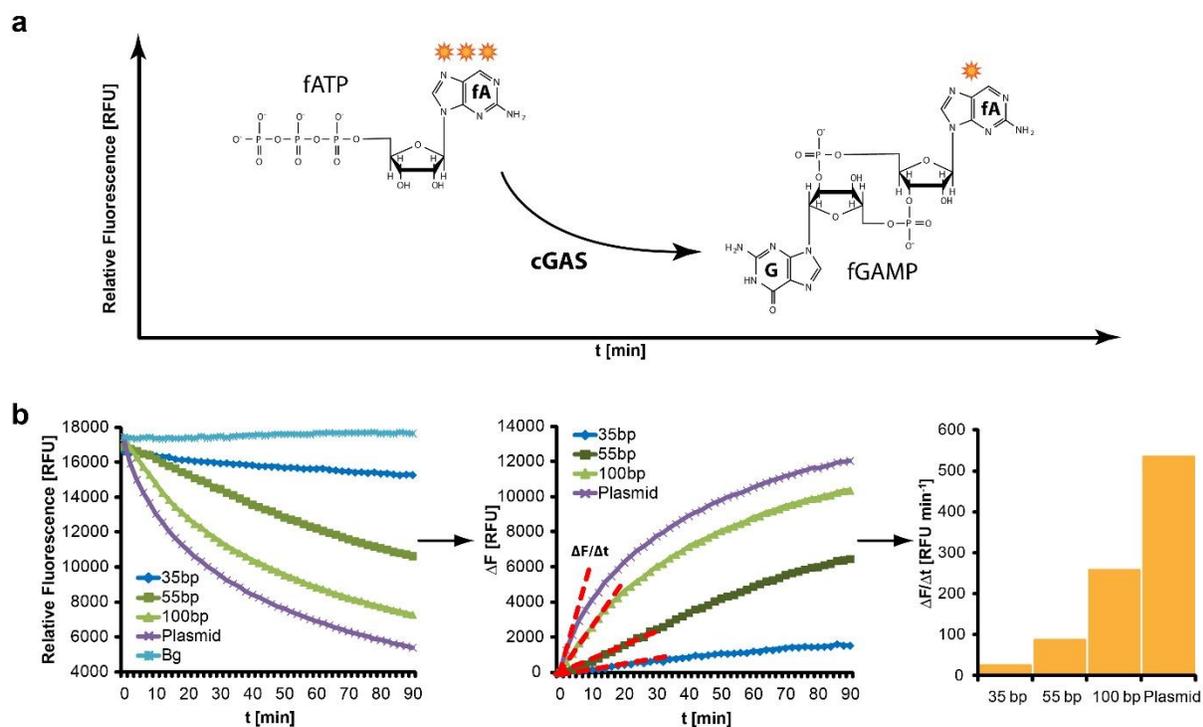
Figure 18 A novel high-throughput fluorescent method for measuring cGAS activity.

(a) The principle of the assay. cGAS incorporates highly fluorescent (three orange stars) ATP analogue 2-aminopurine-riboside-5'-triphosphate (fATP) into less fluorescent (one orange star) cGAMP analogue (fGAMP). Difference between fluorescent intensities corresponds to efficiency of cGAS reaction.

(b) General workflow for calculating initial cGAS reaction rates. Initial curves represent a decrease of fluorescence upon cGAS reaction (on the left). Background fluorescence at each time point is subtracted from fluorescence at corresponding point of cGAS reaction and resulting curves are inverted to obtain positive values (in the middle). The slopes of linear intervals (red dashed lines) are calculated and correspond to cGAS initial reaction rates plotted as bars (on the right).

Adopted from *Andreeva L., et al., 2017* [600].

3'3'-cGAMP. This method is suitable for high-throughput measurements of CDN content of the sample using a simple fluorescence plate reader and is suitable for both CDN-synthase and phosphodiesterase time-resolved activity measurements. 3'3'-cG(d2AP)MP quenching emerging purely from dimerization with CDNs does not occur in presence of NTPs or linear dinucleotides resulting in method specificity towards CDNs. The limitations of this technique arise from the need for Mn^{2+} cations that might not be suitable for some enzymatic reactions and, most importantly, its incapability to detect c-di-AMP and cGAMP with mixed phosphodiester linkages rendering it inapplicable for cGAS activity measurement.

A method presented in this thesis is inspired by the work of B.T. Roembke, *et al.* and, unlike described above, offers an easily accessible high-throughput fluorescent assay for a direct measurement of cGAS activity [600]. Briefly, we discovered that a fluorescent ATP analogue - 2-aminopurine-riboside-5'-triphosphate (fATP) – can be incorporated by cGAS into a fluorescent product (fGAMP), which was accompanied by a quenching of 2-aminopurine base fluorescence within the molecule (Figure 18a). This allows both single-point and time-resolved monitoring of cGAS reaction without any additional components and sensors (Figure 18b). This method thus enables simultaneous and quick comparison of

different reaction compositions, cGAS constructs and DNA-ligands in a large-scale format and is suitable for a broad range of conditions. Furthermore, fATP used in this method is commercially available and no specialized instruments are needed to conduct an experiment, since it can be done in a fluorescent plate reader with tunable excitation and emission wave lengths. Though this method can be applicable for a broad range of CDN-synthases, the enzyme tolerance to fluorescent ATP analogue must be considered. For example, though we tested the assay with all available cGAS constructs to be functional, human cGAS seems to have a stronger preference for native nucleotides and shows less efficiency in incorporating fATP into fGAMP, as compared to mouse and porcine enzymes. In comparison to radiolabeled cGAS activity assays our method requires lager volumes and therefore is best suited for measuring rather high activities, whereas low incorporation rates of fATP would be enzyme-consuming. However, even in this case the replacement of radioactive material with fluorescent analogues might still be favored. One more limitation of the method includes the time frame of the reaction. Since an open plate is preferable for an optimal fluorescence measurement, long-lasting time-resolved kinetics should take into account evaporation rate of the sample. Given that the method described in this work is based on a direct measurement of cGAS activity, it can hardly be applicable for cell-based assays due to the abundance of natural ATP in the living cell, so a combination of the assays might be the best choice for studying cGAS both *in vitro* and in the cells. Following the described principle, further search of environment-sensitive and small fluorescent base analogues that can be metabolized by cGAS and other CDN-synthases like DncV or bacterial c-di-AMP synthases is of great interest. Fluorescent cGAS activity assays based on 2-aminopurine analogue might be advantageous for future studies of cGAS activators and inhibitors and of particular interest for clinical implications.

2. A unique cGAS dimeric structure in context of long stimulatory DNA recognition

Since the first cGAS structures with short DNA fragments from 14 to 18 bp were published, the question why such short DNA fragments fail to activate cGAS *in vivo* and *in vitro* remained open. Indeed, the length of the major cGAS DNA-binding site corresponds to around 18-20 bp DNA stretch and, at least in the crystal structures, such binding is fully capable to introduce a "spine" helix break, N-terminal lobe stabilization, movement of the activation loop and rearrangement and closure of the catalytic site – all the features of an active cGAS [233-236]. The subsequent discovery of cGAS dimerization introduced additional questions that could not be answered with the existing structures. cGAS was shown to form a unique dimer with two cGAS protomers held together by their interactions with each other and each of two DNA molecules [313, 314]. Intriguingly, such dimer is present in all known crystal structures of cGAS that together with mutational analyses supports the idea of its physiological relevance. However, such dimerization does not explain a contradiction between structural and biochemical data, since $cGAS_2:DNA_2$ complexes represent an active cGAS conformation with non-stimulatory DNA ligands. cGAS was reported to require DNA ligands > 40 bp for a full activation *in vivo* and is active with shorter DNA fragments only at unphysiologically high concentrations [68, 112, 232, 314]. Further confusion arises from the relative position of the DNA fragments within such a dimer [313, 314]. Indeed, elongation of the dsDNA strands in a cGAS dimer would result in a steric clash between DNA molecules (Figure 19a). This means that such dimer formation would be reasonable, if cGAS preferentially recognized short DNA fragments in the cell to avoid dimerization on longer DNA species, which is contrast to the biochemical observations. Another plausible explanation of such cGAS dimer would include specific
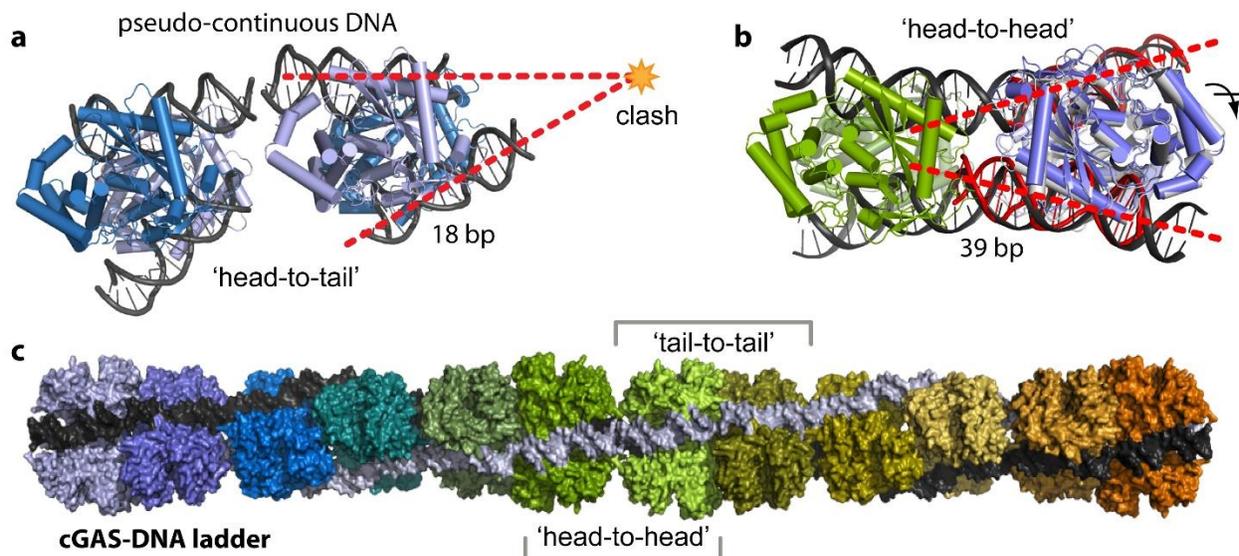
Figure 19 Structures and crystal packing of cGAS in complex with 18 bp and 39 bp DNA.

(a) Relative orientation of cGAS dimers within a complex of cGAS with 18 bp DNA (PDB: 4LEY). cGAS protomers within a dimer are represented in blue and light blue. Elongation of the DNA strands (red dashed lines) would lead to a steric clash of DNA molecules (orange star).

(b) cGAS in complex with 39 bp stimulatory DNA (green and grey, PDB: 5N6I) superposed to a cGAS dimer with 18 bp (blue, PDB: 4LEY). cGAS dimer with 18 bp is slightly rotated relatively to a dimer in cGAS complex with 39 bp (black arrow). DNA bending within 39 bp cGAS complex allows to avoid steric clash otherwise introduced by DNA positioning by a cGAS dimer (red dashed lines).

(c) cGAS-DNA fibril (ladder) emerging from crystal packing of cGAS in complex with 39 bp DNA (PDB: 5N6I). cGAS dimers form a ladder by iterative binding orientation on two nearly parallel pseudo-continuous DNA strands (grey and black). Each cGAS dimer is represented in one color.

Adopted from *Andreeva L., et al., 2017* [600].

recognition of DNA ends by cGAS to avoid the clash. This hypothesis, however, does not explain the difference between long stimulatory and short inactive DNA constructs, since in both cases cGAS would form the same dimer stabilized by the same interactions. Moreover, plasmid DNA that was found to potently activate cGAS would also not fit into this hypothesis [223]. The mystery of a $cGAS_2:DNA_2$ conformation includes the question about the stability of such a complex. Analytical ultracentrifugation, which was performed to show such a complex, required extremely high 80-150 µM cGAS concentrations and its dilution to still non-physiological 30 µM resulted in dissociation of the dimers [313, 314]. Intriguingly, in all crystal structures two DNA fragments from the neighboring cGAS dimers are stacked with each other in crystal packing mimicking a long ~40 bp DNA fragment (Figure 19a). Though the second DNA molecule bound to cGAS is pointing towards the other "pseudo-continuous" DNA strand and would induce the described collision in case of longer DNA segments, such binding of several cGAS along one stimulatory DNA might remotely refer to cGAS arrangement on DNA *in vivo*.

A mechanism combining dimerization and activation of cGAS by DNA > 40 bp is introduced in the current thesis. According to our understanding, cGAS might form fibrils on long DNA by iterating cGAS dimer orientation in contrast to crystal packing of cGAS in complex with 18 bp DNA and shorter (Figure 19c) [314, 600]. Our crystal structure shows how dimer formation occurs on long stimulatory DNA and how steric clash of DNA molecules is avoided without the need of DNA ends (Figure 19b). We also

propose a cooperative mechanism for DNA sensing by cGAS in which a single dimer formation is highly instable and gets mutually stabilized by other dimers within a cGAS-DNA ladder. This mechanism explains why dimerization itself on DNA around 20 bp, though present in crystal structures, is not sufficient for a full cGAS activation and why cGAS activity increases with DNA length at a constant concentration of cGAS DNA-binding sites [600, 601]. In line with previous studies, our cooperativity model shows that a 20 bp segment of DNA introduces all changes in cGAS necessary for its activation. cGAS, however, remains inactive within a single dimer simply because of its high dissociation rate. Such hypothesis of cooperative dimer stabilization provides the easiest explanation of cGAS dynamics. However, one can not exclude additional rearrangements in the active site of cGAS that could not be observed at low resolution. Nevertheless, since organization of the separate dimers in cGAS complex with stimulatory DNA is similar to the better resolved cGAS structures with shorter DNA fragments and since no protein-protein interactions between adjacent dimers that might influence cGAS active site were found, it seems unlikely that some unobserved due to low resolution dramatic changes in the catalytic pocket would occur and cause such a dramatic length-dependent increase in cGAS activity.

The flexibility of the DNA in our structure allows several cGAS dimers to be formed on two parallel DNA strands [600]. Iterating "head-to-head" and "tail-to-tail" (N-termini and C-termini of cGAS facing each other, respectively) orientation of cGAS dimers seems to be thermodynamically preferred, since in this case the DNA has to be bent twice less frequently, compared to a single "head-to-tail" orientation (Figure 20). However, the relative orientation of a cGAS dimer within a fibril in solution and in the cell remains to be investigated. Another question includes the "macroscopic" organization of the proposed oligomers. Though ITC and SEC-RALS experiments confirm the fibril formation up to $cGAS_6:DNA_2$ species in solution, net-like structures formed by several $cGAS_{2n}:DNA_2$ interconnected oligomers and not the compact fibrils with only two participating parallel DNA strands can not be excluded. In case of physiological cGAS ligands like HSV-1 genome or *L. monocytogenes* nucleoid (~ 152 kbp and ~2.9 Mbp) DNA formation of one compact fibril is rather unlikely due to a high flexibility and amount of DNA. It is therefore possible that several interconnected or unconnected cGAS-DNA ladders with different spacing between the cGAS dimers might be formed. cGAS oligomerization, independent of the "macroscopic" organization of the ladder, is further supported by cGAS localization studies, since several groups reported cGAS to form aggregates in the cell upon DNA transfection [222, 223, 602]. Particular organization of cGAS ladder on physiological DNA ligands might be an intriguing subject for future
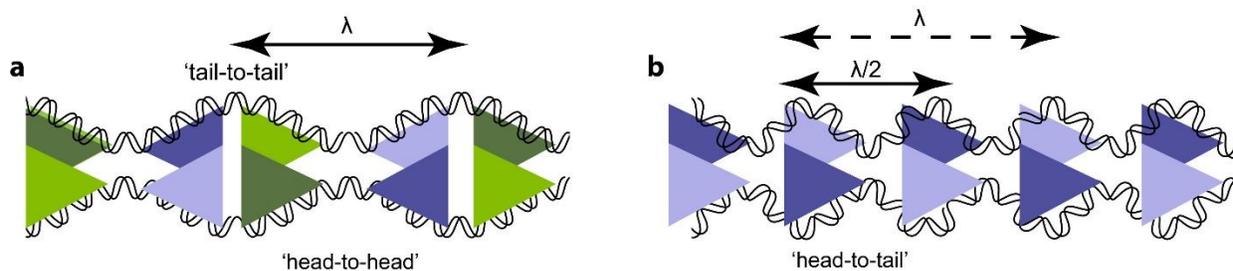
Figure 20 Schematic representation of the possible binding modes of cGAS dimers on long DNA strands. Iterating "head-to-head" and "tail-to-tail" orientation of cGAS dimers (green and blue) on long DNA strands (a) would introduce less frequent bending to DNA compared to a repetitive "head-to-tail" positioning (b). Iterative (a) and repetitive (b) cGAS dimer orientation would induce one DNA curve within λ or $\lambda/2$ bp, respectively, as indicated with black arrows.
Adopted from *Andreeva L., et al., 2017* [600].

studies of cGAS oligomers.

Cooperative mechanism of DNA sensing by cGAS also provides an explanation for low affinity of $cGAS_2:DNA_2$ complex formation. According to our data only oligomeric species bigger than $cGAS_4:DNA_2$ could be observed in solution in nearly physiological conditions and required DNA > 40 bp, whereas cGAS complex with 20 bp DNA dissociated into $cGAS_1:DNA_1$ [600]. In agreement to previous studies, this supports the instability of cGAS dimer with 20 bp or shorter DNA [314, 600]. The nature of a $cGAS_1:DNA_1$ complex, however, is not clear. Though the "major" DNA-binding site (site "A") of cGAS is usually understood to be a "spine" helix and the "platform", the second DNA-binding site seems to be even more important. Mutational studies revealed that a single amino acid substitution in DNA-binding site "B", in contrast to site "A", can fully abolish DNA binding by cGAS [314]. This observation leads to the question, whether cGAS can form $cGAS_1:DNA_1$ complex at all or whether DNA-binding occurs only within a cGAS dimer. However, since our data indicate the presence of $cGAS_1:DNA_1$ in solution, the most probable explanation would be that the binding affinity of site "B" towards DNA is much higher in comparison to a "spine" helix and therefore its disruption results in a weak interaction with DNA. This could mean that cGAS interacts with the long DNA with site "B" first. Afterwards, when the condition is suitable for dimerization, cGAS would bind with the "spine" helix of site "A", which would induce an active site rearrangement and cGAS activation. Such different affinity of the two DNA-binding sites would explain a two-step mechanism of DNA binding by cGAS that, if true, would counteract an aberrant cGAS activation in the context of cGAS dimer. This hypothesis, however, is highly speculative, since a deletion of a "spine" helix analogously to a point mutation in site "B" was reported to completely abolish DNA binding by cGAS in co-immunoprecipitation experiments [223]. Further research is required to evaluate the affinities of both cGAS DNA-binding sites separately and to examine a time-resolved mechanism of cGAS interaction with DNA.

As mentioned above, the cGAS dimer with short DNA is unstable and the first $cGAS_2:DNA_2$ species were observed at high cGAS and DNA concentrations that apparently shifted the equilibrium towards their formation. Remarkably, this correlates with cGAS activity, since 16-20 bp DNA was able to activate cGAS at non-physiologically high enzyme concentrations (10 µM), whereas decreasing cGAS concentrations twice or more resulted in a dramatic decrease of its stimulating capacity [234, 314, 600]. A requirement of cGAS dimerization and a unique structure of cGAS dimer accompanied by its high dissociation rate might have evolved during metazoan evolution in order to distinguish between pathogen- and damage-associated long DNA molecules and short dsDNA products resulting from nuclease cleavage or genome maintenance. Such length-sensitivity provides a sequence-independent mechanism of discriminating between danger-associated and harmless dsDNA products by cGAS needed to avoid the aberrant activation of the cGAS/STING pathway and autoinflammation.

3. Function of N-terminal cGAS domain

The majority of studies concentrated on the structure and function of the conserved Mab21 domain of cGAS comprising both DNA-binding and enzymatic activities, but no robust and mechanistic understanding of functions of the N-terminal part of cGAS was achieved so far [223, 233, 235]. Unlike Mab-21 domain, N-terminus of cGAS (aa 1-160 of human cGAS) is non-conserved and varies in length among cGAS homologs [315, 343]. The secondary structure suggests this cGAS part to be highly flexible [315]. However, some similarities in sequence can be observed among different mammalian cGAS N-

termini that share similar length around 160 aa and are enriched in positively charged residues. In the first studies N-teminus of cGAS was hypothesized to stabilize the protein *in vitro* and participate in autoinhibition of cGAS analogously to other innate immune sensors like RIG-I and AIM2, though no robust confirmation of this functions was obtained so far [235].

In agreement with abundance of positively charged residues in N-terminal part of cGAS it was reported to bind DNA independently of the catalytic domain. Co-immunoprecipitation studies revealed its association with 45 bp DNA with nearly the same efficiency, as Mab21 domain [223]. Further studies reported that the N-terminus increases cGAS affinity towards DNA. According to ITC measurements mouse cGAS catalytic domain binds 20 bp with ~ 20 µM affinity, whereas human full length protein exhibits ~300nM and ~ 90 nM affinity to 20 bp and immunostimulatory 45 bp DNA, respectively [235, 314]. Electromobility shift assays confirmed higher affinity of full length cGAS in comparison to Mab21 domain towards 45 bp DNA for mouse cGAS [314]. A more systematic study revealed an increase in affinity for mouse and human cGAS towards all tested DNA species in a row: catalytic domain, N-terminal part and full length construct, the last having the lowest K_d as determined in electromobility shift assays [315]. Dissociation constants of catalytic domain, N-terminus and full length cGAS to 16 bp DNA were estimated to be ~ 40 µM, 1.5 µM, 0.5 µM and 112 µM, 5 µM and 1 µM for human and mouse cGAS constructs, respectively, as calculated with microscale thermophoresis. The same study reported that the N-terminal domain of cGAS undergoes dramatic changes in its secondary structure upon DNA binding similarly to its catalytic domain [315].

Though the N-terminal domain was quite early shown to participate in DNA binding and cGAS stability, the first studies did not observe any difference in enzymatic activity between full length and truncated cGAS constructs *in vitro* and in the cells [223, 235]. This, however, may be caused by high protein concentration in an *in vitro* system and an overexpression of cGAS constructs in the cells that might have driven cGAS to the maximal activity and thus decreased the sensitivity of the methods. The latter research using milder conditions showed that full length cGAS, having the best affinity to DNA, exhibits the highest enzymatic activity with 45 bp DNA *in vitro* and in human cells among cGAS constructs [315]. Moreover, a cleavage of cGAS by caspase-1 and other inflammatory caspases removing its N-terminal domain was shown to dramatically decrease cGAS-driven cGAMP production in the cells [469]. Intriguingly, this perfectly goes in line with the hypothesis proposed in this thesis that cGAS affinity towards DNA is a primary force increasing cGAS activity. Indeed, the increase of DNA length resulting in higher cGAS activity also directly correlates with affinity and stability of cGAS-DNA complexes in all DNA ranges including weak cGAS stimulators (12-20 bp) and immunoactive DNAs (20-1000 bp) both *in vitro* and in the cells (Figure 21) [68, 112, 232, 235, 314, 600].
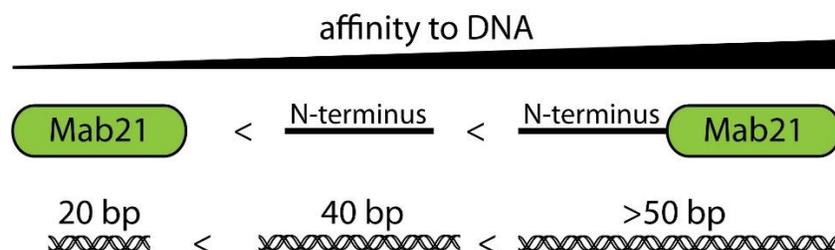
Figure 21 Correlation between affinity of cGAS to DNA (black triangle) and cGAS (on the top) or DNA (on the bottom) constructs.

Since this work proposes the cGAS-Mab21 oligomerization to be a source of increased stability of the cGAS-DNA complex, the question arises, whether the same fibrils are formed in case of the full length cGAS. On the one hand, DNA-binding of the N-terminus might increase the minimal DNA length needed to occupy one cGAS monomer. Though the structural data about N-terminus are missing, it was found to bind 16 bp DNA with high affinity suggesting that together with the catalytic domain a DNA-binding site length would approach ~ 36 bp or more [315, 603]. This DNA length is quite close to transition in cGAS activity around 40-50 bp long DNAs, so one might assume a single full length cGAS dimer formation to be sufficient for a full cGAS activation [68, 112, 232, 600]. On the other hand, however, the DNA length-dependency does not end at the initial point of 40-50 bp and further increase of DNA length results in continuous increase of cGAS activity. Though a composition of full length cGAS complex with 40-50 bp needs further research, these results suggest that also for a full length cGAS the oligomerization mechanism most likely remains relevant for stabilization of the dimers and cGAS activity.

Another intriguing question includes the influence of N-terminal cGAS domain on dimerization and oligomerization and several possibilities might be expected. The N-terminus itself was reported to bind DNA as a 1:1 complex suggesting that it does not induce dimerization by itself, unlike the catalytic domain [315]. Moreover, the N-terminus of cGAS was also proposed rather to counteract dimerization of cGAS as tested by analytic ultracentrifugation [603]. However, the interpretation of the molecular weights in this study, as well as a reverse correlation between DNA length and cGAS activity is debatable. Moreover, the assumption that the N-terminal cGAS part deoligomerizes cGAS can be explained by a competition of cGAS domains for DNA binding resulting in disruption of cGAS dimers held together mostly by cGAS-DNA interactions and any other DNA binding protein would result in the same effect. Together with previously published mutational analysis of cGAS in cell lines and a stimulating effect of DNA-bending proteins on full length cGAS, discovered in this work, formation of dimers by the full length cGAS seems to be the best explanation of the existing data [314, 600]. However, the mechanism of such dimer formation remains enigmatic in regard of full length cGAS. One possibility
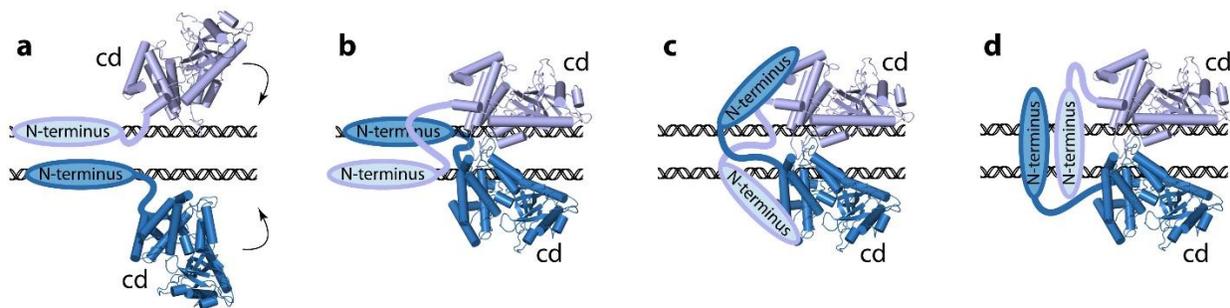
Figure 22 Schematic representation of possible mechanisms of dimer stabilization by cGAS N-terminus.
  (a) The N-terminus might bind DNA and recruit the catalytic domain (cd) of cGAS to the DNA strands.
  (b) The N-termini might stabilize cGAS dimerization by crossed binding of the DNA strands opposite to the corresponding catalytic domain.
  (c) The N-termini might "hug" the cGAS dimer by protein-protein interactions and additionally interact with DNA.
  (d) The N-termini might bind both DNA strands simultaneously stabilizing a parallel arrangement of DNA strands.

would be that the N-terminal cGAS domain simply increases cGAS affinity to DNA and thus brings weak DNA-binding catalytic domains to its proximity facilitating DNA-binding and dimerization of Mab21 domains (Figure 22a). Another possibility would be that the N-terminus of one cGAS protomer binds the neighboring parallel DNA strand resulting in additional dimer stabilization (Figure 22b). Since a co-expression of the N-terminal cGAS fragment together with a catalytic core restored cGAS-Mab21 activity to that of a full length protein, following two functions of the cGAS N-terminus could also be imagined [603]. The N-terminus could interact directly with a neighboring cGAS protomer thus "hugging" a cGAS dimer and stabilizing cGAS active conformation (Figure 22c). To investigate this possibility protein-protein interactions between cGAS catalytic core and N-terminus must be explored. Moreover, a single N-terminal domain could act as a hook binding two neighboring DNAs with its ends thus holding them together and pre-arranging DNA in a parallel manner (Figure 22d). All these possibilities, however, are highly speculative and further research is required to evaluate each of them.

Intrinsic capability of full length cGAS to measure DNA length as observed for Mab21 domain and the formation of the aggregates as observed in electromobility shift assays with the cGAS full length construct rise the question, whether N-terminal cGAS domain specifically participates in further oligomerization of cGAS and fibrils formation [314]. Despite DNA-binding, the N-terminal domain of cGAS might participate in protein-protein interactions between adjacent cGAS dimers within a fibril or in a net formation by protein-protein or DNA-mediated interactions with other fibrils. Aggregation and oligomerization of the full length cGAS on different DNA constructs need further research, since electromobility shift assays, though showing increased aggregation for cGAS full length construct compared to Mab21 domain in the same condition, also show aggregate formation of a catalytic domain with DNA in some studies [233, 314]. Since a complex of the N-terminus with only a short 14 bp DNA was tested, its intrinsic capability of filament formation along DNA strand would be also of a particular interest [315]. Intriguingly, the N-terminus was found to possess a high sliding mobility along a DNA strand, which was reduced by the addition of a catalytic domain in full length cGAS construct [315]. Together with the oligomerization mechanism proposed in this thesis it is tempting to speculate that cGAS might first bind DNA with its N-terminus and slide towards the previously formed cGAS-DNA ladder or fibril nucleation point where the DNA strands are already prearranged, so dimerization of the catalytic domain and cGAS activation may occur (Figure 23). Further structural and biochemical studies, however, are necessary to address all the points concerning oligomerization mechanism of the full length cGAS, as well as structure and function of its N-terminus in this process.

Figure 23 A model for a ladder formation by full length cGAS.
cGAS might first interact with the DNA with its N-terminus (hexagons) and slide towards the end of the formed ladder. The existing ladder parallelizes DNA strands in its proximity and therefore facilitates DNA binding and dimerization of the Mab21 catalytic domains (circles). Protomers within cGAS dimers are colored with different shades of blue.
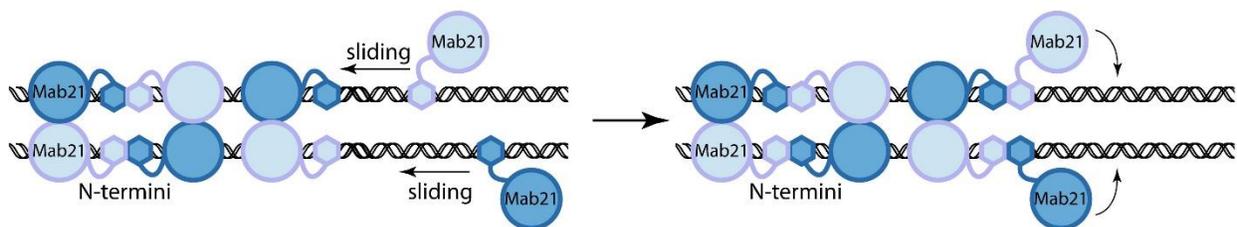
4. cGAS co-factors and physiological ligands: genome products, mitochondrial DNA and bacterial nucleoids

Following the first discovery of cGAS as a sensor of foreign DNA emerging from DNA-viruses and intracellular bacteria the complexity of its function evolved rapidly. The first break-through was made by the finding of mitochondrial DNA (mtDNA) leakage and its recognition by cGAS upon mitochondrial stress induction or viral infection [225]. Though the exact mechanism of mtDNA escape into the cytosol is unknown, TFAM depletion was found to result in accumulation of aberrant mtDNA and mitochondrial hyperfusive phenotype. Induction of the mitochondrial fission abrogated mtDNA leakage and IFN response. Mitochondrial stress was also shown to be induced by HSV-1 and (+)ssRNA Dengue virus explaining the involvement of cGAS/STING axis in recognition of RNA viruses [225, 604]. These findings indicate that even in the presence of viral DNA in the cytosol pathogen-induced self-DNA translocation may serve as the most potent cGAS activator. A crucial role of mtDNA in cGAS activation was further supported by the discovery of a cGAS-inhibiting role of apoptotic caspases [226, 227]. R.A. Flavell's and D.C.S. Huang's groups simultaneously published that inhibition of apoptotic caspases together with activation of apoptosis results in mtDNA escape from mitochondria and cGAS activation. According to the authors, dying cells release mtDNA into the cytosol in a Bax/Bak-dependent manner and trigger type I IFN response in the absence of apoptotic progression. However, the exact mechanism of mtDNA leakage into the cytosol remains unclear and needs further research. Mitochondrial hyperfusion may lead to a simple rupture of mitochondrial membrane followed by exposure of mitochondrial material into the cytosol. On the other side, involvement of Bax/Bak in mtDNA transition by apoptosis inhibition may indicate that mtDNA leaks through protein channels like Bax/Bak pores [226, 227]. The latter is, however, enigmatic given a small size of Bax/Bak channel – 5-6nm – that might not be sufficient for a transport of 16 nm in diameter supercoiled plasmid, but would suit for a single 2nm in diameter dsDNA fragment resulting from a preceding mtDNA cleavage [605-607]. Whether the nucleoid DNA interacts with the cytosolic compartment fully or partially by exposed loops, also remains to be uncovered. Based on the presence of mtDNA in cytosolic fractions free of mitochondrial contaminations, however, the leakage of the whole mitochondrial nucleoid into cytosol seems to be most likely [225].

Another source of self-DNA emerges from nucleus. Cytosolic chromatin fragments were shown to activate cGAS in cancer and senescent cells triggering IFN response [230, 231, 608]. Moreover, a mechanism explaining how nuclear instability and double-stranded DNA breaks (DSBs) induced by radio- and chemotherapies in cancer elicits IFN response was recently discovered [228, 229]. A.P. Jackson's and R.A. Greenberg's groups discovered that not DSBs themselves, but a mitotic progression following DNA damage results in formation of aberrant enveloped chromatin fragments – micronuclei – that stimulate cGAS activation and IFN signaling upon membrane rupture. Similar to mitochondrial stress, genome instability may be caused by such viruses as KSHV or Epstein-Barr virus, further supporting the role of self-DNA in cGAS activation [609-611]. Accumulation of aberrant mitochondrial nucleoids and chromosomes seems to fit into a general mechanism of cGAS activation by stress- or virus-induced instability of genomic material and attracts a huge interest to further studies of cGAS activation by these DNA species. Moreover, the mechanism including membrane permeabilization followed by chromatin or mitochondrial nucleoid exposure is tempting to apply for cGAS sensing of intracellular bacteria, though precise mechanism of DNA leakage into cytosol in this case remains to be elucidated.

Based on these spectacular findings the question arises, if such a self-DNA-mediated cGAS activation evolved separately to cGAS recognition of dsDNA or if cGAS is specifically tailored for sensitive self-DNA recognition. Or, formulated in another way, does self-DNA trigger a potent cGAS response only as an additional source of DNA, due to its length shown to be important for cGAS activation, or through special structures introduced by associated proteins? The work presented in this thesis offers a hypothesis explaining a particular role of self-DNA in cGAS activation. Since chromatin, mitochondrial and bacterial DNA are all structured by a range of DNA-binding proteins the function of these proteins as cGAS co-factors was tempting to suggest. Among others, DNA-bending mitochondrial TFAM, nuclear HMGB1 and bacterial nucleoid-packaging proteins HU attract special attention. HMGB1 was previously found to be involved in innate immune response as a cytokine activating TLR2, TLR4, TLR9 and receptor for advanced glycation end products (RAGE), as well as in IFN response to cytosolic DNA and poly(I:C) [295, 298, 612, 613]. Indeed, all mentioned proteins were found to strongly stimulate cGAS activity *in vitro* and mutations disabling DNA-bending properties were found to abolish this effect at least for HMGB1 [600]. According to the previously published structures U-turns introduced by these proteins can be perfectly aligned to a cGAS fibrils observed with 39 bp stimulatory DNA [600, 614]. Thus cGAS oligomerization on long DNA combined with DNA-bending cGAS co-factors stimulating its activity generate a novel view on nucleation-oligomerization mechanism of cGAS sensing [600]. In this hypothesis DNA-bending proteins pre-arrange long DNA in a way favorable for cGAS dimer binding and may serve as high-affinity starting points for cGAS fibril formation (Figure 24). In a physiological context this would mean that pre-structured TFAM/HMGB1/HU-bound long DNA species rather than naked DNA are preferable cGAS ligands. Such DNA structures may represent leaked into cytosol partly unfolded mitochondrial and bacterial nucleoids (Figure 24). However, further studies are needed to explore the physiological relevance of the proposed mechanism and the effect of these co-factors on cGAS activity in the cells and *in vivo*. For example, the activating effect of HU needs further clarification, since HU seems to introduce U-turns only in the context of specific distorted DNA [615]. Composition of HUα and HUβ *E. coli* proteins was found to influence their binding mode of bacterial plasmids switching between a single filament state and net-like DNA structures formed by bridging several DNA molecules into a network [616]. If not DNA-bending, formation of parallelized DNA strands by interacting HU proteins might explain the cGAS stimulation effect of these proteins. Structural studies on cGAS binding mode in context of HMGB-bound DNA are also of particular interest. Furthermore, a necessity of cGAS fibrils formation in the case of bent DNA must be further explored. Though nucleation-oligomerization mechanism of cGAS activation seems to be the most plausible explanation of the biochemical data obtained so far and was previously reported for other innate immune sensors like NAIP2/NLRC4 inflammasome assembly [617], recognition of structured DNA and cGAS fibril formation might also occur separately in a biological context. Like for MDA5 dsRNA receptor [618], cGAS oligomerization might be a major mechanism for discriminating long pathogenic DNA molecules from short non-stimulating digestion products independently from the DNA U-turns and DNA-bending proteins. The recognition of structured DNA by cGAS, on the other side, might function independently of cGAS oligomerization to specifically detect HMGB-bound self-DNA or bacterial nucleoids and require a single cGAS dimer stabilized by its proximity to a DNA-bending protein. However, the fact that TFAM stimulating activity on cGAS increases with DNA-length up to 200 bp rather supports the hypothesis of nucleation followed by cGAS oligomerization [600]. Intriguingly, a rapid increase in the TFAM-stimulating capacity around 200 bp coincides with a previously reported DNA-length, which is needed for an optimal TFAM dimerization
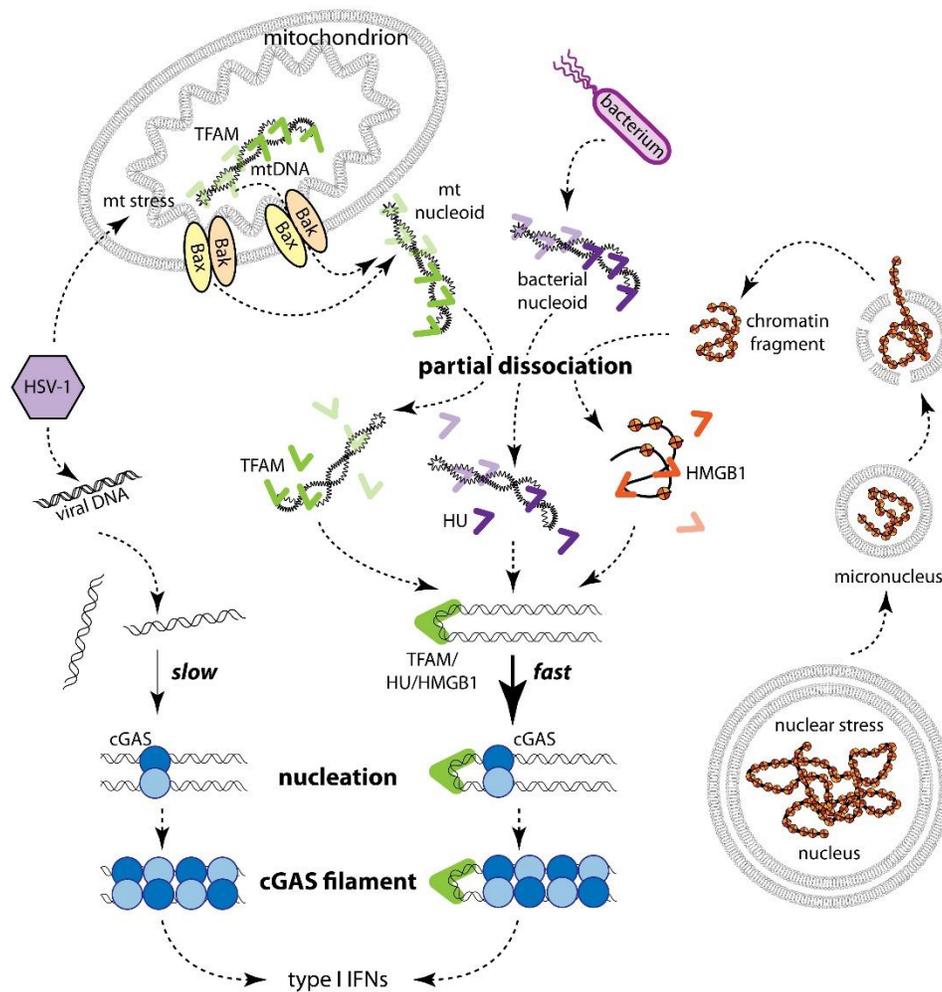
Figure 24 Nucleation-oligomerization mechanism of cGAS sensing.
Leakage of mitochondrial (mt) and bacterial nucleoids, as well as chromatin fragments into cytosol results in their decondensation and partial dissociation of the associated proteins. Remaining DNA-bound TFAM, HU and HMGB1 (green, violet and orange corners, respectively) prearrange the DNA in a nearly parallel way favorable for cGAS dimer binding (solid bold arrow) and thus efficiently nucleate cGAS filament formation. "Naked" DNA species like viral DNA are not structurally prearranged and thus potentiate cGAS filament formation slower (thin solid arrow). Consequent events are connected with black dashed arrows.

[619]. This raises the question whether DNA U-turns or the loops introduced by TFAM dimerization and nucleoid packaging are the driving force for cGAS stimulation by TFAM. Given that TFAM dimer-induced DNA compactization was reported to gradually increase with DNA length, the latter would result in further increase of TFAM capacity to stimulate cGAS on the DNA fragments above 200 bp, which is not the case [600, 619]. Moreover, since no further growth of TFAM-stimulated cGAS activity was observed above 200 bp threshold unlike TFAM-free system, cGAS oligomer growth from TFAM-nuclei might be either limited or unnecessary for the observed TFAM stimulation. Therefore such discrimination between DNA-structuring and cGAS oligomerization would be one of the important directions of the future studies.

Given that only a particular concentration of DNA-bending proteins resulted in stimulation of cGAS activity and above this concentration competed with cGAS for DNA binding and counteracted cGAS reaction, DNA compactization seems to play an important role in regulating cGAS activity. It would explain the immunosilence of mitosis despite the fact that cGAS co-localizes with chromatin after nuclear membrane disassembly and immunostimulatory activity of the micronuclei, on the other side [228, 229, 608]. The need for a partly dissociation of chromatin or mtDNA structuring proteins for cGAS activation and its inactivity on fully condensed self-DNA would thus allow a "healthy"-self vs. "damaged"-self discrimination, as well as delayed cGAS response in the case of self-DNA leakage. The latter might be needed for activation of other interconnected pathways in order to fine-tune a final cellular response. The same principle could also be applied for bacterial nucleoids covered with HU DNA-binding proteins (Figure 24).

The discovery of TFAM and HMGB1 in the context of self-DNA recognition by cGAS offers a broad field for searching for other potential cellular co-factors of cGAS. As such, nuclear proteins MRE11 and IFI16 previously shown to be involved in STING-dependent signaling represent promising potential candidates for assisting cGAS activation [274, 307]. Nuclear transcription factors that might reside on chromatin material within cGAS-stimulating micronuclei are another batch of potential cGAS DNA-arranging co-factors. Moreover, chromatin and nucleosome structure itself may arrange DNA in a way favorable for cGAS dimerization and thus needs further characterization as cGAS activator. A recent study revealing the interferogenic potential of neutrophil extracellular traps (NETs) that consist of chromatin material bound to cytosolic and granular content attracts attention to these structures as cGAS activators [620]. Reactive oxygen species production by mitochondria that drive the exposure of NETs and mtDNA oxidation provide an additional link between chromatin and mitochondrial DNA recognition and is of particular interest as potential structured cGAS ligand due to its involvement in STING-dependent IFN signaling and systemic lupus erythematosus [620].

Whether cGAS is capable to specifically attract the described and putative co-factors, is another intriguing question. Though no protein-protein interactions between human full length cGAS and TFAM could be detected, it can be different for other co-factors. Among others, the function of the N-terminal cGAS domain in cGAS oligomerization and its interaction with cellular co-factors is of particular interest. The involvement of described cGAS activators in recognition of initially naked transfected or viral DNA is another open question. Though TFAM normally localizes to mitochondria and leaks to the cytosol only under mitochondrial stress conditions, it might partly dissociate and bind viral DNA further increasing cGAS response to viral DNA [600]. Moreover, nuclear protein HMGB1 was found to co-localize with cGAS and transfected plasmid. The presence of cGAS does not seem to be necessary for HMGB1 translocation from the nucleus. Indeed, DNA transfection itself was sufficient to induce HMGB1 relocalization and resulted in its association with cytosolic DNA loci [600]. Based on the data presented in this thesis, no direct interaction of cGAS with its DNA-binding co-factors is necessary to assemble active co-factor-bound cGAS complexes. Instead, a dramatic increase of cytosolic DNA concentration during infection might lead to a shift in the equilibrium and attract HMGB1 and other co-factors into the cytosol. In this case a constant shuttling of the factors between their compartments and cytosol that was already proposed for HMGB1 would favor such association and accumulate low amounts of DNA-binding cGAS co-factors on cytosolic DNA species thus increasing their potency as cGAS activators [621]. However, further studies on cGAS co-factors and their localization are needed to evaluate this hypothesis.

5. HIV genome and structured DNA as specific cGAS ligands

The hypothesis that cGAS specifically recognizes structured self-DNA like mitochondrial nucleoids and chromatin fragments raises the question whether some host factors act in the same way in context of viral DNA. Since viral genomes often have secondary structured elements and are often covered with viral proteins, such structured DNA species also attract particular interest as potential strong cGAS activators, similar to that discovered for other innate immune sensors. For example, RIG-I specifically recognizes secondary structured elements within defective DenV viral genomes, which are produced when the viral polymerase slips from the template and copies back the short RNA strand [622]. Moreover, circular RNA species evolving during back-splicing from host mRNAs are also potent RIG-I activators and are proposed to be transmitted between the cells to amplify innate immune response [623, 624].

Though cGAS is mostly known for sensing dsDNA viruses, recent discoveries introduced retroviruses as activators of the cGAS/STING axis [220, 509, 512]. In the series of knock-out and mutational studies the group of James Chen showed that cGAS is responsible for IFN induction in various human cell lines in response to murine leukemia virus (MLV), simian immunodeficiency virus (SIV) and HIV-1 [220]. Intriguingly, TREX1 knock-out decreased but not abolished this response. cGAS activation was shown to depend strongly on reverse transcription since its chemical inhibition or expression of SAMHD1 in human macrophages and dendritic cells abolished IFN production, which means that retroviral DNA species are responsible for immune response. An observation that HIV-1 elicits immune response independently of viral genome integration together with the finding that SAMHD1 activity leads to accumulation of HIV-1 reverse transcription intermediates activating STING and IRF3 attracted attention to these ssDNA species as potential cGAS activators [512, 625]. Indeed, ssDNA generated from HIV-1 genome by reverse transcription was capable of inducing IFN response due to its secondary structure comprising dsDNA stems [237]. However, such dsDNA regions within HIV-1 ssDNA being ~20 bp long are below the minimal length needed for cGAS activation. Further studies, presented in this thesis, revealed that the presence of G-bases flanking a dsDNA hairpin are needed for HIV-1 ssDNA immunostimulatory capacity and designed synthetic short dsDNA sequences are capable of cGAS activation (Figure 25) [626]. In contrast to blunt dsDNA fragments that have to exceed ~ 45 bp to be efficient cGAS ligands, the dsDNA constructs with single or double G_2-G_3-overhangs on both sides – Y-shaped DNA (YSD) - reduce this threshold up to 14 bp. Such YSDs represent a novel specific DNA structure of viral origin recognized by cGAS [626]. The mechanism of sensitive cGAS recognition of such DNA species, however, remains to be clarified. Since several hairpin regions are present within HIV-1 ssDNA, they might occur close to each other forming nearly parallel dsDNA strands that would stimulate cGAS dimer formation analogically to U-turn DNA species, which were described above (Figure 25a). Guanines at the end of dsDNA regions in this case might influence an overall fold of HIV-1 ssDNA contributing to positioning of two dsDNA regions in close proximity favorable for cGAS activation. More enigmatic is a mechanism of cGAS activation by synthetic YSD constructs in which dsDNA regions are not connected within one molecule. Given that YSDs are potent cGAS activators not only in the cells, but also in *in vitro* system, their intrinsic capability of cGAS activation needs further research. Since cGAS active conformation corresponds to a dimer that is unstable in complex with 20 bp DNA, G-overhangs might contribute to cGAS dimer stabilization via guanine-protein interactions (Figure 25c). Alternatively, given that guanines are prone of forming G-quadruplex structures by chelating a K^+ cation, short YSDs might stick together through guanine interactions thus mimicking a long stimulatory DNA (Figure 25d). Though this hypothesis would go in line with cGAS oligomerization mechanism on
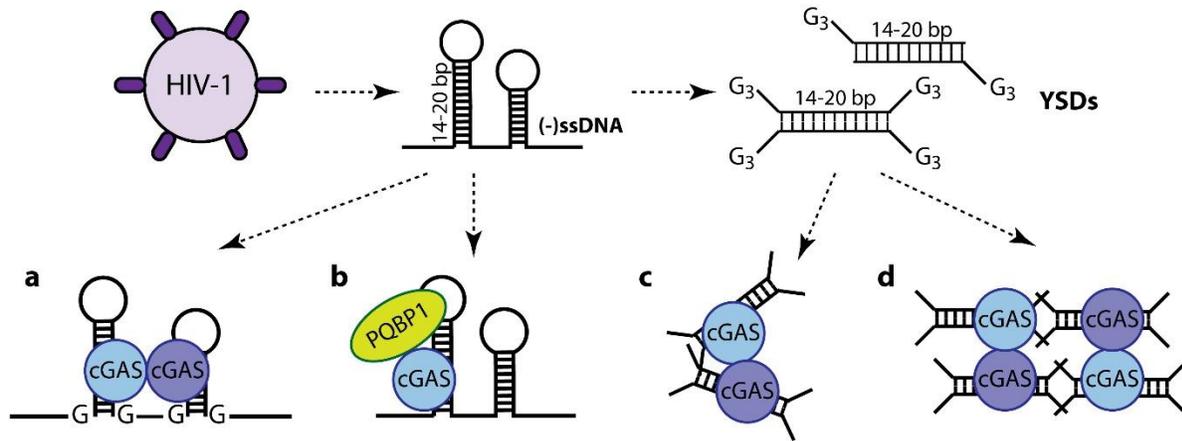
Figure 25 Nucleic acid species activating cGAS during HIV-1 infection and possible modes of their recognition.
Reverse transcription of HIV-1 genome results in negative ssDNA with guanines flanking double-stranded stem regions recognized due to presumable specific parallel arrangement of dsDNA regions (a) or with help of PQBP1 (b). G-overhangs of Y-shaped DNA (YSD) – an analogue of HIV-1 (-)ssDNA stem regions – might stabilize the cGAS dimer (c) or mimic long DNA by stacking of several short dsDNA fragments (d).

long stimulatory DNA, it does not explain why a replacement of guanines with 7-deazaguanines or inosines that are incapable of G-quandruplex formation retain their high potential to activate cGAS in the cell system [600, 626]. Therefore further research is needed to clarify the mechanism of cGAS activation by YSDs and HIV-1 in general.

Though the ssDNA structure itself is a prerequisite for cGAS activation by HIV-1, other cellular factors might also contribute to cGAS-mediated immune response to the virus *in vivo*. One might expect cellular co-factors to specifically bind HIV-1 ssDNA or G-overhangs and thus stimulate HIV-1 and YSD recognition by cGAS. Remarkably, parallel to the discovery of cGAS as HIV-1 sensor, IFI16 was also shown to be involved in HIV-1 sensing [237]. Though the later studies found a role of IFI16 in mediating TBK1 recruitment by STING in the same cGAS/STING axis, its direct influence on cGAS remains unclear, thus offering a possibility of its involvement in recognition of structured HIV-1 ssDNA [306, 307]. Though the role of IFI16 in HIV-1 DNA binding is highly speculative, another cellular co-factor was shown to specifically bind HIV-1 ssDNA and facilitate cGAS activation upon HIV-1 infection. Polyglutamine binding protein 1 (PQBP1) was found to be an essential co-factor for immune response to HIV-1 infection, as discovered during RNAi screening of monocyte-derived dendritic cells (MDDCs) infected with HIV-1 (Figure 25b) [219]. PQBP1 was shown to interact directly with both HIV-1 ssDNA and cGAS, as shown in co-immunoprecipitation studies, and to be dispensable for viral and synthetic dsDNA sensing. The exact mechanism of PQBP1 cGAS stimulation upon HIV-1 infection, as well as importance of HIV-1 ssDNA secondary structure and guanines flanking HIV-1 hairpins for PQBP1 binding represent one of the directions for future studies. Whether PQBP1 is able to recognize synthetic YSDs and trigger cGAS activation by these DNA species is another important question. Moreover, the discovery of PQBP1 further attracts attention to other putative co-factors stimulating cGAS activation in context of various DNA species.

The finding of cGAS activation by HIV-1 indicates a general mechanism of cellular response to retroviruses and raises the question whether retrotransposons might be another source of cellular DNA molecules triggering IFN response through cGAS. The majority of the retroelements in human genome are represented by long interspersed nuclear elements (LINEs, ~ 6kb in length) and short interspersed nuclear elements (SINEs, ~ 3kb in length) that are normally epigenetically silenced but upon transcription into RNA are reverse transcribed into cDNA and dsDNA and inserted into human genome [627, 628]. Though intermediates of retroelements are normally present in the nucleus and are not accessible to cytosolic sensors, their increased activation leads to accumulation of cytosolic nucleic acid species like dsRNA and dsDNA that were proposed to be recognized by RLRs and cGAS, respectively (Figure 26). Increased expression of retroelements was observed in human prostate cancer and metastatic colon tumor cells [629, 630]. Furthermore, tumor tissues are characterized by an increased level of demethylation of retroelement loci presumably leading to a loss of their epigenetic silencing in various cancers [631]. RNA intermediates of SINE replication were shown to activate TLR7 and contribute to SLE, whereas DNA methyltransferase inhibitors used as chemotherapeutics were shown to upregulate retroelements and induce MDA5 and TLR3 via their dsRNA intermediates [632-634]. Moreover, some viruses like MHV68
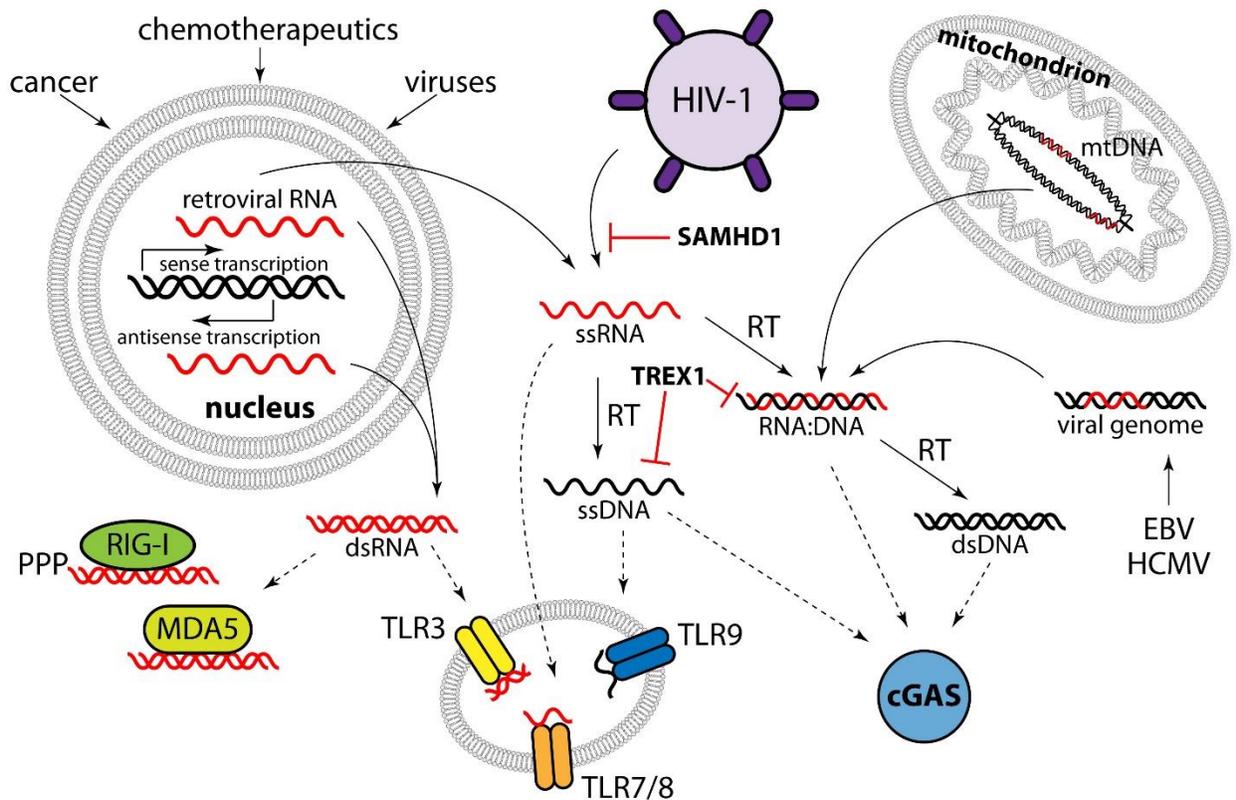
Figure 26 Innate immune sensing of retroelements and sources of RNA:DNA hybrids.
HIV-1 infection, as well as activation of retroelements by chemotherapeutics, viruses or cancer leads to enhanced transcription of retroviral ssRNA (red wave), which is reverse transcribed (RT) into ssDNA (black wave), RNA:DNA hybrids (black and red double wave) and dsDNA (black double wave) recognized by cGAS. Other sources of RNA:DNA hybrids include mitochondrial replication products and herpesviral genomes. dsRNA (red double wave) generated by annealing of sense- and antisense transcription of nuclear retroelements is detected by RIG-I, MDA5 and TLR3. Transition between nucleic acid species is shown with black solid arrows, detection – with dashed black arrows, inhibition – with red lines with blunt ends.

were described to induce SINE expression and therefore to activate MAVS-dependent NF-κB signaling [635]. Collectively, these findings indicate an important role of retroelements in innate immune response. Additionally to RNA species evolving from retroelements, their DNA-containing replication intermediates can also be linked to induction of IFN pathways. SAMHD1 inhibiting retroelement replication and exonuclease TREX1 that metabolizes reverse transcribed retroviral DNA control propagation of retroelements and their deficiency was associated with AGS and SLE autoimmune diseases making retroelements an important source of self-DNA and potential ligand for cGAS recognition [261, 636]. Apart from that deficiency of RNase H2 was linked to autoimmunity by inducing genome instability and formation of micronuclei that trigger cGAS activation [228, 526, 527]. Of note, activation of retroelements also results in chromosomal rearrangements and nuclear stress providing an additional link between their elevated expression and IFN response induced by cGAS activation [637, 638]. RNase H2 dysfunction resulting in incorporation of ribonucleotides into genomic DNA, as well as RNA:DNA species that are produced as intermediates of retroelement replication point to RNA:DNA species as a specific immunogenic nucleic acid. Indeed, *in vitro* and in human cell lines RNA:DNA hybrids could activate cGAS though less efficiently than dsDNA [639]. RNA:DNA hybrids were also described to occur during herpesvirus infection and as part of mtDNA arising from mtDNA replication (Figure 26) [640-642]. However, the sources of RNA:DNA hybrids and an exact nature of retroviral DNA-containing species that stimulate cGAS remain to be discovered. Similar to recognition of HIV-1 ssDNA, retroviral replication intermediates may possess specific secondary structures or could require other cellular co-factors for efficient cGAS activation. Despite a number of open questions it is tempting to broaden the research field of cGAS ligands from a single dsDNA to structured ssDNA and RNA:DNA hybrids. The involvement of cGAS in retrotransposon recognition, however, needs further research. Nevertheless, enhanced retroelement expression induced by type I IFN treatment [632], activation of retroelements upon viral infection and cancer, stress-induced mtDNA and chromatin exposure are tempting to be linked together into a general mechanism of cGAS activation initiated by viral infection and amplified using intracellular sources of DNA. Thus, a cross-talk between these mechanisms, as well as sources of cellular cGAS ligands and unique nucleic acid structures and co-factors are of a great interest for further studies.

6. Oligomerization in innate immunity

Signal transduction is usually understood as a chain of reactions or interactions serving as second messengers that lead to activation of enzymes and transcriptional or non-transcriptional effects. One of the well-studied examples includes β-adrenergic G-protein-coupled receptor (GPCR) which upon ligand binding undergoes conformational changes resulting in dissociation of G-proteins and activation of adenylyl cyclase that amplifies signals by producing several cAMP molecules and thus activates several molecules of protein kinase A that in turn phosphorylates cellular proteins [643]. Receptor tyrosine kinases (RTKs), on the other hand, possess a kinase activity that is stimulated upon ligand binding and directly initiate downstream kinase cascade [644]. Such pathways result in signal amplification through several rounds of second messenger production and a direct correlation between the induction strength, activated receptor lifetime and response. Unlike GPCRs and RTKs, innate immune receptors do not have enzymatic activity or are directly coupled to an enzyme. Instead, the majority of innate immune receptors initiate an assembly of huge signalosomes resulting in cooperative oligomerization of proteins after initial seed formation (Figure 27). Though GPCRs can also form oligomers on the cell surface, a monomeric
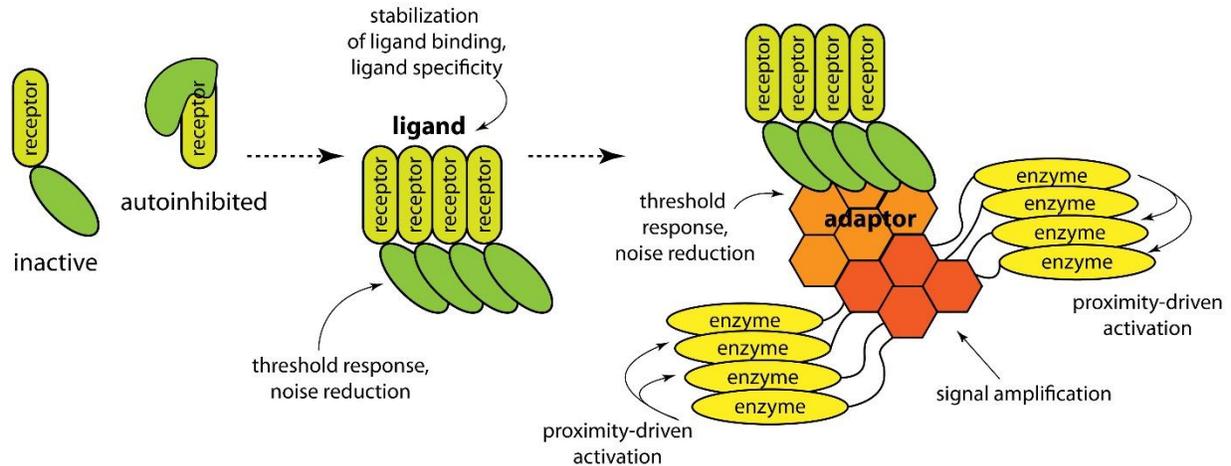
Figure 27 Schematic representation of innate immune signaling.
Oligomerization of innate immune receptors upon ligand binding stabilizes such interaction and ensures specific ligand recognition. Receptor oligomerization further nucleates adaptor clustering and recruitment of multiple effector enzymes resulting in signal amplification and enzyme transactivation by the neighboring enzymes. Receptor and downstream clustering ensures noise reduction and a potent dose-threshold response upon specific, strong and persistent stimulus.

receptor is fully capable of G protein activation and no propagation of its oligomeric state was observed in the downstream signaling events [645, 646]. Signal amplification in case of innate immune receptors emerges from subsequent oligomerization steps of receptor, adaptor and effector proteins resulting in recruitment of overstoichiometric numbers of signaling enzymes (Figure 27). Such a cooperativity leads to a sharp transition from non-active to activate states and dose-threshold response [647]. Since aberrant activation of innate immune programs might have deadly consequences, cooperative sensing ensures that the activating stimulus is strong enough to trigger the signaling. Time delay of such type of response provides additional mechanism to overcome transient and stochastic stimuli. Thus, cooperative signaling of innate immune receptors reduces noise signals and elicits immune response only upon strong and persistent activation.

Toll-like receptor (TLR) signaling starts from TIR-TIR homotypic domain interaction between activated TLR and MyD88 adaptor protein [648]. Oligomeric clustering of MyD88 TIR-domains around TLR promotes an assembly of its N-terminal death domains (DD) and their interaction with IRAK family members resulting in formation of DD signaling scaffold also called the "MyDDosome" [649, 650]. According to a crystal structure the MyDDosome represents a left-handed helix characteristic for all DD-fibrils with 6 MyD88 DDs followed by 4 IRAK4 and 4 IRAK2 kinases [649]. Increase of the local concentration of IRAK4 kinase domains results in its dimerization, autophosphorylation and enables IRAK4-mediated phosphorylation of IRAK1 or IRAK 2 [651, 652]. Intriguingly, formation of such oligomeric complex is extremely hierarchical and is determined by the shape and charge of interacting surfaces [649]. MyD88-DD is prone of oligomerization in higher protein concentrations whereas IRAK4-DD is monomeric in solution and assembles into a helix only using pre-formed MyD88-DD platform. Moreover, IRAK2-DD does not form a stable complex with IRAK4 or MyD88 alone, meaning the sequential propagation of DD-helix from more stable MyD88-DD association to addition of IRAK4 and IRAK2 or IRAK1 DD-domains that together stabilize the MyDDosome [649]. Oligomerization scaffold further contributes to recruitment of TRAF6 to the membrane and its trimerization. Unlike helical DD-

assembly, C-terminal TRAF6 domains forms trimers, whereas its N-terminal RING and ZF domains are dimeric [653-655]. This forms a 2-dimentional lattice that facilitates K63-linked polyubiquitination of IRAK1, NEMO and TRAF6 and downstream signaling [655, 656]. The TLR pathway represents a conjunction of three different types of oligomerization for signal amplification and proximity-driven activation of the participating enzymes and catalysis.

Unlike TLR signaling, inflammasome assembly depends mostly on helical fibrils formed by different types of DDs including PYD and CARD domains. Nucleated by substoichiometric amount of AIM2 or NLRs oligomers bringing their CARD or PYD domains into close proximity, oligomerization hierarchically propagates through adaptor ASC PYD filament formation and further pro-caspase-1 recruitment [87]. Strikingly, ASC PYD filaments have different helical symmetry from AIM2 PYD filaments [657]. Moreover, ASC PYD oligomers are able to nucleate oligomerization of death-effector domains (DEDs) of pro-caspase-8 indicating a high level of plasticity of this adaptor protein needed to accommodate different upstream and downstream proteins [658]. PYD domains of ASC create a cylinder-like structure surrounded by a ring of flexibly attached CARD domains [87, 89]. The CARD domains in turn nucleate oligomerization of pro-caspase-1 into CARD-filaments thus stimulating its autoactivation. Unlike right-handed ASC-PYD helical filament, CARD domains of caspase-1 assemble into a left-handed helix similar to MyDDosome [659]. Remarkably, propagation of ASC-PYD filament was reported to be limited by CARD domains likely due to CARD aggregation [87]. A current model of inflammasome activation thus includes NLR-nucleated aggregation of a short ASC filament with a consequent assembly and growth of pro-caspase-1 fibrils resulting in a nucleated star-shaped ternary complex.

Similar to inflammasomes, RLRs also utilize helical assemblies of CARD domains for signal transduction. RNA-bound RIG-I and MDA5 nucleate MAVS CARD filaments with their tetrameric CARDs resulting in aggregates of MAVS on mitochondria that recruit TBK1 and IRF3 directly or initiate NF-κB signaling through TRAF6 adaptor recruiting IKK and TAK1-kinase complexes. Intriguingly, MAVS has several features similar to cGAS signaling adaptor STING. Both STING and MAVS initiate the IFN production in overexpression system, require their membrane association for downstream cascades under physiological conditions and form punctate structures upon activation in perinuclear regions and mitochondria, respectively [141, 142, 215, 244]. Though oligomerization of MAVS and its localization in rod-shaped clusters on one side of mitochondria is well-studied, the nature of STING punctate structures remains enigmatic and needs further research [660]. STING oligomerization was observed *in vitro* for a full length protein, however, such aggregation might have emerged from stacking of the hydrophobic transmembrane domains [331]. Since STING itself does not possess any death domains and its cytosolic CDN-binding part is a stable dimer in solution, other factors or STING transmembrane region could influence its higher-order assemblies. Given direct binding of IRF3 by both STING and MAVS and a need of MAVS oligomerization for IRF3 activation [141, 661], the same mechanism is very tempting to propose for STING. For example, CTT clustering was proposed to increase the local concentrations of TBK1 and IRF3 and to enhance kinase activity and signal transduction similar to IRAK1 and IRAK4 proximity-driven activation in the MyDDosome [331, 649].

Downstream adaptor oligomerization in innate immune pathways is initiated by PRR clustering upon ligand binding. For example, dimerization of active TLRs or their conformational change upon ligand binding followed by clustering transmits oligomerization on MyDD88 and downstream adaptor assemblies [662-664]. Cytosolic PRRs do also aggregate for signal transduction. For example,

NAIP2/NLRC4 disc structures were shown to assemble hierarchically starting from ligand-bound NAIP2 followed by subsequent binding of NLRC4 monomers resulting in CARD domains clustering for ASC nucleation [617]. Unlike TLRs and NLRC4/NAIP2 recognizing small PAMPs, nucleic acid recognizing PRRs often use DNA or RNA molecules as a scaffold for oligomerization. AIM2 was found to form long filaments around dsDNA through interactions between its DNA-binding HIN domains and further aggregation of PYD domains [79, 657]. Negative stain micrographs also clearly show IFI16 fibrils formed on DNA [80]. Unlike AIM2, IFI16 DNA-binding HIN-domains alone do not interact within a fibril and its oligomerization emerges purely from PYD-domains association [80]. Such clustering of PYD domains results in nucleation of ASC filament and inflammasome activation [87]. Oligomerization was also observed for MDA5 and RIG-I that form continuous and limited filaments, respectively, through interactions between their ATPase domains [136, 138]. The importance of such an oligomerization for CARD-domains association can be illustrated by the additional requirement of K63-linked polyubiquitin chains for signal transduction by RIG-I in case of short dsRNAs (< 60 bp) [130, 665]. Short dsRNA and ssRNA do not introduce RIG-I oligomerization and therefore released CARD domains need additional co-factor – K63-linked polyubiquitin chains – to stabilize their tetramerization. In case of longer dsRNA, however, several RIG-I molecules are clustered in close proximity thus resulting in increased local concentration of CARD domains and their tetramerization without need for stabilizing ubiquitin binding [138].

Apart from driving downstream oligomerization events, such filament formation on nucleic acids plays an important role in discrimination between short non-stimulatory nucleic acids and long pathogen-derived or damage-associated species. For example, AIM2 was shown to recognize dsDNA fragments > 70 bp and binds 250-300 bp DNA fragments with highest affinity [79]. DNA-binding of AIM2 was found to sigmoidally increase with DNA length suggesting cooperativity [79]. *In vivo* studies showing cooperative increase of IL-1β production with DNA between 10 and 80 bp further confirm length-dependent activation of AIM2 inflammasome [76]. Similarly to AIM2 filaments, IFI16 was also found to cooperatively bind dsDNA with an optimal binding efficiency starting from 150 bp corresponding to 10 IFI16 protomers in cluster [80]. Interestingly, IFI16 was proposed to act in a switch-like manner: no binding was observed with < 60 bp DNA, however, all IFI16 molecules were bound with 60 bp DNA corresponding to IFI16 tetramer assembling with high cooperativity [80]. The RNA-sensor MDA5 that is known to stimulate IFN response uses a similar oligomerization mechanism to measure DNA length. MDA5 is activated with dsRNA fragments exceeding 1 kbp and shows increasing activation with increasing dsDNA length between 1 kbp and 7 kbp [115]. MDA5 was shown to have a slow filament nucleation and a fast dissociation initiated by ATP hydrolysis (Figure 28a) [125, 135, 618]. Though ATP hydrolysis occurs throughout the whole filament, MDA5 is stabilized in the middle by neighboring protomers and the dissociation mostly occurs at the filament ends (Figure 28b) [618]. On the short dsRNA such destabilization would lead to a complete MDA5 dissociation and inefficient nucleation, thus limiting MDA5 fibril formation and signaling. On the long dsRNA, however, MDA5 forms long dynamic filaments with less fraction of unstably bound terminal MDA5 molecules and therefore slower dissociation and higher association rates, because no limiting *de novo* nucleation is needed (Figure 28b) [618]. Thus, different PRRs utilize the same oligomerization principle to specifically recognize pathogen-derived nucleic acids and to avoid recognition of non-pathogenic nucleic acid species.

Surprisingly, previously proposed to be monomeric or form dimers, cGAS also seems to adapt oligomerization mechanism for self- vs. non-self discrimination. In the study presented in this thesis

cGAS was found to oligomerize into fibrils on long DNA fragments [600]. Analogically to PRRs described above, we showed that such oligomerization enhances DNA-binding and thus the active conformation of cGAS. Similar to dynamic MDA5 assembly, the binding of the first cGAS dimer on DNA may represent a slow nucleation step that is followed by more rapid fibril elongation. Though some sensors like AIM2, RIG-I , NAIP2 and NLRC4 were reported to reside in autoinhibited state before ligand binding, others like MDA5, IFI16 and cGAS seem to use the same oligomerization mechanism to counteract spontaneous activation, since only formation of higher-order assemblies were shown to enhance the ligand binding and receptor activation [77, 80, 129, 133, 618, 666]. Instead of ATP hydrolysis stimulating MDA5 dissociation and driving its specificity to long (> 1kb) dsRNA fragments, cGAS may use weak dimeric interaction as a source of filament destabilization resulting in specific recognition of long pathogenic dsDNA species (Figure 28a). Based on MDA5 RNA-sensing model, we hypothesize that cGAS dimers may also partly disassemble within the fibril but not dissociate due to the stabilization by the neighboring cGAS dimers (Figure 28b) [600]. Given a number of similarities between cGAS and MDA5 oligomerization, a possibility of cGAS fibril editing is of great interest. MDA5 oligomerization occurring from different seeds may result in oligomers proceeding in different directions. These MDA5 stretches were found to undergo disassembly in order to generate a unidirectional stable oligomer [618]. Since a cGAS filament might also be initiated from different starting points, it is
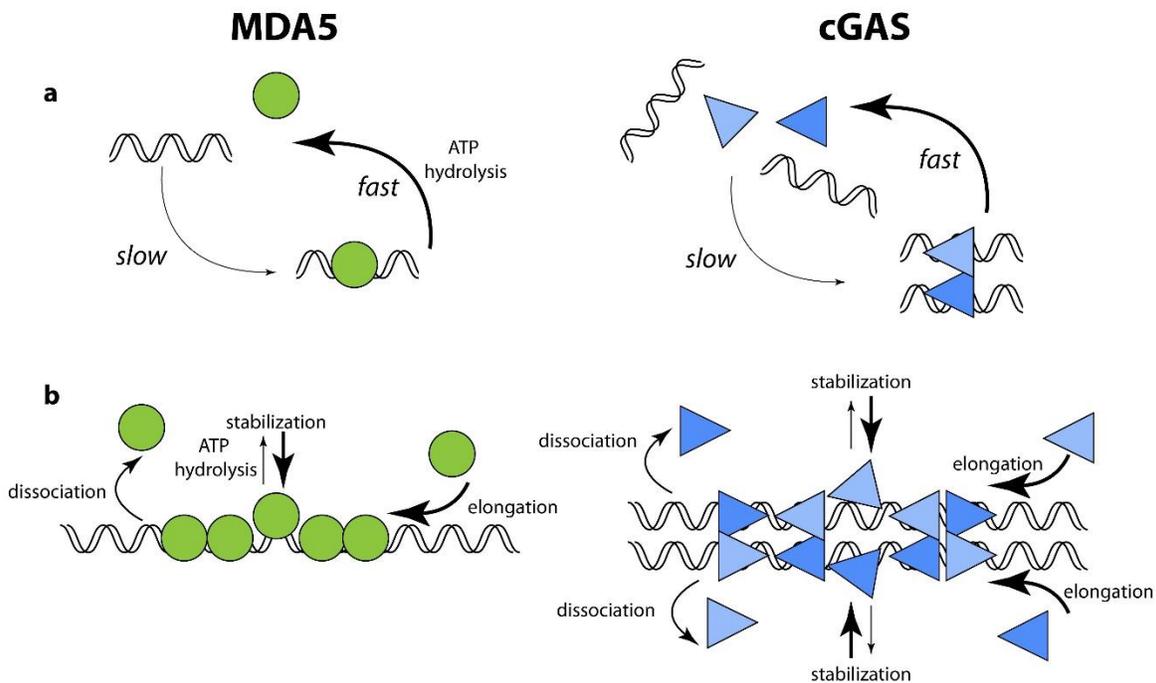
Figure 28 Similarities between MDA5 and cGAS assemblies on short and long nucleic acid species.
    (a)   Slow association and fast dissociation rates of MDA5 (green circles) and cGAS dimers (blue triangles) binding result in transient and unstable binding of short dsRNA and dsDNA fragments, respectively.
    (b)   Longer nucleic acid species allow oligomerization of MDA5 on dsRNA and cGAS dimers on two parallel dsDNA strands. Monomers within the fibrils are mutually stabilized and dissociation mostly occurs on the ends of oligomers. MDA5 and cGAS oligomers nucleate fibril growth and result in effective association of further monomers or dimers.
Thick and thin black arrows represent fast and slow processes, respectively. Dashed arrows represent a direction of events.
Partly adopted from *del Toro Duany Y., et al., 2015* [667].

intriguing to know, whether similar processes exist to remove gaps and inconsistencies within the cGAS ladders.

Apart from stabilization of the active conformation common for nearly all PRRs, oligomerization may have additional effector functions. For example, RIG-I and MDA5 were found to remove viral proteins like influenza A NS1 from dsRNA and restrict viral replication [116]. Moreover, RIG-I was recently found to bind 5' end of hepatitis B virus (HBV) and therefore to counteract viral polymerase recruitment and reverse transcription of HBV [105]. Similar to this, higher-order oligomers of PRRs might interfere with viral replication due to competition for nucleic acid binding. Whether such activities can be featured to cGAS and other PRRs, remains an open question and needs further research.

Oligomerization mechanism is also applied for regulation of innate immune response. For example, cellular anti-apoptotic FLICE/Casp-8-like inhibitory proteins (cFLIP) containing tandem DED domains were shown to incorporate into caspase-8 filament formed upon Fas receptor activation and death-inducing signaling complex (DISC) or inflammasome activation [668]. A short version of cFLIP – cFLIP$_S$ – was found to interfere with pro-caspase-8 clustering and to reduce local concentration of caspase-8 in the filament and therefore counteracting its dimerization and activation. A long form of cFLIP – cFLIP$_L$ – on the other side, additionally to DED possesses a caspase-like domain able to heterodimerize with caspase-8. Incorporation of cFLIP$_L$ in caspase-8 filament can therefore stimulate caspase-8 activation and promote cell death. However, since such heterodimers have lower activity in comparison to caspase-8 dimers, cFLIP$_L$ inhibits caspase-8 activity in higher concentrations. Viral anti-apoptotic protein MC159 from *Molluscum contagiosum* virus can also associate with caspase-8 filament, but, unlike cFLIP$_S$, blocks the filament propagation and caspase-8 activation due to its defective interfaces that prevent further oligomerization [668]. The similar capping mechanism was shown for human inhibitor of CARD (INCA), which inhibits the inflammasome assembly. Having only a CARD domain with defective interfaces
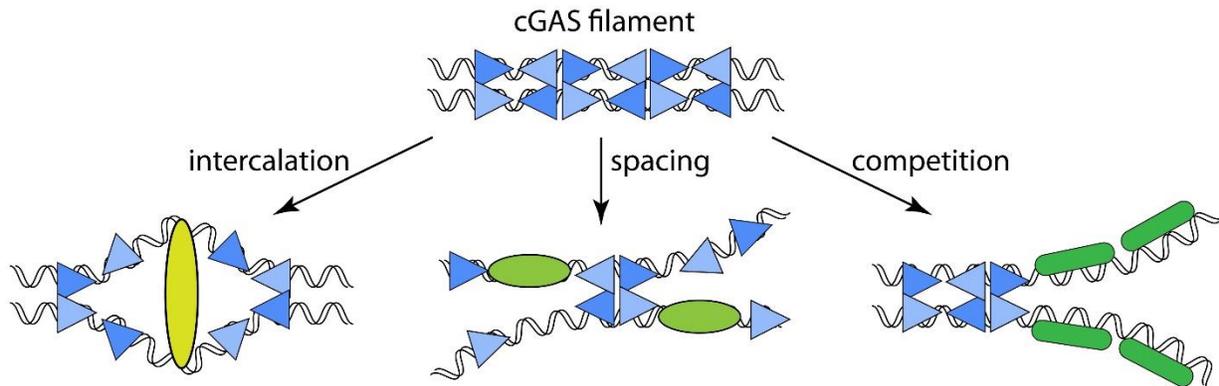
Figure 29 Schematic representation of possible mechanisms of cGAS inhibition based on interfering with cGAS oligomerization.
Cellular or viral inhibitor proteins (green) might intercalate between two DNA strands in cGAS oligomer (on the left) increasing the distance and thus disrupting cGAS dimerization and activation. They might incorporate within cGAS filament (in the middle) increasing a spacing between cGAS dimers (blue triangles) leading to oligomer destabilization and dissociation of dimers. Inhibitory proteins might also compete for DNA binding with cGAS (on the right) thus preventing cGAS filament growth and disassembling existing oligomers.

INCA interacts with caspase-1 filament and blocks further filament growth [659]. Inhibition via oligomerization interfering was also found for mouse HIN-domain containing protein p202. P202 was shown to interact with HIN-domain of AIM2 within AIM2 oligomer and to separate AIM2 PYD domains preventing PYD clustering and inflammasome activation [669]. Though no p202 orthologs could be found in humans, a similar mechanism of inflammasome inactivation may be used by other human proteins. Since oligomerization destabilization is a common mechanism for inhibition of diverse cellular pathways, it is an important question whether cGAS can also be regulated in a similar way (Figure 29).

As a single $cGAS_2:DNA_2$ complex is rather unstable, some cellular or viral cGAS inhibitors might exist that interfere with cGAS packing by intercalating between two DNA strands within cGAS filaments and thus increase the distance between DNA and disrupt cGAS dimers. Another option, analogous to other oligomers, would be an incorporation of inhibitory factors into cGAS filament and increasing spacing between neighboring cGAS dimers that would lead to a less efficient DNA positioning and thus to destabilization of cGAS dimers and cGAS inactivation. Moreover, a simple competition for DNA binding or an active translocation along DNA could be used to disassemble cGAS oligomers and inhibit cGAMP production and IFN response (Figure 29). However, these mechanisms are highly speculative and further research is needed to study cGAS inhibition and involvement of cGAS oligomerization in its regulation.

Taken together, nucleation-oligomerization mechanism of cGAS activation goes in line with similar mechanisms proposed for various PRRs. Unlike other receptors using their aggregation to nucleate downstream adaptors that in turn stimulate proximity-driven enzyme activation, cGAS combines both pattern recognition and enzymatic function. For this reason its oligomerization is only required for stabilization of its active conformation. Another distinct feature of cGAS includes its higher-order assembly without protein-protein contacts. A unique cGAS dimeric structure seems to be sufficient for positioning of DNA strands in nearly parallel manner that is favorable for subsequent cGAS dimer binding and therefore mediates cGAS oligomerization through DNA arrangement rather than direct lateral cGAS contacts [600]. Despite the unique features described above, the oligomerization mechanism of cGAS further confirms importance of higher-order filaments in innate immune signaling and suggests a strong, specific and dose-threshold response to pathogen- or damage-associated cytosolic DNA.

References

1.   Du, B., et al., *Targeting Toll-like receptors against cancer.* Journal of Cancer Metastasis and Treatment; Vol 2, No 12 (2016), 2016.

2.   Hoffmann, J.A., et al., *Phylogenetic Perspectives in Innate Immunity.* Science, 1999. 284(5418): p. 1313-1318.

3.   Hartmann, G., *Nucleic acid immunity.* Adv. Immunol., 2017. 133: p. 121-169.

4.   Medzhitov, R. and C. Janeway, Jr., *Innate immune recognition: mechanisms and pathways.* Immunological Reviews, 2000. 173(1): p. 89-97.

5.   Long, E.O., *Regulation of immune responses through inhibitory receptors.* Annual Review of Immunology, 1999. 17(1): p. 875-904.

6.   Charles A. Janeway, J. and R. Medzhitov, *Innate Immune Recognition.* Annual Review of Immunology, 2002. 20(1): p. 197-216.

7.   Fearon, D.T. and R.M. Locksley, *The Instructive Role of Innate Immunity in the Acquired Immune Response.* Science, 1996. 272(5258): p. 50-54.

8.   Medzhitov , R. and C.J. Janeway *Innate Immunity.* New England Journal of Medicine, 2000. 343(5): p. 338-344.

9.   Akira S., T.K., Kaisho T., *Toll-like receptors: critical proteins linking innate and acquired immunity.* Nat. Immunol., 2001. 2: p. 675-680.

10.   Medzhitov, R. and C.A. Janeway, *Innate immunity: impact on the adaptive immune response.* Current Opinion in Immunology, 1997. 9(1): p. 4-9.

11.   Charles A Janeway, J., Paul Travers, Mark Walport, and Mark J Shlomchik, *Immunobiology: The Immune System in Health and Disease. 5th edition.* 2001, New York: Garland Science.

12.   Ahn, J. and G. Barber, *Self-DNA, STING-dependent signaling and the origins of autoinflammatory disease*. Vol. 31. 2014.

13.   Sharma, S., et al., *Triggering the Interferon Antiviral Response Through an IKK-Related Pathway.* Science, 2003. 300(5622): p. 1148-1151.

14.   Sato, M., et al., *Distinct and Essential Roles of Transcription Factors IRF-3 and IRF-7 in Response to Viruses for IFN-α/β Gene Induction.* Immunity, 2000. 13(4): p. 539-548.

15.   Fitzgerald, K.A., et al., *LPS-TLR4 Signaling to IRF-3/7 and NF-κB Involves the Toll Adapters TRAM and TRIF*. The Journal of Experimental Medicine, 2003. 198(7): p. 1043-1055.

16.   Biron, C.A., *Interferons α and β as Immune Regulators—A New Look.* Immunity, 2001. 14(6): p. 661-664.

17.   Darnell, J., I. Kerr, and G. Stark, *Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins.* Science, 1994. 264(5164): p. 1415-1421.

18.   Le Page, C., et al., *Interferon activation and innate immunity.* Reviews in immunogenetics, 2000. 2(3): p. 374-386.

19.   Goodman, A.G., et al., *The Alpha/Beta Interferon Receptor Provides Protection against Influenza Virus Replication but Is Dispensable for Inflammatory Response Signaling.* Journal of Virology, 2010. 84(4): p. 2027-2037.

20.   Thornberry, N.A., et al., *A novel heterodimeric cysteine protease is required for interleukin-1[beta]processing in monocytes.* Nature, 1992. 356(6372): p. 768-774.

21.   Ghayur, T., et al., *Caspase-1 processes IFN-γ-inducing factor and regulates LPS-induced IFN-γ production*. Vol. 386. 1997. 619-23.

22.   Brennan, M.A. and B.T. Cookson, *Salmonella induces macrophage death by caspase-1-dependent necrosis.* Molecular Microbiology, 2000. 38(1): p. 31-40.

23.   Aliprantis, A.O., et al., *Cell Activation and Apoptosis by Bacterial Lipoproteins Through Toll-like Receptor-2.* Science, 1999. 285(5428): p. 736-739.

24.   Aliprantis, A.O., et al., *The apoptotic signaling pathway activated by Toll-like receptor-2.* The EMBO Journal, 2000. 19(13): p. 3325-3336.

25.   Sun, R., et al., *Toll-like Receptor 3 (TLR3) Induces Apoptosis via Death Receptors and Mitochondria by Up-regulating the Transactivating p63 Isoform α (TAP63α).* The Journal of Biological Chemistry, 2011. 286(18): p. 15918-15928.

26.   Zhang, Y. and J.B. Bliska, *Role of Toll-Like Receptor Signaling in the Apoptotic Response of Macrophages to Yersinia Infection.* Infection and Immunity, 2003. 71(3): p. 1513-1519.

27.   Ruckdeschel, K., et al., *Signaling of Apoptosis through TLRs Critically Involves Toll/IL-1 Receptor Domain-Containing Adapter Inducing IFN-β, but Not MyD88, in Bacteria-Infected Murine Macrophages.* The Journal of Immunology, 2004. 173(5): p. 3320-3328.

28.   Upton, J.W., W.J. Kaiser, and E.S. Mocarski, *DAI complexes with RIP3 to mediate virus-induced programmed necrosis that is targeted by murine cytomegalovirus vIRA.* Cell host & microbe, 2012. 11(3): p. 290-297.

29.   Gaidt, M.M., et al., *The DNA Inflammasome in Human Myeloid Cells Is Initiated by a STING-Cell Death Program Upstream of NLRP3.* Cell, 2017. 171(5): p. 1110-1124.e18.

30.   Ikeda, K., et al., *Serum lectin with known structure activates complement through the classical pathway.* Journal of Biological Chemistry, 1987. 262(16): p. 7451-7454.

31.   Galiana-Arnoux, D., et al., *Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila.* Nat Immunol, 2006. 7(6): p. 590-597.

32.   van Rij, R.P., et al., *The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster.* Genes & Development, 2006. 20(21): p. 2985-2995.

33.   Seo, Gil J., et al., *Reciprocal Inhibition between Intracellular Antiviral Signaling and the RNAi Machinery in Mammalian Cells.* Cell host & microbe, 2013. 14(4): p. 435-445.

34.   Schlee, M. and G. Hartmann, *Discriminating self from non-self in nucleic acid sensing.* Nat Rev Immunol, 2016. 16(9): p. 566-580.

35.   Anderson, K.V., L. Bokla, and C. Nüsslein-Volhard, *Establishment of dorsal-ventral polarity in the drosophila embryo: The induction of polarity by the Toll gene product.* Cell, 1985. 42(3): p. 791-798.

36. Lemaitre, B., et al., *The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in Drosophila adults.* Cell, 1996. 86(6): p. 973-983.

37. Akira, S., S. Uematsu, and O. Takeuchi, *Pathogen Recognition and Innate Immunity.* Cell, 2006. 124(4): p. 783-801.

38. Alexopoulou, L., et al., *Recognition of double-stranded RNA and activation of NF-[kappa]B by Toll-like receptor 3.* Nature, 2001. 413(6857): p. 732-738.

39. Diebold, S.S., et al., *Innate Antiviral Responses by Means of TLR7-Mediated Recognition of Single-Stranded RNA.* Science, 2004. 303(5663): p. 1529-1531.

40. Heil, F., et al., *Species-Specific Recognition of Single-Stranded RNA via Toll-like Receptor 7 and 8.* Science, 2004. 303(5663): p. 1526-1529.

41. Hemmi, H., et al., *A Toll-like receptor recognizes bacterial DNA.* Nature, 2000. 408(6813): p. 740-745.

42. Dinarello, C.A., *Immunological and Inflammatory Functions of the Interleukin-1 Family.* Annual Review of Immunology, 2009. 27(1): p. 519-550.

43. Sims, J.E. and D.E. Smith, *The IL-1 family: regulators of immunity.* Nature Reviews Immunology, 2010. 10: p. 89.

44. Yarovinsky, F., et al., *TLR11 Activation of Dendritic Cells by a Protozoan Profilin-Like Protein.* Science, 2005. 308(5728): p. 1626-1629.

45. Mathur, R., et al., *A mouse model of Salmonella typhi infection.* Cell, 2012. 151(3): p. 590-602.

46. Koblansky, A., *Recognition of profilin by Toll-like receptor 12 is critical for host resistance to Toxoplasma gondii.* Vol. 30. 2013. 119-130.

47. Hidmark, A., A. von Saint Paul, and A.H. Dalpke, *Cutting Edge: TLR13 Is a Receptor for Bacterial RNA.* The Journal of Immunology, 2012. 189(6): p. 2717-2721.

48. Matsushima, N., et al., *Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors.* Vol. 8. 2007. 124.

49. Ozinsky, A., et al., *The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between Toll-like receptors.* Proceedings of the National Academy of Sciences, 2000. 97(25): p. 13766-13771.

50. Takada, E., et al., *C-terminal LRRs of human Toll-like receptor 3 control receptor dimerization and signal transmission.* Molecular Immunology, 2007. 44(15): p. 3633-3640.

51. Zhang, H., et al., *Integrin-nucleated Toll-like receptor (TLR) dimerization reveals subcellular targeting of TLRs and distinct mechanisms of TLR4 activation and signaling.* FEBS Letters, 2002. 532(1): p. 171-176.

52. Ivičak-Kocjan, K., et al., *Determination of the physiological 2:2 TLR5:flagellin activation stoichiometry revealed by the activity of a fusion receptor.* Vol. 435. 2013.

53. Ohto, U., et al., *Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9.* Nature, 2015. 520(7549): p. 702-705.

54.   Wang, J., et al., *The Functional Effects of Physical Interactions among Toll-like Receptors 7, 8, and 9.* Journal of Biological Chemistry, 2006. 281(49): p. 37427-37434.

55.   G Motshwene, P., et al., *An Oligomeric Signaling Platform Formed by the Toll-like Receptor Signal Transducers MyD88 and IRAK-4*. Vol. 284. 2009. 25404-11.

56.   Kawasaki, T. and T. Kawai, *Toll-Like Receptor Signaling Pathways.* Frontiers in Immunology, 2014. 5: p. 461.

57.   Honda, K., et al., *Role of a transductional-transcriptional processor complex involving MyD88 and IRF-7 in Toll-like receptor signaling.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(43): p. 15416-15421.

58.   Oshiumi, H., et al., *TICAM-1, an adaptor molecule that participates in Toll-like receptor 3-mediated interferon-[beta] induction.* Nat Immunol, 2003. 4(2): p. 161-167.

59.   Fitzgerald, K.A., et al., *IKK[epsi] and TBK1 are essential components of the IRF3 signaling pathway.* Nat Immunol, 2003. 4(5): p. 491-496.

60.   Cheng, Z., et al., *Distinct single-cell signaling characteristics are conferred by the MyD88 and TRIF pathways during TLR4 activation.* Science Signaling, 2015. 8(385): p. ra69-ra69.

61.   Harton, J.A., et al., *Cutting Edge: CATERPILLER: A Large Family of Mammalian Genes Containing CARD, Pyrin, Nucleotide-Binding, and Leucine-Rich Repeat Domains.* The Journal of Immunology, 2002. 169(8): p. 4088-4093.

62.   Kummer, J.A., et al., *Inflammasome Components NALP 1 and 3 Show Distinct but Separate Expression Profiles in Human Tissues Suggesting a Site-specific Role in the Inflammatory Response.* Journal of Histochemistry & Cytochemistry, 2007. 55(5): p. 443-452.

63.   Hamada, M., et al., *The Complex NOD-Like Receptor Repertoire of the Coral Acropora digitifera Includes Novel Domain Combinations.* Molecular Biology and Evolution, 2013. 30(1): p. 167-176.

64.   Initiative, T.A.G., *Analysis of the genome sequence of the flowering plant Arabidopsis thaliana.* Nature, 2000. 408(6814): p. 796-815.

65.   Hornung, V., et al., *AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC.* Nature, 2009. 458(7237): p. 514-518.

66.   Fernandes-Alnemri, T., et al., *AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA.* Nature, 2009. 458(7237): p. 509-513.

67.   Kerur, N., et al., *IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi sarcoma associated herpesvirus infection.* Cell host & microbe, 2011. 9(5): p. 363-375.

68.   Unterholzner, L., et al., *IFI16 is an innate immune sensor for intracellular DNA.* Nat Immunol, 2010. 11(11): p. 997-1004.

69.   Miao, E.A., et al., *Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome.* Proceedings of the National Academy of Sciences, 2010. 107(7): p. 3076-3080.

70.   Gross, O., et al., *Syk kinase signalling couples to the Nlrp3 inflammasome for anti-fungal host defence.* Nature, 2009. 459(7245): p. 433-436.

71.   Mariathasan, S., et al., *Cryopyrin activates the inflammasome in response to toxins and ATP.* Nature, 2006. 440(7081): p. 228-232.

72.   Craven, R.R., et al., *Staphylococcus aureus α-Hemolysin Activates the NLRP3-Inflammasome in Human and Mouse Monocytic Cells.* PLoS ONE, 2009. 4(10): p. e7446.

73.   Halle, A., et al., *The NALP3 inflammasome is involved in the innate immune response to amyloid-[beta].* Nat Immunol, 2008. 9(8): p. 857-865.

74.   Gasse, P., et al., *Uric Acid Is a Danger Signal Activating NALP3 Inflammasome in Lung Injury Inflammation and Fibrosis.* American Journal of Respiratory and Critical Care Medicine, 2009. 179(10): p. 903-913.

75.   Zhou, R., et al., *Thioredoxin-interacting protein links oxidative stress to inflammasome activation.* Nat Immunol, 2010. 11(2): p. 136-140.

76.   Jin, T., et al., *Structures of the HIN Domain:DNA Complexes Reveal Ligand Binding and Activation Mechanisms of the AIM2 Inflammasome and IFI16 Receptor.* Immunity, 2012. 36(4): p. 561-571.

77.   Hu, Z., et al., *Crystal Structure of NLRC4 Reveals Its Autoinhibition Mechanism.* Science, 2013. 341(6142): p. 172-175.

78.   Poyet, J.-L., et al., *Identification of Ipaf, a Human Caspase-1-activating Protein Related to Apaf-1.* Journal of Biological Chemistry, 2001. 276(30): p. 28309-28313.

79.   Morrone, S.R., et al., *Assembly-driven activation of the AIM2 foreign-dsDNA sensor provides a polymerization template for downstream ASC.* Nature Communications, 2015. 6: p. 7827.

80.   Morrone, S.R., et al., *Cooperative assembly of IFI16 filaments on dsDNA provides insights into host defense strategy.* Proceedings of the National Academy of Sciences, 2014. 111(1): p. E62-E71.

81.   Riedl, S.J., et al., *Structure of the apoptotic protease-activating factor 1 bound to ADP.* Nature, 2005. 434(7035): p. 926-933.

82.   Qi, S., et al., *Crystal Structure of the Caenorhabditis elegans Apoptosome Reveals an Octameric Assembly of CED-4.* Cell, 2010. 141(3): p. 446-457.

83.   Faustin, B., et al., *Reconstituted NALP1 Inflammasome Reveals Two-Step Mechanism of Caspase-1 Activation.* Molecular Cell, 2007. 25(5): p. 713-724.

84.   Martinon, F., K. Burns, and J. Tschopp, *The Inflammasome.* Molecular Cell, 2002. 10(2): p. 417-426.

85.   Broz, P., et al., *Redundant roles for inflammasome receptors NLRP3 and NLRC4 in host defense against Salmonella.* The Journal of Experimental Medicine, 2010. 207(8): p. 1745-1755.

86.   Cai, X., et al., *Prion-like Polymerization Underlies Signal Transduction in Antiviral Immune Defense and Inflammasome Activation.* Cell, 2014. 156(6): p. 1207-1222.

87.   Lu, A., et al., *Unified Polymerization Mechanism for the Assembly of ASC-Dependent Inflammasomes.* Cell, 2014. 156(6): p. 1193-1206.

88.   Proell, M., et al., *The CARD plays a critical role in ASC foci formation and inflammasome signaling.* The Biochemical journal, 2013. 449(3): p. 613-621.

89.   Sborgi, L., et al., *Structure and assembly of the mouse ASC inflammasome by combined NMR spectroscopy and cryo-electron microscopy.* Proceedings of the National Academy of Sciences of the United States of America, 2015. 112(43): p. 13237-13242.

90.   Mariathasan, S., et al., *Differential activation of the inflammasome by caspase-1 adaptors ASC and Ipaf.* Nature, 2004. 430(6996): p. 213-218.

91.   Ketelut-Carneiro, N., et al., *IL-18 Triggered by the Nlrp3 Inflammasome Induces Host Innate Resistance in a Pulmonary Model of Fungal Infection.* The Journal of Immunology, 2015. 194(9): p. 4507-4517.

92.   Shi, J., et al., *Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death.* Nature, 2015. 526(7575): p. 660-665.

93.   Liu, X., et al., *Inflammasome-activated gasdermin D causes pyroptosis by forming membrane pores.* Nature, 2016. 535(7610): p. 153-158.

94.   Man, S.M. and T.-D. Kanneganti, *Converging roles of caspases in inflammasome activation, cell death and innate immunity.* Nature reviews. Immunology, 2016. 16(1): p. 7-21.

95.   von Moltke, J., et al., *Rapid induction of inflammatory lipid mediators by the inflammasome in vivo.* Nature, 2012. 490(7418): p. 107-111.

96.   Sokolovska, A., et al., *Activation of caspase-1 by the NLRP3 inflammasome regulates the NADPH oxidase NOX2 to control phagosome function.* Nature immunology, 2013. 14(6): p. 543-553.

97.   Jabir, Majid S., et al., *Caspase-1 Cleavage of the TLR Adaptor TRIF Inhibits Autophagy and β-Interferon Production during Pseudomonas aeruginosa Infection.* Cell host & microbe, 2014. 15(2): p. 214-227.

98.   Shao, W., et al., *The Caspase-1 Digestome Identifies the Glycolysis Pathway as a Target during Infection and Septic Shock.* Journal of Biological Chemistry, 2007. 282(50): p. 36321-36329.

99.   Wilson, J.E., et al., *Inflammasome-independent role of AIM2 in suppressing colon tumorigenesis via DNA-PK and Akt.* Nat Med, 2015. 21(8): p. 906-913.

100.   Bruchard, M., et al., *The receptor NLRP3 is a transcriptional regulator of TH2 differentiation.* Nat Immunol, 2015. 16(8): p. 859-870.

101.   Anand, P.K., et al., *NLRP6 negatively regulates innate immunity and host defence against bacterial pathogens.* Nature, 2012. 488(7411): p. 389-393.

102.   Wang, P., et al., *Nlrp6 regulates intestinal antiviral innate immunity.* Science, 2015. 350(6262): p. 826-830.

103.   Liu, G., et al., *Influenza A Virus Panhandle Structure Is Directly Involved in RIG-I Activation and Interferon Induction.* Journal of Virology, 2015. 89(11): p. 6067-6079.

104.   Choi, M.K., et al., *A selective contribution of the RIG-I-like receptor pathway to type I interferon responses activated by cytosolic DNA.* Proceedings of the National Academy of Sciences, 2009. 106(42): p. 17870-17875.

105.   Sato, S., et al., *The RNA Sensor RIG-I Dually Functions as an Innate Sensor and Direct Antiviral Factor for Hepatitis B Virus.* Immunity, 2015. 42(1): p. 123-132.

106.   Yoneyama, M., et al., *Shared and Unique Functions of the DExD/H-Box Helicases RIG-I, MDA5, and LGP2 in Antiviral Innate Immunity.* The Journal of Immunology, 2005. 175(5): p. 2851-2858.

107.   Binder, M., et al., *Molecular mechanism of signal perception and integration by the innate immune sensor retinoic acid-inducible gene-I (RIG-I).* Journal of Biological Chemistry, 2011. 286(31): p. 27278-27287.

108.   Hornung, V., et al., *5'-Triphosphate RNA is the ligand for RIG-I.* Science, 2006.

109.   Goubau, D., et al., *Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates.* Nature, 2014. 514(7522): p. 372-375.

110.   Yoneyama, M., et al., *The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses.* Nat Immunol, 2004. 5(7): p. 730-737.

111.   Weber, M., et al., *Incoming RNA virus nucleocapsids containing a 5'-triphosphorylated genome activate RIG-I and antiviral signaling.* Cell host & microbe, 2013. 13(3): p. 336-346.

112.   Ablasser, A., et al., *RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate.* Nat Immunol, 2009. 10(10): p. 1065-1072.

113.   Gitlin, L., et al., *Essential role of mda-5 in type I IFN responses to polyriboinosinic:polyribocytidylic acid and encephalomyocarditis picornavirus.* Proceedings of the National Academy of Sciences, 2006. 103(22): p. 8459-8464.

114.   Runge, S., et al., *In Vivo Ligands of MDA5 and RIG-I in Measles Virus-Infected Cells.* PLOS Pathogens, 2014. 10(4): p. e1004081.

115.   Kato, H., et al., *Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid–inducible gene-I and melanoma differentiation–associated gene 5.* The Journal of Experimental Medicine, 2008. 205(7): p. 1601-1610.

116.   Yao, H., et al., *ATP-Dependent Effector-like Functions of RIG-I-like Receptors.* Molecular Cell, 2015. 58(3): p. 541-548.

117.   Li, X., et al., *The RIG-I-like Receptor LGP2 Recognizes the Termini of Double-stranded RNA.* Journal of Biological Chemistry, 2009. 284(20): p. 13881-13891.

118.   Pippig, D.A., et al., *The regulatory domain of the RIG-I family ATPase LGP2 senses double-stranded RNA.* Nucleic Acids Research, 2009. 37(6): p. 2014-2025.

119.   Deddouche, S., et al., *Identification of an LGP2-associated MDA5 agonist in picornavirus-infected cells.* eLife, 2014. 3: p. e01535.

120.   Bruns, Annie M., et al., *The Innate Immune Sensor LGP2 Activates Antiviral Signaling by Regulating MDA5-RNA Interaction and Filament Assembly.* Molecular Cell, 2014. 55(5): p. 771-781.

121.   Lu, C., et al., *The Structural Basis of 5' Triphosphate Double-stranded RNA Recognition by RIG-I C-terminal Domain.* Structure (London, England : 1993), 2010. 18(8): p. 1032-1043.

122.   Wang, Y., et al., *Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I.* Nat Struct Mol Biol, 2010. 17(7): p. 781-787.

123.   Uchikawa, E., et al., *Structural Analysis of dsRNA Binding to Anti-viral Pattern Recognition Receptors LGP2 and MDA5.* Molecular Cell, 2016. 62(4): p. 586-602.

124.   Li, X., et al., *Structural basis of double-stranded RNA recognition by the RIG-I like receptor MDA5.* Archives of Biochemistry and Biophysics, 2009. 488(1): p. 23-33.

125.	Peisley, A., et al., *Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition.* Proceedings of the National Academy of Sciences, 2011. 108(52): p. 21010-21015.

126.	Louber, J., et al., *Kinetic discrimination of self/non-self RNA by the ATPase activity of RIG-I and MDA5.* BMC Biology, 2015. 13(1): p. 54.

127.	Lässig, C., et al., *ATP hydrolysis by the viral RNA sensor RIG-I prevents unintentional recognition of self-RNA.* eLife, 2015. 4: p. e10859.

128.	Jiang, F., et al., *Structural basis of RNA recognition and activation by innate immune receptor RIG-I.* Nature, 2011. 479(7373): p. 423-427.

129.	Kowalinski, E., et al., *Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA.* Cell, 2011. 147.

130.	Zeng, W., et al., *Reconstitution of the RIG-I Pathway Reveals a Signaling Role of Unanchored Polyubiquitin Chains in Innate Immunity.* Cell, 2010. 141(2): p. 315-330.

131.	Gack, M.U., et al., *TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity.* Nature, 2007. 446(7138): p. 916-920.

132.	Peisley, A., et al., *Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I.* Nature, 2014. 509(7498): p. 110-114.

133.	Saito, T., et al., *Regulation of innate antiviral defenses through a shared repressor domain in RIG-I and LGP2.* Proceedings of the National Academy of Sciences of the United States of America, 2007. 104(2): p. 582-587.

134.	Yoneyama, M., et al., *Viral RNA detection by RIG-I-like receptors.* Curr Opin Immunol., 2015. 32C.

135.	Berke, I.C. and Y. Modis, *MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA.* The EMBO Journal, 2012. 31(7): p. 1714-1726.

136.	Wu, B., et al., *Structural Basis for dsRNA Recognition, Filament Formation, and Antiviral Signal Activation by MDA5.* Cell, 2012. 152(1): p. 276-289.

137.	Uchikawa, E., et al., *Structural Analysis of dsRNA Binding to Anti-viral Pattern Recognition Receptors LGP2 and MDA5.* Molecular Cell. 62(4): p. 586-602.

138.	Peisley, A., et al., *RIG-I Forms Signaling-Competent Filaments in an ATP-Dependent, Ubiquitin-Independent Manner.* Molecular Cell, 2013. 51(5): p. 573-583.

139.	Seth, R.B., et al., *Identification and Characterization of MAVS, a Mitochondrial Antiviral Signaling Protein that Activates NF-κB and IRF3.* Cell, 2005. 122(5): p. 669-682.

140.	Wu, B., et al., *Molecular Imprinting as a Signal-Activation Mechanism of the Viral RNA Sensor RIG-I.* Molecular Cell, 2014. 55(4): p. 511-523.

141.	Hou, F., et al., *MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response.* Cell., 2011. 146.

142.	Kawai, T., et al., *IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction.* Nat Immunol, 2005. 6(10): p. 981-988.

143.   Liu, S., et al., *Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation.* Science, 2015. 347(6227).

144.   Meylan, E., et al., *Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus.* Nature, 2005. 437(7062): p. 1167-1172.

145.   Saha, S.K., et al., *Regulation of antiviral responses by a direct and specific interaction between TRAF3 and Cardif.* The EMBO Journal, 2006. 25(14): p. 3257-3263.

146.   Guo, B. and G. Cheng, *Modulation of the Interferon Antiviral Response by the TBK1/IKKi Adaptor Protein TANK.* Journal of Biological Chemistry, 2007. 282(16): p. 11817-11826.

147.   Liu, S., et al., *MAVS recruits multiple ubiquitin E3 ligases to activate antiviral signaling cascades.* eLife, 2013. 2: p. e00785.

148.   Hovanessian, A.G., R.E. Brown, and I.M. Kerr, *Synthesis of low molecular weight inhibitor of protein synthesis with enzyme from interferon-treated cells.* Nature, 1977. 268(5620): p. 537-540.

149.   Kerr, I.M. and R.E. Brown, *pppA2'p5'A2'p5'A: an inhibitor of protein synthesis synthesized with an enzyme fraction from interferon-treated cells.* Proceedings of the National Academy of Sciences, 1978. 75(1): p. 256-260.

150.   Clemens, M.J. and B.R.G. Williams, *Inhibition of cell-free protein synthesis by pppA2' p5' A2' p5' A: a novel oligonucleotide synthesized by interferon-treated L cell extracts.* Cell, 1978. 13(3): p. 565-572.

151.   Williams, B.R.G. and I.M. Kerr, *Inhibition of protein synthesis by 2[prime]-5[prime] linked adenine oligonucleotides in intact cells.* Nature, 1978. 276(5683): p. 88-90.

152.   Baglioni, C., M.A. Minks, and P.A. Maroney, *Interferon action may be mediated by activation of a nuclease by pppA2[prime]p5[prime]A2[prime]p5[prime]A.* Nature, 1978. 273(5664): p. 684-687.

153.   Stark, G.R., et al., *2-5A synthetase: assay, distribution and variation with growth or hormone status.* Nature, 1979. 278(5703): p. 471-473.

154.   Coccia, E.M., et al., *A full-length murine 2-5A synthetase cDNA transfected in NIH-3T3 cells impairs EMCV but not VSV replication.* Virology, 1990. 179(1): p. 228-233.

155.   Chebath, J., et al., *Constitutive expression of (2[prime]-5[prime]) oligo A synthetase confers resistance to picornavirus infection.* Nature, 1987. 330(6148): p. 587-588.

156.   Li, Y., et al., *Activation of RNase L is dependent on OAS3 expression during infection with diverse human viruses.* Proceedings of the National Academy of Sciences, 2016. 113(8): p. 2241-2246.

157.   Han, J.-Q. and D.J. Barton, *Activation and evasion of the antiviral 2'-5' oligoadenylate synthetase/ribonuclease L pathway by hepatitis C virus mRNA.* RNA, 2002. 8(4): p. 512-525.

158.   Schröder, H.C., et al., *(2'-5')Oligoadenylate and intracellular immunity against retrovirus infection.* International Journal of Biochemistry, 1992. 24(1): p. 55-63.

159.   Cayley, P.J., et al., *Distribution of the ppp(A2'p)nA-binding protein and interferon-related enzymes in animals, plants, and lower organisms.* Biochemical and Biophysical Research Communications, 1982. 108(3): p. 1243-1250.

160.   Torralba, S., J. Sojat, and R. Hartmann, *2'-5' Oligoadenylate synthetase shares active site architecture with the archaeal CCA-adding enzyme.* Cellular and Molecular Life Sciences, 2008. 65(16): p. 2613-2620.

161.   Wiens, M., et al., *Origin of the interferon-inducible (2'-5')oligoadenylate synthetases: cloning of the (2'-5')oligoadenylate synthetase from the marine sponge Geodia cydonium 1.* FEBS Letters, 1999. 462(1-2): p. 12-18.

162.   Chebath, J., et al., *Four different forms of interferon-induced 2',5'-oligo(A) synthetase identified by immunoblotting in human cells.* Journal of Biological Chemistry, 1987. 262(8): p. 3852-7.

163.   Hovanessian, A.G., et al., *Identification of 69-kd and 100-kd forms of 2-5A synthetase in interferon-treated human cells by specific monoclonal antibodies.* The EMBO Journal, 1987. 6(5): p. 1273-1280.

164.   Rebouillat, D., I. Marié, and A.G. Hovanessian, *Molecular cloning and characterization of two related and interferon-induced 56-kDa and 30-kDa proteins highly similar to 2'-5' oligoadenylate synthetase.* European Journal of Biochemistry, 1998. 257(2): p. 319-330.

165.   Hartmann, R., et al., *p59OASL, a 2'-5' oligoadenylate synthetase like protein: a novel human gene related to the 2'-5' oligoadenylate synthetase family.* Nucleic Acids Research, 1998. 26(18): p. 4121-4128.

166.   Justesen, J., R. Hartmann, and N.O. Kjeldgaard, *Gene structure and function of the 2'-5'-oligoadenylate synthetase family.* Cellular and Molecular Life Sciences CMLS, 2000. 57(11): p. 1593-1612.

167.   Benech, P., et al., *Structure of two forms of the interferon-induced (2'-5') oligo A synthetase of human cells based on cDNAs and gene sequences.* The EMBO Journal, 1985. 4(9): p. 2249-2256.

168.   Marié, I. and A.G. Hovanessian, *The 69-kDa 2-5A synthetase is composed of two homologous and adjacent functional domains.* Journal of Biological Chemistry, 1992. 267(14): p. 9933-9939.

169.   Hovnanian, A., et al., *The Human 2',5'-Oligoadenylate Synthetase Locus Is Composed of Three Distinct Genes Clustered on Chromosome 12q24.2 Encoding the 100-, 69-, and 40-kDa Forms.* Genomics, 1998. 52(3): p. 267-277.

170.   Marié, I., et al., *Differential expression and distinct structure of 69- and 100-kDa forms of 2-5A synthetase in human cells treated with interferon.* Journal of Biological Chemistry, 1990. 265(30): p. 18601-18607.

171.   Besse, S., et al., *Ultrastructural Localization of Interferon-Inducible Double-Stranded RNA-Activated Enzymes in Human Cells.* Experimental Cell Research, 1998. 239(2): p. 379-392.

172.   Shiojiri, S., et al., *Structure and Expression of a Cloned cDNA for Human (2'-5') Oligoadenylate Synthetase.* The Journal of Biochemistry, 1986. 99(5): p. 1455-1464.

173.   Wathelet, M., et al., *Full-length sequence and expression of the 42 kDa 2–5A synthetase induced by human interferon.* FEBS Letters, 1986. 196(1): p. 113-120.

174.   Kristiansen, H., et al., *Extracellular 2'-5' Oligoadenylate Synthetase Stimulates RNase L-Independent Antiviral Activity: a Novel Mechanism of Virus-Induced Innate Immunity.* Journal of Virology, 2010. 84(22): p. 11898-11904.

175. Bordignon, J., et al., *Expression profile of interferon stimulated genes in central nervous system of mice infected with dengue virus Type-1.* Virology, 2008. 377(2): p. 319-329.

176. Warke, R.V., et al., *Dengue Virus Induces Novel Changes in Gene Expression of Human Umbilical Vein Endothelial Cells.* Journal of Virology, 2003. 77(21): p. 11822-11832.

177. Alagarasu, K., et al., *Polymorphisms in the oligoadenylate synthetase gene cluster and its association with clinical outcomes of dengue virus infection.* Infection, Genetics and Evolution, 2013. 14(Supplement C): p. 390-395.

178. Donovan, J., M. Dufner, and A. Korennykh, *Structural basis for cytosolic double-stranded RNA surveillance by human oligoadenylate synthetase 1.* Proceedings of the National Academy of Sciences, 2013. 110(5): p. 1652-1657.

179. Lohöfener, J., et al., *The Activation Mechanism of 2'-5'-Oligoadenylate Synthetase Gives New Insights Into OAS/cGAS Triggers of Innate Immunity.* Structure, 2015. 23(5): p. 851-862.

180. Desai, S.Y. and G.C. Sen, *Effects of Varying Lengths of Double-Stranded RNA on Binding and Activation of 2'-5'-Oligoadenylate Synthetase.* Journal of Interferon & Cytokine Research, 1997. 17(9): p. 531-536.

181. Maitra, R.K., et al., *HIV-1 TAR RNA Has an Intrinsic Ability to Activate Interferon-Inducible Enzymes.* Virology, 1994. 204(2): p. 823-827.

182. Steitz, T., et al., *A unified polymerase mechanism for nonhomologous DNA and RNA polymerases.* Science, 1994. 266(5193): p. 2022-2025.

183. Steitz, T.A., *DNA Polymerases: Structural Diversity and Common Mechanisms.* Journal of Biological Chemistry, 1999. 274(25): p. 17395-17398.

184. Balbo, P.B. and A. Bohm, *Mechanism of Poly(A) Polymerase: Structure of the Enzyme-MgATP-RNA Ternary Complex and Kinetic Analysis.* Structure, 2007. 15(9): p. 1117-1131.

185. Marié, I., et al., *69-kDa and 100-kDa Isoforms of Interferon-Induced (2'-5')Oligoadenylate Synthetase Exhibit Differential Catalytic Parameters.* European Journal of Biochemistry, 1997. 248(2): p. 558-566.

186. Ball, A., *Induction, purification, and properties of 2'5' oligoadenylate synthetase.* Vol. 350. 1980. 486-96.

187. Cayley, P.J. and I.M. Kerr, *Synthesis, Characterisation and Biological Significance of (2'-5')Oiigoadenylate Derivatives of NAD+, ADP-Ribose and Adenosine(5')Tetraphosph0(5')Adenosine.* European Journal of Biochemistry, 1982. 122(3): p. 601-608.

188. Justesen, J., et al., *The interferon-induced enzyme 2-5A synthetase adenylates tRNA.* Vol. 67. 1985. 651-5.

189. Ghosh, A., et al., *Enzymatic Activity of 2'-5'-Oligoadenylate Synthetase Is Impaired by Specific Mutations that Affect Oligomerization of the Protein.* Vol. 272. 1998. 33220-6.

190. Sarkar, S.N., S. Pal, and G.C. Sen, *Crisscross Enzymatic Reaction between the Two Molecules in the Active Dimeric P69 Form of the 2'-5' Oligodenylate Synthetase.* Journal of Biological Chemistry, 2002. 277(47): p. 44760-44764.

191.   Justesen, J., D. Ferbus, and M.N. Thang, *Elongation mechanism and substrate specificity of 2',5'-oligoadenylate synthetase.* Proceedings of the National Academy of Sciences of the United States of America, 1980. 77(8): p. 4618-4622.

192.   Dougherty, J.P., et al., *Interferon, double-stranded RNA, and RNA degradation. Isolation of homogeneous pppA(2'p5'A)n-1 synthetase from Ehrlich ascites tumor cells.* Journal of Biological Chemistry, 1980. 255(9): p. 3813-6.

193.   Dong, B., et al., *Intrinsic molecular activities of the interferon-induced 2-5A-dependent RNase.* Journal of Biological Chemistry, 1994. 269(19): p. 14153-14158.

194.   Rebouillat, D., et al., *The 100-kDa 2',5'-Oligoadenylate Synthetase Catalyzing Preferentially the Synthesis of Dimeric pppA2'p5'A Molecules Is Composed of Three Homologous Domains.* Journal of Biological Chemistry, 1999. 274(3): p. 1557-1565.

195.   Dong, B. and R.H. Silverman, *A Bipartite Model of 2-5A-dependent RNase L.* Journal of Biological Chemistry, 1997. 272(35): p. 22236-22242.

196.   Dong, B. and R.H. Silverman, *2-5A-dependent RNase Molecules Dimerize during Activation by 2-5A.* Journal of Biological Chemistry, 1995. 270(8): p. 4133-4137.

197.   Silverman, R.H., et al., *rRNA cleavage as an index of ppp(A2'p)nA activity in interferon-treated encephalomyocarditis virus-infected cells.* Journal of Virology, 1983. 46(3): p. 1051-1055.

198.   Li, X.-L., J.A. Blackford, and B.A. Hassel, *RNase L Mediates the Antiviral Effect of Interferon through a Selective Reduction in Viral RNA during Encephalomyocarditis Virus Infection.* Journal of Virology, 1998. 72(4): p. 2752-2759.

199.   Floyd-Smith, G., E. Slattery, and P. Lengyel, *Interferon action: RNA cleavage pattern of a (2'-5')oligoadenylate--dependent endonuclease.* Science, 1981. 212(4498): p. 1030-1032.

200.   Wreschner, D., et al., *Interferon action: Sequence specificity of the ppp(A2'p)(n)A-dependent ribonuclease.* Vol. 289. 1981. 414-7.

201.   Malathi, K., et al., *Small self-RNA generated by RNase L amplifies antiviral innate immunity.* Nature, 2007. 448(7155): p. 816-819.

202.   Nilsen, T.W. and C. Baglioni, *Mechanism for discrimination between viral and host mRNA in interferon-treated cells.* Proceedings of the National Academy of Sciences of the United States of America, 1979. 76(6): p. 2600-2604.

203.   Castelli, J.C., et al., *A Study of the Interferon Antiviral Mechanism: Apoptosis Activation by the 2–5A System.* The Journal of Experimental Medicine, 1997. 186(6): p. 967-972.

204.   Castelli, J.C., et al., *The role of 2'-5' oligoadenylate-activated ribonuclease L in apoptosis.* Cell Death Differ, 1998. 5(4): p. 313-20.

205.   Marques, J., et al., *The p59 oligoadenylate synthetase-like protein possesses antiviral activity that requires the C-terminal ubiquitin-like domain.* Journal of General Virology, 2008. 89(11): p. 2767-2772.

206.   Zhu, J., et al., *Antiviral Activity of Human OASL Protein Is Mediated by Enhancing Signaling of the RIG-I RNA Sensor.* Immunity, 2014. 40(6): p. 936-948.

207.   Lee, M.S., et al., *OASL1 inhibits translation of the type I interferon–regulating transcription factor IRF7.* 2013. 14: p. 346.

208.   Lee, M.S., et al., *Negative Regulation of Type I IFN Expression by OASL1 Permits Chronic Viral Infection and CD8+ T-Cell Exhaustion.* PLOS Pathogens, 2013. 9(7): p. e1003478.

209.   Bréhin, A.-C., et al., *The large form of human 2',5'-Oligoadenylate Synthetase (OAS3) exerts antiviral effect against Chikungunya virus.* Virology, 2009. 384(1): p. 216-222.

210.   Sorgeloos, F., et al., *Evasion of Antiviral Innate Immunity by Theiler's Virus L* Protein through Direct Inhibition of RNase L.* PLOS Pathogens, 2013. 9(6): p. e1003474.

211.   Zhao, L., et al., *Antagonism of the Interferon-Induced OAS-RNase L Pathway by Murine Coronavirus ns2 Protein Is Required for Virus Replication and Liver Pathology.* Cell host & microbe, 2012. 11(6): p. 607-616.

212.   Zhang, R., et al., *Homologous 2',5'-phosphodiesterases from disparate RNA viruses antagonize antiviral innate immunity.* Proceedings of the National Academy of Sciences of the United States of America, 2013. 110(32): p. 13114-13119.

213.   Li, X.-D., et al., *Pivotal Roles of cGAS-cGAMP Signaling in Antiviral Defense and Immune Adjuvant Effects.* Science (New York, N.Y.), 2013. 341(6152): p. 10.1126/science.1244040.

214.   Gray, E. and D.B. Stetson, *The AIM2-like receptors are dispensable for activation of the interferon stimulatory DNA (ISD) pathway.* The Journal of Immunology, 2016. 196(1 Supplement): p. 203.10-203.10.

215.   Ishikawa, H., Z. Ma, and G.N. Barber, *STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity.* Nature, 2009. 461(7265): p. 788-792.

216.   Schoggins, J.W., et al., *Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity.* Nature, 2014. 505(7485): p. 691-695.

217.   Ma, Z., et al., *Modulation of the cGAS-STING DNA sensing pathway by gammaherpesviruses.* Proceedings of the National Academy of Sciences, 2015. 112(31): p. E4306-E4315.

218.   Shravanthi, G.V., et al., *HBV Induce Cytosolic DNA Sensor cGAS but Suppress IRF3 to Prevent Generation of Antiviral Defense Mechanism in Infected Cell.* Journal of Clinical and Experimental Hepatology, 2015. 5(Supplement 2): p. S4.

219.   Yoh, Sunnie M., et al., *PQBP1 Is a Proximal Sensor of the cGAS-Dependent Innate Response to HIV-1.* Cell, 2015. 161(6): p. 1293-1305.

220.   Gao, D., et al., *Cyclic GMP-AMP Synthase Is an Innate Immune Sensor of HIV and Other Retroviruses.* Science, 2013. 341(6148): p. 903-906.

221.   Manzanillo, Paolo S., et al., *Mycobacterium Tuberculosis Activates the DNA-Dependent Cytosolic Surveillance Pathway within Macrophages.* Cell host & microbe, 2012. 11(5): p. 469-480.

222.   Hansen, K., et al., *Listeria monocytogenes induces IFNβ expression through an IFI16-, cGAS- and STING-dependent pathway.* The EMBO Journal, 2014. 33(15): p. 1654-1666.

223.   Sun, L., et al., *Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway.* Science, 2013. 339(6121): p. 786-791.

224.   Gehrke, N., et al., *Oxidative Damage of DNA Confers Resistance to Cytosolic Nuclease TREX1 Degradation and Potentiates STING-Dependent Immune Sensing.* Immunity, 2013. 39(3): p. 482-495.

225.   West, A.P., et al., *Mitochondrial DNA stress primes the antiviral innate immune response.* Nature, 2015. 520(7548): p. 553-557.

226.   Rongvaux, A., et al., *Apoptotic Caspases Prevent the Induction of Type I Interferons by Mitochondrial DNA.* Cell, 2014. 159(7): p. 1563-1577.

227.   White, Michael J., et al., *Apoptotic Caspases Suppress mtDNA-Induced STING-Mediated Type I IFN Production.* Cell, 2014. 159(7): p. 1549-1562.

228.   Mackenzie, K.J., et al., *cGAS surveillance of micronuclei links genome instability to innate immunity.* Nature, 2017. 548(7668): p. 461-465.

229.   Harding, S.M., et al., *Mitotic progression following DNA damage enables pattern recognition within micronuclei.* Nature, 2017. 548(7668): p. 466-470.

230.   Glück, S., et al., *Innate immune sensing of cytosolic chromatin fragments through cGAS promotes senescence.* Nature Cell Biology, 2017. 19: p. 1061.

231.   Dou, Z., et al., *Cytoplasmic chromatin triggers inflammation in senescence and cancer.* Nature, 2017. 550: p. 402.

232.   Stetson, D.B. and R. Medzhitov, *Recognition of Cytosolic DNA Activates an IRF3-Dependent Innate Immune Response.* Immunity, 2006. 24(1): p. 93-103.

233.   Civril, F., et al., *Structural mechanism of cytosolic DNA sensing by cGAS.* Nature, 2013. 498(7454): p. 332-337.

234.   Gao, P., et al., *Cyclic [G(2',5')pA(3',5')p] Is the Metazoan Second Messenger Produced by DNA-Activated Cyclic GMP-AMP Synthase.* Cell, 2013. 153(5): p. 1094-1107.

235.   Kranzusch, Philip J., et al., *Structure of Human cGAS Reveals a Conserved Family of Second-Messenger Enzymes in Innate Immunity.* Cell Reports, 2013. 3(5): p. 1362-1368.

236.   Kato, K., et al., *Structural and Functional Analyses of DNA-Sensing and Immune Activation by Human cGAS.* PLoS ONE, 2013. 8(10): p. e76983.

237.   Jakobsen, M.R., et al., *IFI16 senses DNA forms of the lentiviral replication cycle and controls HIV-1 replication.* Proceedings of the National Academy of Sciences, 2013. 110(48): p. E4571-E4580.

238.   Wu, J., et al., *Cyclic GMP-AMP Is an Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA.* Science, 2013. 339(6121): p. 826-830.

239.   Diner, Elie J., et al., *The Innate Immune DNA Sensor cGAS Produces a Noncanonical Cyclic Dinucleotide that Activates Human STING.* Cell Reports, 2013. 3(5): p. 1355-1361.

240.   Ablasser, A., et al., *cGAS produces a 2'-5'-linked cyclic dinucleotide second messenger that activates STING.* Nature, 2013. 498(7454): p. 380-384.

241.   Ablasser, A., et al., *Cell intrinsic immunity spreads to bystander cells via the intercellular transfer of cGAMP.* Nature, 2013. 503(7477): p. 530-534.

242.   Bridgeman, A., et al., *Viruses transfer the antiviral second messenger cGAMP between cells.* Science, 2015. 349(6253): p. 1228-1232.

243.   Gentili, M., et al., *Transmission of innate immune signaling by packaging of cGAMP in viral particles.* Science, 2015. 349(6253): p. 1232-1236.

244. Ishikawa, H. and G.N. Barber, *STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling.* Nature, 2008. 455(7213): p. 674-678.

245. Zhong, B., et al., *The Adaptor Protein MITA Links Virus-Sensing Receptors to IRF3 Transcription Factor Activation.* Immunity, 2008. 29(4): p. 538-550.

246. Sun, W., et al., *ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization.* Proceedings of the National Academy of Sciences, 2009. 106(21): p. 8653-8658.

247. Tanaka, Y. and Z.J. Chen, *STING Specifies IRF3 Phosphorylation by TBK1 in the Cytosolic DNA Signaling Pathway.* Science Signaling, 2012. 5(214): p. ra20-ra20.

248. Ouyang, S., et al., *Structural Analysis of the STING Adaptor Protein Reveals a Hydrophobic Dimer Interface and Mode of Cyclic di-GMP Binding.* Immunity, 2012. 36(6): p. 1073-1086.

249. Gao, P., et al., *Structure-Function Analysis of STING Activation by c[G(2',5')pA(3',5')p] and Targeting by Antiviral DMXAA.* Cell, 2013. 154(4): p. 748-762.

250. Zhang, X., et al., *Cyclic GMP-AMP Containing Mixed Phosphodiester Linkages Is An Endogenous High-Affinity Ligand for STING.* Molecular Cell, 2013. 51(2): p. 226-235.

251. Saitoh, T., et al., *Atg9a controls dsDNA-driven dynamic translocation of STING and the innate immune response.* Proceedings of the National Academy of Sciences of the United States of America, 2009. 106(49): p. 20842-20846.

252. Chen, H., et al., *Activation of STAT6 by STING Is Critical for Antiviral Innate Immunity.* Cell, 2011. 147(2): p. 436-446.

253. Abe, T. and G.N. Barber, *Cytosolic-DNA-Mediated, STING-Dependent Proinflammatory Gene Induction Necessitates Canonical NF-κB Activation through TBK1.* Journal of Virology, 2014. 88(10): p. 5328-5341.

254. Ma, F., et al., *Positive Feedback Regulation of Type I IFN Production by the IFN-Inducible DNA Sensor cGAS.* The Journal of Immunology, 2015. 194(4): p. 1545-1554.

255. Ma, F., et al., *Positive feedback regulation of type I interferon by the interferon-stimulated gene STING.* EMBO Reports, 2015. 16(2): p. 202-212.

256. Jin, L., et al., *MPYS is required for IRF3 activation and type I IFN production in the response of cultured phagocytes to bacterial second messengers c-di-AMP and c-di-GMP().* Journal of immunology (Baltimore, Md. : 1950), 2011. 187(5): p. 2595-2601.

257. Sauer, J.-D., et al., *The N-Ethyl-N-Nitrosourea-Induced Goldenticket Mouse Mutant Reveals an Essential Function of Sting in the In Vivo Interferon Response to Listeria monocytogenes and Cyclic Dinucleotides.* Infection and Immunity, 2011. 79(2): p. 688-694.

258. Woodward, J.J., A.T. Iavarone, and D.A. Portnoy, *c-di-AMP secreted by intracellular Listeria monocytogenes activates a host type I interferon response.* Science, 2010. 328(5986): p. 1703-1705.

259. Dey, B., et al., *A bacterial cyclic dinucleotide activates the cytosolic surveillance pathway and mediates innate resistance to tuberculosis.* 2015. 21: p. 401.

260. Barker, J.R., et al., *STING-Dependent Recognition of Cyclic di-AMP Mediates Type I Interferon Responses during Chlamydia trachomatis Infection.* mBio, 2013. 4(3): p. e00018-13.

261.   Stetson, D.B., et al., *Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity.* Cell, 2008. 134(4): p. 587-598.

262.   Kawane, K., et al., *Chronic polyarthritis caused by mammalian DNA that escapes from degradation in macrophages.* Nature, 2006. 443(7114): p. 998-1002.

263.   Napirei, M., et al., *Features of systemic lupus erythematosus in Dnase1-deficient mice.* Nature Genetics, 2000. 25: p. 177.

264.   Yoshida, H., et al., *Lethal anemia caused by interferon-β produced in mouse embryos carrying undigested DNA.* Nature immunology, 2004. 6: p. 49.

265.   Takaoka, A., et al., *DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response.* Nature, 2007. 448(7152): p. 501-505.

266.   DeFilippis, V.R., et al., *Human Cytomegalovirus Induces the Interferon Response via the DNA Sensor ZBP1.* Journal of Virology, 2010. 84(1): p. 585-598.

267.   Ishii, K.J., et al., *TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines.* Nature, 2008. 451(7179): p. 725-729.

268.   Upton, Jason W., William J. Kaiser, and Edward S. Mocarski, *DAI/ZBP1/DLM-1 Complexes with RIP3 to Mediate Virus-Induced Programmed Necrosis that Is Targeted by Murine Cytomegalovirus vIRA.* Cell host & microbe, 2012. 11(3): p. 290-297.

269.   Lieber, M.R., et al., *Mechanism and regulation of human non-homologous DNA end-joining.* Nat Rev Mol Cell Biol, 2003. 4(9): p. 712-720.

270.   Ferguson, B.J., et al., *DNA-PK is a DNA sensor for IRF-3-dependent innate immunity.* eLife, 2012. 1: p. e00047.

271.   Zhang, X., et al., *Cutting Edge: Ku70 Is a Novel Cytosolic DNA Sensor That Induces Type III Rather Than Type I IFN.* The Journal of Immunology, 2011. 186(8): p. 4541-4545.

272.   Sui, H., et al., *STING is an essential mediator of the Ku70-mediated production of IFN-λ1 in response to exogenous DNA.* Science Signaling, 2017. 10(488).

273.   Stracker, T.H. and J.H.J. Petrini, *The MRE11 complex: starting from the ends.* Nat Rev Mol Cell Biol, 2011. 12(2): p. 90-103.

274.   Kondo, T., et al., *DNA damage sensor MRE11 recognizes cytosolic double-stranded DNA and induces type I interferon by regulating STING trafficking.* Proceedings of the National Academy of Sciences, 2013. 110(8): p. 2969-2974.

275.   Mariggiò, G., et al., *Kaposi Sarcoma Herpesvirus (KSHV) Latency-Associated Nuclear Antigen (LANA) recruits components of the MRN (Mre11-Rad50-NBS1) repair complex to modulate an innate immune signaling pathway and viral latency.* PLOS Pathogens, 2017. 13(4): p. e1006335.

276.   Roth, S., et al., *Rad50-CARD9 interactions link cytosolic DNA sensing to IL-1[beta] production.* Nat Immunol, 2014. 15(6): p. 538-545.

277.   Stein, S.C. and E. Falck-Pedersen, *Sensing Adenovirus Infection: Activation of Interferon Regulatory Factor 3 in RAW 264.7 Cells.* Journal of Virology, 2012. 86(8): p. 4527-4537.

278.   Zhang, Z., et al., *The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells.* Nat Immunol, 2011. 12(10): p. 959-965.

279.   Parvatiyar, K., et al., *DDX41 recognizes bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response.* Nature immunology, 2012. 13(12): p. 1155-1161.

280.   Zhang, Z., et al., *The E3 ubiquitin ligase TRIM21 negatively regulates the innate immune response to intracellular double-stranded DNA.* Nat Immunol, 2013. 14(2): p. 172-178.

281.   Abe, T., et al., *STING Recognition of Cytoplasmic DNA Instigates Cellular Defense.* Molecular Cell, 2013. 50(1): p. 5-15.

282.   Kim, T., et al., *Aspartate-glutamate-alanine-histidine box motif (DEAH)/RNA helicase A helicases sense microbial DNA in human plasmacytoid dendritic cells.* Proceedings of the National Academy of Sciences, 2010. 107(34): p. 15181-15186.

283.   Zhang, Z., et al., *DHX9 Pairs with IPS-1 To Sense Double-Stranded RNA in Myeloid Dendritic Cells.* The Journal of Immunology, 2011. 187(9): p. 4501-4508.

284.   Zhang, Z., et al., *DDX1, DDX21, and DHX36 Helicases Form a Complex with the Adaptor Molecule TRIF to Sense dsRNA in Dendritic Cells.* Immunity, 2011. 34(6): p. 866-878.

285.   Sarkar, A. and K. Hochedlinger, *The Sox Family of Transcription Factors: Versatile Regulators of Stem and Progenitor Cell Fate.* Cell Stem Cell, 2013. 12(1): p. 15-30.

286.   Arnold, K., et al., *Sox2+ Adult Stem and Progenitor Cells Are Important for Tissue Regeneration and Survival of Mice.* Cell Stem Cell, 2011. 9(4): p. 317-329.

287.   Xia, P., et al., *Sox2 functions as a sequence-specific DNA sensor in neutrophils to initiate innate immunity against microbial infection.* Nat Immunol, 2015. 16(4): p. 366-375.

288.   Bonaldi, T., et al., *The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding.* The EMBO Journal, 2002. 21(24): p. 6865-6873.

289.   Ura, K., K. Nightingale, and A.P. Wolffe, *Differential association of HMG1 and linker histones B4 and H1 with dinucleosomal DNA: structural transitions and transcriptional repression.* The EMBO Journal, 1996. 15(18): p. 4959-4969.

290.   Calogero, S., et al., *The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice.* Nat Genet, 1999. 22(3): p. 276-280.

291.   Harris, H.E., U. Andersson, and D.S. Pisetsky, *HMGB1: A multifunctional alarmin driving autoimmune and inflammatory disease.* 2012. 8: p. 195.

292.   Scaffidi, P., T. Misteli, and M.E. Bianchi, *Release of chromatin protein HMGB1 by necrotic cells triggers inflammation.* Nature, 2002. 418(6894): p. 191-195.

293.   Hreggvidsdóttir, H.S., et al., *High Mobility Group Box Protein 1 (HMGB1)-Partner Molecule Complexes Enhance Cytokine Production by Signaling Through the Partner Molecule Receptor.* Molecular Medicine, 2012. 18(1): p. 224-230.

294.   Campana, L., et al., *Requirement of HMGB1 for stromal cell-derived factor-1/CXCL12-dependent migration of macrophages and dendritic cells.* Journal of Leukocyte Biology, 2009. 86(3): p. 609-615.

295.   Park, J.S., et al., *High mobility group box 1 protein interacts with multiple Toll-like receptors.* American Journal of Physiology - Cell Physiology, 2006. 290(3): p. C917-C924.

296.   Yun, J., et al., *The HMGB1–CXCL12 Complex Promotes Inflammatory Cell Infiltration in Uveitogenic T Cell-Induced Chronic Experimental Autoimmune Uveitis.* Frontiers in Immunology, 2017. 8: p. 142.

297.   Andersson, U., et al., *High Mobility Group 1 Protein (Hmg-1) Stimulates Proinflammatory Cytokine Synthesis in Human Monocytes.* The Journal of Experimental Medicine, 2000. 192(4): p. 565-570.

298.   Yanai, H., et al., *HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses.* Nature, 2009. 462(7269): p. 99-103.

299.   Schattgen, S.A. and K.A. Fitzgerald, *The PYHIN protein family as mediators of host defenses.* Immunological Reviews, 2011. 243(1): p. 109-118.

300.   Stratmann, S.A., et al., *The innate immune sensor IFI16 recognizes foreign DNA in the nucleus by scanning along the duplex.* eLife, 2015. 4: p. e11721.

301.   Ansari, M.A., et al., *Herpesvirus Genome Recognition Induced Acetylation of Nuclear IFI16 Is Essential for Its Cytoplasmic Translocation, Inflammasome and IFN-β Responses.* PLOS Pathogens, 2015. 11(7): p. e1005019.

302.   Conrady, C.D., et al., *Resistance to HSV-1 infection in the epithelium resides with the novel innate sensor, IFI-16.* Mucosal Immunol, 2012. 5(2): p. 173-183.

303.   Diner, B.A., et al., *Viral DNA Sensors IFI16 and Cyclic GMP-AMP Synthase Possess Distinct Functions in Regulating Viral Gene Expression, Immune Defenses, and Apoptotic Responses during Herpesvirus Infection.* mBio, 2016. 7(6).

304.   Orzalli, M.H., et al., *cGAS-mediated stabilization of IFI16 promotes innate signaling during herpes simplex virus infection.* Proceedings of the National Academy of Sciences, 2015. 112(14): p. E1773-E1781.

305.   Orzalli, M.H., N.A. DeLuca, and D.M. Knipe, *Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein.* Proceedings of the National Academy of Sciences, 2012. 109(44): p. E3008–E3017.

306.   Almine, J.F., et al., *IFI16 and cGAS cooperate in the activation of STING during DNA sensing in human keratinocytes.* 2017. 8: p. 14392.

307.   Jønsson, K.L., et al., *IFI16 is required for DNA sensing in human macrophages by promoting production and function of cGAMP.* 2017. 8: p. 14391.

308.   Aravind, L. and E.V. Koonin, *DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history.* Nucleic Acids Research, 1999. 27(7): p. 1609-1618.

309.   Holm, L. and C. Sander, *DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily.* Trends in Biochemical Sciences, 1995. 20(9): p. 345-347.

310.   Kuchta, K., et al., *Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human.* Nucleic Acids Research, 2009. 37(22): p. 7701-7714.

311.   Chow, K.L., D.H. Hall, and S.W. Emmons, *The mab-21 gene of Caenorhabditis elegans encodes a novel protein required for choice of alternate cell fates.* Development, 1995. 121(11): p. 3615-3626.

312.   Hall, J., et al., *The catalytic mechanism of cyclic GMP-AMP synthase (cGAS) and implications for innate immunity and inhibition.* Protein Science, 2017: p. n/a-n/a.

313.   Zhang, X., et al., *The Cytosolic DNA Sensor cGAS Forms an Oligomeric Complex with DNA and Undergoes Switch-like Conformational Changes in the Activation Loop.* Cell Reports, 2014. 6(3): p. 421-430.

314.   Li, X., et al., *Cyclic GMP-AMP Synthase Is Activated by Double-Stranded DNA-Induced Oligomerization.* Immunity, 2013. 39(6): p. 1019-1031.

315.   Tao, J., et al., *Nonspecific DNA Binding of cGAS N Terminus Promotes cGAS Activation.* The Journal of Immunology, 2017. 198(9): p. 3627-3636.

316.   Hartmann, R., et al., *Crystal Structure of the 2'-Specific and Double-Stranded RNA-Activated Interferon-Induced Antiviral Protein 2'-5'-Oligoadenylate Synthetase.* Molecular Cell, 2003. 12(5): p. 1173-1185.

317.   Danilchanka, O. and John J. Mekalanos, *Cyclic Dinucleotides and the Innate Immune Response.* Cell, 2013. 154(5): p. 962-970.

318.   Chan, C., et al., *Structural basis of activity and allosteric control of diguanylate cyclase.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(49): p. 17084-17089.

319.   De, N., et al., *Phosphorylation-Independent Regulation of the Diguanylate Cyclase WspR.* PLOS Biology, 2008. 6(3): p. e67.

320.   Witte, G., et al., *Structural Biochemistry of a Bacterial Checkpoint Protein Reveals Diadenylate Cyclase Activity Regulated by DNA Recombination Intermediates.* Molecular Cell, 2008. 30(2): p. 167-178.

321.   Davies, Bryan W., et al., *Coordinated Regulation of Accessory Genetic Elements Produces Cyclic Di-Nucleotides for V. cholerae Virulence.* Cell, 2012. 149(2): p. 358-370.

322.   Kranzusch, Philip J., et al., *Structure-Guided Reprogramming of Human cGAS Dinucleotide Linkage Specificity.* Cell, 2014. 158(5): p. 1011-1021.

323.   Zhu, D., et al., *Structural Biochemistry of a Vibrio cholerae Dinucleotide Cyclase Reveals Cyclase Activity Regulation by Folates.* Molecular Cell, 2014. 55(6): p. 931-937.

324.   Kato, K., et al., *Structural Basis for the Catalytic Mechanism of DncV, Bacterial Homolog of Cyclic GMP-AMP Synthase.* Structure, 2015. 23(5): p. 843-850.

325.   Jin, L., et al., *MPYS, a Novel Membrane Tetraspanner, Is Associated with Major Histocompatibility Complex Class II and Mediates Transduction of Apoptotic Signals.* Molecular and Cellular Biology, 2008. 28(16): p. 5014-5026.

326.   Burdette, D.L., et al., *STING is a direct innate immune sensor of cyclic di-GMP.* Nature, 2011. 478(7370): p. 515-518.

327.   Surpris, G., et al., *Novel Tmem173 allele reveals importance of STING N-terminus in trafficking and type I IFN production.* Journal of immunology (Baltimore, Md. : 1950), 2016. 196(2): p. 547-552.

328.   Huang, Y.-H., et al., *The structural basis for the sensing and binding of cyclic di-GMP by STING.* Nature Structural &Amp; Molecular Biology, 2012. 19: p. 728.

329.   Shu, C., et al., *Structure of STING bound to cyclic di-GMP reveals the mechanism of cyclic dinucleotide recognition by the immune system.* Nature Structural &Amp; Molecular Biology, 2012. 19: p. 722.

330.   Shang, G., et al., *Crystal structures of STING protein reveal basis for recognition of cyclic di-GMP.* Nature Structural &Amp; Molecular Biology, 2012. 19: p. 725.

331.   Yin, Q., et al., *Cyclic di-GMP Sensing via the Innate Immune Signaling Protein STING.* Molecular Cell, 2012. 46(6): p. 735-745.

332.   Shi, H., et al., *Molecular basis for the specific recognition of the metazoan cyclic GMP-AMP by the innate immune adaptor protein STING.* Proceedings of the National Academy of Sciences, 2015. 112(29): p. 8947-8952.

333.   Cavlar, T., et al., *Species-specific detection of the antiviral small-molecule compound CMA by STING.* The EMBO Journal, 2013. 32(10): p. 1440-1450.

334.   Yi, G., et al., *Single Nucleotide Polymorphisms of Human STING Can Affect Innate Immune Response to Cyclic Dinucleotides.* PLoS ONE, 2013. 8(10): p. e77846.

335.   Chin, K.-H., et al., *Novel c-di-GMP recognition modes of the mouse innate immune adaptor protein STING.* Acta Crystallographica Section D, 2013. 69(3): p. 352-366.

336.   Chin, K.H., et al., *Structural Insights into the Distinct Binding Mode of Cyclic Di-AMP with SaCpaA_RCK.* Biochemistry, 2015. 54(31): p. 4936-51.

337.   Conlon, J., et al., *Mouse, but not Human STING, Binds and Signals in Response to the Vascular Disrupting Agent 5,6-Dimethylxanthenone-4-Acetic Acid.* The Journal of Immunology, 2013. 190(10): p. 5216-5225.

338.   Zhao, B., et al., *Structural basis for concerted recruitment and activation of IRF-3 by innate immune adaptor proteins.* Proceedings of the National Academy of Sciences, 2016. 113(24): p. E3403-E3412.

339.   Qin, B.Y., et al., *Crystal structure of IRF-3 reveals mechanism of autoinhibition and virus-induced phosphoactivation.* Nature Structural Biology, 2003. 10: p. 913.

340.   Arnold, M.M., M. Barro, and J.T. Patton, *Rotavirus NSP1 Mediates Degradation of Interferon Regulatory Factors through Targeting of the Dimerization Domain.* Journal of Virology, 2013. 87(17): p. 9813-9821.

341.   Barro, M. and J.T. Patton, *Rotavirus nonstructural protein 1 subverts innate immune response by inducing degradation of IFN regulatory factor 3.* Proceedings of the National Academy of Sciences of the United States of America, 2005. 102(11): p. 4114-4119.

342.   Tsuchiya, Y., et al., *Ligand-induced Ordering of the C-terminal Tail Primes STING for Phosphorylation by TBK1.* EBioMedicine, 2016. 9: p. 87-96.

343.   Wu, X., et al., *Molecular evolutionary and structural analysis of the cytosolic DNA sensor cGAS and STING.* Nucleic Acids Research, 2014. 42(13): p. 8243-8257.

344.   Schaap, P., *Cyclic di-nucleotide signaling enters the eukaryote domain.* IUBMB Life, 2013. 65(11): p. 897-903.

345.   Kranzusch, Philip J., et al., *Ancient Origin of cGAS-STING Reveals Mechanism of Universal 2',3' cGAMP Signaling.* Molecular Cell, 2015. 59(6): p. 891-903.

346.   Margolis, S.R., S.C. Wilson, and R.E. Vance, *Evolutionary Origins of cGAS-STING Signaling.* Trends in Immunology, 2017. 38(10): p. 733-743.

347.   Ge, R., et al., *Conservation of the STING-Mediated Cytosolic DNA Sensing Pathway in Zebrafish.* Journal of Virology, 2015. 89(15): p. 7696-7706.

348.   de Oliveira Mann, C.C., et al., *Structural and biochemical characterization of the cell fate determining nucleotidyltransferase fold protein MAB21L1.* Scientific Reports, 2016. 6: p. 27498.

349.   Huang, Z.X., et al., *The Male Abnormal Gene Family 21 (Mab21) Members Regulate Eye Development.* Current Molecular Medicine, 2016. 16(7): p. 660-667.

350.   Wong, R.L.Y. and K.L. Chow, *Depletion of Mab21l1 and Mab21l2 messages in mouse embryo arrests axial turning, and impairs notochord and neural tube differentiation.* Teratology, 2002. 65(2): p. 70-77.

351.   Mozzi, A., et al., *OASes and STING: Adaptive Evolution in Concert.* Genome Biology and Evolution, 2015. 7(4): p. 1016-1032.

352.   Hancks, D.C., et al., *Overlapping Patterns of Rapid Evolution in the Nucleic Acid Sensors cGAS and OAS1 Suggest a Common Mechanism of Pathogen Antagonism and Escape.* PLOS Genetics, 2015. 11(5): p. e1005203.

353.   Kao, C.J., et al., *CBAP interacts with the un-liganded common β-subunit of the GM-CSF/IL-3/IL-5 receptor and induces apoptosis via mitochondrial dysfunction.* Oncogene, 2007. 27: p. 1397.

354.   Chiang, Y.-J., et al., *CBAP Functions as a Novel Component in Chemokine-Induced ZAP70-Mediated T-Cell Adhesion and Migration.* PLoS ONE, 2013. 8(4): p. e61761.

355.   Ho, K.C., et al., *CBAP promotes thymocyte negative selection by facilitating T-cell receptor proximal signaling.* Cell Death & Disease, 2014. 5(11): p. e1518.

356.   van Rossum, D.B., et al., *DANGER, a Novel Regulatory Protein of Inositol 1,4,5-Trisphosphate-Receptor Activity.* Journal of Biological Chemistry, 2006. 281(48): p. 37111-37116.

357.   Kang, B.N., et al., *Death-Associated Protein Kinase-Mediated Cell Death Modulated by Interaction with DANGER.* The Journal of Neuroscience, 2010. 30(1): p. 93-98.

358.   Kwon, T., et al., *DANGER is involved in high glucose-induced radioresistance through inhibiting DAPK-mediated anoikis in non-small cell lung cancer.* Oncotarget, 2016. 7(6): p. 7193-7206.

359.   Karagoz, K., et al., *Proteomic and Metabolic Signatures of Esophageal Squamous Cell Carcinoma.* Curr Cancer Drug Targets, 2016. 2: p. 2.

360.   Simpson, J.C., et al., *Systematic subcellular localization of novel proteins identified by large-scale cDNA sequencing.* EMBO Reports, 2000. 1(3): p. 287-292.

361.   Palmer, C.S., et al., *MiD49 and MiD51, new components of the mitochondrial fission machinery.* EMBO Reports, 2011. 12(6): p. 565-573.

362.   Losón, O.C., et al., *Fis1, Mff, MiD49, and MiD51 mediate Drp1 recruitment in mitochondrial fission.* Molecular Biology of the Cell, 2013. 24(5): p. 659-667.

363.   Osellame, L.D., et al., *Cooperative and independent roles of the Drp1 adaptors Mff, MiD49 and MiD51 in mitochondrial fission.* Journal of Cell Science, 2016. 129(11): p. 2170-2181.

364.   Palmer, C.S., et al., *Adaptor Proteins MiD49 and MiD51 Can Act Independently of Mff and Fis1 in Drp1 Recruitment and Are Specific for Mitochondrial Fission.* Journal of Biological Chemistry, 2013. 288(38): p. 27584-27593.

365.   Liu, T., et al., *The mitochondrial elongation factors MIEF1 and MIEF2 exert partially distinct functions in mitochondrial dynamics.* Experimental Cell Research, 2013. 319(18): p. 2893-2904.

366.   Zhao, J., et al., *Human MIEF1 recruits Drp1 to mitochondrial outer membranes and promotes mitochondrial fusion rather than fission.* The EMBO Journal, 2011. 30(14): p. 2762-2778.

367.   Otera, H., et al., *Drp1-dependent mitochondrial fission via MiD49/51 is essential for apoptotic cristae remodeling.* The Journal of Cell Biology, 2016. 212(5): p. 531-544.

368.   Xu, S., et al., *Mitochondrial E3 ubiquitin ligase MARCH5 controls mitochondrial fission and cell sensitivity to stress-induced apoptosis through regulation of MiD49 protein.* Molecular Biology of the Cell, 2016. 27(2): p. 349-359.

369.   Cherok, E., et al., *Novel regulatory roles of Mff and Drp1 in E3 ubiquitin ligase MARCH5–dependent degradation of MiD49 and Mcl1 and control of mitochondrial dynamics.* Molecular Biology of the Cell, 2017. 28(3): p. 396-410.

370.   Baird, S.E., et al., *Pattern formation in the nematode epidermis: determination of the arrangement of peripheral sense organs in the C. elegans male tail.* Development, 1991. 113(2): p. 515-526.

371.   Morita, K., K.L. Chow, and N. Ueno, *Regulation of body length and male tail ray pattern formation of Caenorhabditis elegans by a member of TGF-beta family.* Development, 1999. 126(6): p. 1337-1347.

372.   Savage, C., et al., *Caenorhabditis elegans genes sma-2, sma-3, and sma-4 define a conserved family of transforming growth factor beta pathway components.* Proceedings of the National Academy of Sciences of the United States of America, 1996. 93(2): p. 790-794.

373.   Heldin, C.-H., K. Miyazono, and P. ten Dijke, *TGF-β signalling from cell membrane to nucleus through SMAD proteins.* Nature, 1997. 390: p. 465.

374.   Choy, S.W., et al., *C. elegans SIN-3 and its associated HDAC corepressor complex act as mediators of male sensory ray development.* Biochemical and Biophysical Research Communications, 2007. 358(3): p. 802-807.

375.   Lau, G.T.C., et al., *Embryonic XMab21l2 Expression Is Required for Gastrulation and Subsequent Neural Development.* Biochemical and Biophysical Research Communications, 2001. 280(5): p. 1378-1384.

376.   Mariani, M., et al., *Two Murine and Human Homologs of Mab-21, a Cell Fate Determination Gene Involved in Caenorhabditis Elegans Neural Development.* Human Molecular Genetics, 1999. 8(13): p. 2397-2406.

377.   Wong, R.L.Y., H.T. Wong, and K. Chow, *Genomic cloning and chromosomal localization of the mouse Mab21I2 locus.* Vol. 86. 1999. 21-4.

378.   Yamada, R., et al., *Cell-autonomous involvement of Mab21l1 is essential for lens placode development.* Development, 2003. 130(9): p. 1759-1770.

379.   Yamada, R., et al., *Requirement for Mab21l2 during development of murine retina and ventral body wall.* Developmental Biology, 2004. 274(2): p. 295-307.

380.   Wong, Y.-M. and K.L. Chow, *Expression of zebrafish mab21 genes marks the differentiating eye, midbrain and neural tube.* Mechanisms of Development, 2002. 113(2): p. 149-152.

381.   Mariani, M., et al., *Mab21, the mouse homolog of a C. elegans cell-fate specification gene, participates in cerebellar, midbrain and eye development.* Mech Dev, 1998. 79(1-2): p. 131-5.

382.   Wong, R.L.Y., K.K.L. Chan, and K.L. Chow, *Developmental expression of Mab21l2 during mouse embryogenesis.* Mechanisms of Development, 1999. 87(1–2): p. 185-188.

383.   Cederlund, M.L., et al., *Mab21l2 transgenics reveal novel expression patterns of mab21l1 and mab21l2, and conserved promoter regulation without sequence conservation.* Developmental Dynamics, 2011. 240(4): p. 745-754.

384.   Margolis, R.L., et al., *cDNA Cloning of a Human Homologue of the Caenorhabditis Elegans Cell Fate-Determining Gene Mab -21: Expression, Chromosomal Localization and Analysis of a Highly Polymorphic (CAG) n Trinucleotide Repeat.* Human Molecular Genetics, 1996. 5(5): p. 607-616.

385.   Deml, B., et al., *Mutations in MAB21L2 Result in Ocular Coloboma, Microcornea and Cataracts.* PLOS Genetics, 2015. 11(2): p. e1005002.

386.   Horn, D., et al., *A Novel Oculo-Skeletal syndrome with intellectual disability caused by a particular MAB21L2 mutation.* European Journal of Medical Genetics, 2015. 58(8): p. 387-391.

387.   Rainger, J., et al., *Monoallelic and Biallelic Mutations in MAB21L2 Cause a Spectrum of Major Eye Malformations.* The American Journal of Human Genetics, 2014. 94(6): p. 915-923.

388.   Sridharan, J., et al., *Xmab21l3 mediates dorsoventral patterning in Xenopus laevis.* Mechanisms of Development, 2012. 129(5-8): p. 136-146.

389.   Takahashi, C., et al., *Mab21-l3 regulates cell fate specification of multiciliate cells and ionocytes.* Nature Communications, 2015. 6: p. 6017.

390.   Losón, Oliver C., et al., *The Mitochondrial Fission Receptor MiD51 Requires ADP as a Cofactor.* Structure, 2014. 22(3): p. 367-377.

391.   Richter, V., et al., *Structural and functional analysis of MiD51, a dynamin receptor required for mitochondrial fission.* The Journal of Cell Biology, 2014. 204(4): p. 477-486.

392.   Losón, O.C., et al., *Crystal structure and functional analysis of MiD49, a receptor for the mitochondrial fission protein Drp1.* Protein Science, 2015. 24(3): p. 386-394.

393.   Manning, B.D. and L.C. Cantley, *AKT/PKB Signaling: Navigating Downstream.* Cell, 2007. 129(7): p. 1261-1274.

394.   Seo, Gil J., et al., *Akt Kinase-Mediated Checkpoint of cGAS DNA Sensing Pathway.* Cell Reports, 2015. 13(2): p. 440-449.

395.   Kaur, S., et al., *Role of the Akt pathway in mRNA translation of interferon-stimulated genes.* Proceedings of the National Academy of Sciences, 2008. 105(12): p. 4808-4813.

396.   Ou, Y.-H., et al., *TBK1 Directly Engages Akt/PKB Survival Signaling to Support Oncogenic Transformation.* Molecular Cell, 2011. 41(4): p. 458-470.

397.   Xia, P., et al., *Glutamylation of the DNA sensor cGAS regulates its binding and synthase activity in antiviral immunity.* Nature immunology, 2016. 17: p. 369.

398.   Wang, Q., et al., *The E3 ubiquitin ligase RNF185 facilitates the cGAS-mediated innate immune response.* PLOS Pathogens, 2017. 13(3): p. e1006264.

399.   Chen, M., et al., *TRIM14 Inhibits cGAS Degradation Mediated by Selective Autophagy Receptor p62 to Promote Innate Immune Responses.* Molecular Cell, 2016. 64(1): p. 105-119.

400.   Hu, M.-M., et al., *Sumoylation Promotes the Stability of the DNA Sensor cGAS and the Adaptor STING to Regulate the Kinetics of Response to DNA Virus.* Immunity, 2016. 45(3): p. 555-569.

401.   Cui, Y., et al., *SENP7 Potentiates cGAS Activation by Relieving SUMO-Mediated Inhibition of Cytosolic DNA Sensing.* PLOS Pathogens, 2017. 13(1): p. e1006156.

402.   Li, Z., et al., *PPM1A Regulates Antiviral Signaling by Antagonizing TBK1-Mediated STING Phosphorylation and Aggregation.* PLOS Pathogens, 2015. 11(3): p. e1004783.

403.   Konno, H., K. Konno, and Glen N. Barber, *Cyclic Dinucleotides Trigger ULK1 (ATG1) Phosphorylation of STING to Prevent Sustained Innate Immune Signaling.* Cell, 2013. 155(3): p. 688-698.

404.   Tsuchida, T., et al., *The Ubiquitin Ligase TRIM56 Regulates Innate Immune Responses to Intracellular Double-Stranded DNA.* Immunity, 2010. 33(5): p. 765-776.

405.   Zhang, J., et al., *TRIM32 Protein Modulates Type I Interferon Induction and Cellular Antiviral Response by Targeting MITA/STING Protein for K63-linked Ubiquitination.* Journal of Biological Chemistry, 2012. 287(34): p. 28646-28655.

406.   Wang, Q., et al., *The E3 Ubiquitin Ligase AMFR and INSIG1 Bridge the Activation of TBK1 Kinase by Modifying the Adaptor STING.* Immunity, 2014. 41(6): p. 919-933.

407.   Zhong, B., et al., *The Ubiquitin Ligase RNF5 Regulates Antiviral Responses by Mediating Degradation of the Adaptor Protein MITA.* Immunity, 2009. 30(3): p. 397-407.

408.   Qin, Y., et al., *RNF26 Temporally Regulates Virus-Triggered Type I Interferon Induction by Two Distinct Mechanisms.* PLOS Pathogens, 2014. 10(9): p. e1004358.

409.   Luo, W.-W., et al., *iRhom2 is essential for innate immunity to DNA viruses by mediating trafficking and stability of the adaptor STING.* Nature immunology, 2016. 17: p. 1057.

410.   Mukai, K., et al., *Activation of STING requires palmitoylation at the Golgi.* Nature Communications, 2016. 7: p. 11932.

411.   Yarbrough, M.L., et al., *Primate-specific miR-576-3p sets host defense signalling threshold.* Nature Communications, 2014. 5: p. 4963.

412.   Wu, M.-Z., et al., *miR-25/93 mediates hypoxia-induced immunosuppression by repressing cGAS.* Nature Cell Biology, 2017. 19: p. 1286.

413.   Škrnjug, I., C.A. Guzmán, and C. Ruecker, *Cyclic GMP-AMP Displays Mucosal Adjuvant Activity in Mice.* PLoS ONE, 2014. 9(10): p. e110150.

414.   Wang, J., P. Li, and M.X. Wu, *Natural STING Agonist as an "Ideal" Adjuvant for Cutaneous Vaccination.* Journal of Investigative Dermatology, 2016. 136(11): p. 2183-2191.

415.   Chen, Q., et al., *Carcinoma–astrocyte gap junctions promote brain metastasis by cGAMP transfer.* Nature, 2016. 533: p. 493.

416.   Mankan, A.K., et al., *Cyclic Dinucleotides in the Scope of the Mammalian Immune System*, in *Non-canonical Cyclic Nucleotides*, R. Seifert, Editor. 2017, Springer International Publishing: Cham. p. 269-289.

417.   Berg, R.K., et al., *T Cells Detect Intracellular DNA but Fail to Induce Type I IFN Responses: Implications for Restriction of HIV Replication.* PLoS ONE, 2014. 9(1): p. e84513.

418.   Xu, S., et al., *cGAS-Mediated Innate Immunity Spreads Intercellularly through HIV-1 Env-Induced Membrane Fusion Sites.* Cell host & microbe, 2016. 20(4): p. 443-457.

419.   Valadi, H., et al., *Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells.* Nature Cell Biology, 2007. 9: p. 654.

420.   Dreux, M., et al., *Short-Range Exosomal Transfer of Viral RNA from Infected Cells to Plasmacytoid Dendritic Cells Triggers Innate Immunity.* Cell host & microbe, 2012. 12(4): p. 558-570.

421.   Gao, J., et al., *Identification and characterization of phosphodiesterases that specifically degrade 3'3′-cyclic GMP-AMP.* Cell Research, 2015. 25: p. 539.

422.   Kubota, K., et al., *Identification of 2'-Phosphodiesterase, Which Plays a Role in the 2-5A System Regulated by Interferon.* Journal of Biological Chemistry, 2004. 279(36): p. 37832-37841.

423.   Li, L., et al., *Hydrolysis of 2'3'-cGAMP by ENPP1 and design of nonhydrolyzable analogs.* Nature Chemical Biology, 2014. 10: p. 1043.

424.   Kato, K., et al., *Crystal structure of Enpp1, an extracellular glycoprotein involved in bone mineralization and insulin signaling.* Proceedings of the National Academy of Sciences of the United States of America, 2012. 109(42): p. 16876-16881.

425.   Namasivayam, V., S.-Y. Lee, and C.E. Müller, *The promiscuous ectonucleotidase NPP1: molecular insights into substrate binding and hydrolysis.* Biochimica et Biophysica Acta (BBA) - General Subjects, 2017. 1861(3): p. 603-614.

426.   Gao, D., et al., *Activation of cyclic GMP-AMP synthase by self-DNA causes autoimmune diseases.* Proceedings of the National Academy of Sciences of the United States of America, 2015. 112(42): p. E5699-E5705.

427.   Tao, J., X. Zhou, and Z. Jiang, *cGAS-cGAMP-STING: The three musketeers of cytosolic DNA sensing and signaling.* IUBMB Life, 2016. 68(11): p. 858-870.

428.   Wang, F., et al., *S6K-STING interaction regulates cytosolic DNA–mediated activation of the transcription factor IRF3.* Nature immunology, 2016. 17: p. 514.

429.   Zhou, Q., et al., *The ER-Associated Protein ZDHHC1 Is a Positive Regulator of DNA Virus-Triggered, MITA/STING-Dependent Innate Immune Signaling.* Cell host & microbe, 2014. 16(4): p. 450-461.

430. Lahouassa, H., et al., *SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates.* Nature immunology, 2012. 13: p. 223.

431. Goldstone, D.C., et al., *HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase.* Nature, 2011. 480: p. 379.

432. Berger, A., et al., *SAMHD1-Deficient CD14+ Cells from Individuals with Aicardi-Goutières Syndrome Are Highly Susceptible to HIV-1 Infection.* PLOS Pathogens, 2011. 7(12): p. e1002425.

433. Guo, H., et al., *NLRX1 Sequesters STING to Negatively Regulate the Interferon Response, Thereby Facilitating the Replication of HIV-1 and DNA Viruses.* Cell host & microbe, 2016. 19(4): p. 515-528.

434. Abdul-Sater, A.A., et al., *Enhancement of Reactive Oxygen Species Production and Chlamydial Infection by the Mitochondrial Nod-like Family Member NLRX1.* The Journal of Biological Chemistry, 2010. 285(53): p. 41637-41645.

435. Tattoli, I., et al., *NLRX1 is a mitochondrial NOD-like receptor that amplifies NF-κB and JNK pathways by inducing reactive oxygen species production.* EMBO Reports, 2008. 9(3): p. 293-300.

436. Allen, Irving C., et al., *NLRX1 Protein Attenuates Inflammatory Responses to Infection by Interfering with the RIG-I-MAVS and TRAF6-NF-κB Signaling Pathways.* Immunity, 2011. 34(6): p. 854-865.

437. Xia, X., et al., *NLRX1 Negatively Regulates TLR-Induced NF-κB Signaling by Targeting TRAF6 and IKK.* Immunity, 2011. 34(6): p. 843-853.

438. Lo Cigno, I., et al., *The Nuclear DNA Sensor IFI16 Acts as a Restriction Factor for Human Papillomavirus Replication through Epigenetic Modifications of the Viral Promoters.* Journal of Virology, 2015. 89(15): p. 7506-7520.

439. Gariano, G.R., et al., *The Intracellular DNA Sensor IFI16 Gene Acts as Restriction Factor for Human Cytomegalovirus Replication.* PLOS Pathogens, 2012. 8(1): p. e1002498.

440. Iqbal, J., et al., *Histone H2B-IFI16 Recognition of Nuclear Herpesviral Genome Induces Cytoplasmic Interferon-β Responses.* PLOS Pathogens, 2016. 12(10): p. e1005967.

441. Takahama, M., et al., *The RAB2B-GARIL5 Complex Promotes Cytosolic DNA-Induced Innate Immune Responses.* Cell Reports, 2017. 20(12): p. 2944-2954.

442. Fang, R., et al., *NEMO–IKKβ Are Essential for IRF3 and NF-κB Activation in the cGAS–STING Pathway.* The Journal of Immunology, 2017. 199(9): p. 3222-3233.

443. Chiu, Y.-H., J.B. MacMillan, and Z.J. Chen, *RNA Polymerase III Detects Cytosolic DNA and Induces Type I Interferons through the RIG-I Pathway.* Cell, 2009. 138(3): p. 576-591.

444. Nazmi, A., et al., *STING Mediates Neuronal Innate Immune Response Following Japanese Encephalitis Virus Infection.* Scientific Reports, 2012. 2: p. 347.

445. Suzuki, T., et al., *Cell Type-Specific Subcellular Localization of Phospho-TBK1 in Response to Cytoplasmic Viral DNA.* PLoS ONE, 2013. 8(12): p. e83639.

446. Tattoli, I., et al., *Amino Acid Starvation Induced by Invasive Bacterial Pathogens Triggers an Innate Host Defense Program.* Cell host & microbe, 2012. 11(6): p. 563-575.

447.   Xia, M., et al., *Mitophagy Enhances Oncolytic Measles Virus Replication by Mitigating DDX58/RIG-I-Like Receptor Signaling.* Journal of Virology, 2014. 88(9): p. 5152-5164.

448.   Lei, Y., et al., *The Mitochondrial Proteins NLRX1 and TUFM Form a Complex that Regulates Type I Interferon and Autophagy.* Immunity, 2012. 36(6): p. 933-946.

449.   Watson, R.O., P.S. Manzanillo, and J.S. Cox, *Extracellular M. tuberculosis DNA Targets Bacteria for Autophagy by Activating the Host DNA-Sensing Pathway.* Cell, 2012. 150(4): p. 803-815.

450.   McFarlane, S., et al., *Early Induction of Autophagy in Human Fibroblasts after Infection with Human Cytomegalovirus or Herpes Simplex Virus 1.* Journal of Virology, 2011. 85(9): p. 4212-4221.

451.   Rasmussen, S.B., et al., *Activation of autophagy by alpha-herpesviruses in myeloid cells is mediated by cytoplasmic viral DNA through a mechanism dependent on STING.* Journal of immunology (Baltimore, Md. : 1950), 2011. 187(10): p. 5268-5276.

452.   Liang, C., et al., *Autophagic and tumour suppressor activity of a novel Beclin1-binding protein UVRAG.* Nature Cell Biology, 2006. 8: p. 688.

453.   Matsunaga, K., et al., *Two Beclin 1-binding proteins, Atg14L and Rubicon, reciprocally regulate autophagy at different stages.* Nature Cell Biology, 2009. 11: p. 385.

454.   Zhong, Y., et al., *Distinct regulation of autophagic activity by Atg14L and Rubicon associated with Beclin 1–phosphatidylinositol-3-kinase complex.* Nature Cell Biology, 2009. 11: p. 468.

455.   Pattingre, S., et al., *Bcl-2 Antiapoptotic Proteins Inhibit Beclin 1-Dependent Autophagy.* Cell, 2005. 122(6): p. 927-939.

456.   Liang, Q., et al., *Crosstalk between the cGAS DNA Sensor and Beclin-1 Autophagy Protein Shapes Innate Antimicrobial Immune Responses.* Cell host & microbe, 2014. 15(2): p. 228-238.

457.   Mathew, R., et al., *Functional Role of Autophagy-Mediated Proteome Remodeling in Cell Survival Signaling and Innate Immunity.* Molecular Cell, 2014. 55(6): p. 916-930.

458.   Grimm, W.A., et al., *The Thr300Ala variant in ATG16L1 is associated with improved survival in human colorectal cancer and enhanced production of type I interferon.* Gut, 2016. 65(3): p. 456-464.

459.   Marchiando, Amanda M., et al., *A Deficiency in the Autophagy Gene Atg16L1 Enhances Resistance to Enteric Bacterial Infection.* Cell host & microbe, 2013. 14(2): p. 216-224.

460.   Lan, Yuk Y., et al., *Dnase2a Deficiency Uncovers Lysosomal Clearance of Damaged Nuclear DNA via Autophagy.* Cell Reports, 2014. 9(1): p. 180-192.

461.   Allen, Irving C., et al., *NLRP12 Suppresses Colon Inflammation and Tumorigenesis through the Negative Regulation of Noncanonical NF-κB Signaling.* Immunity, 2012. 36(5): p. 742-754.

462.   Zaki, M.H., et al., *The NOD-Like Receptor NLRP12 Attenuates Colon Inflammation and Tumorigenesis.* Cancer Cell, 2011. 20(5): p. 649-660.

463.   Cui, J., et al., *NLRC5 Negatively Regulates the NF-κB and Type I Interferon Signaling Pathways.* Cell, 2010. 141(3): p. 483-496.

464.   Zhang, L., et al., *NLRC3, a Member of the NLR Family of Proteins, Is a Negative Regulator of Innate Immune Signaling Induced by the DNA Sensor STING.* Immunity, 2014. 40(3): p. 329-341.

465.	Schneider, M., et al., *The innate immune sensor NLRC3 attenuates Toll-like receptor signaling via modification of the signaling adaptor TRAF6 and transcription factor NF-κB.* Nature immunology, 2012. 13(9): p. 823-831.

466.	Cui, J., et al., *NLRP4 negatively regulates type I interferon signaling by targeting the kinase TBK1 for degradation via the ubiquitin ligase DTX4.* Nature immunology, 2012. 13(4): p. 387-395.

467.	Abe, T., et al., *Germ-Cell-Specific Inflammasome Component NLRP14 Negatively Regulates Cytosolic Nucleic Acid Sensing to Promote Fertilization.* Immunity, 2017. 46(4): p. 621-634.

468.	Corrales, L., et al., *Antagonism of the STING pathway via activation of the AIM2 inflammasome by intracellular DNA.* Journal of immunology (Baltimore, Md. : 1950), 2016. 196(7): p. 3191-3198.

469.	Wang, Y., et al., *Inflammasome Activation Triggers Caspase-1-Mediated Cleavage of cGAS to Regulate Responses to DNA Virus Infection.* Immunity, 2017. 46(3): p. 393-404.

470.	Monroe, K.M., et al., *IFI16 DNA Sensor Is Required for Death of Lymphoid CD4 T Cells Abortively Infected with HIV.* Science, 2014. 343(6169): p. 428-432.

471.	Ghosh, S., et al., *The PYHIN Protein p205 Regulates the Inflammasome by Controlling Asc Expression.* The Journal of Immunology, 2017. 199(9): p. 3249-3260.

472.	Rathinam, Vijay A.K., et al., *TRIF Licenses Caspase-11-Dependent NLRP3 Inflammasome Activation by Gram-Negative Bacteria.* Cell, 2012. 150(3): p. 606-619.

473.	Broz, P., et al., *Caspase-11 increases susceptibility to Salmonella infection in the absence of caspase-1.* Nature, 2012. 490: p. 288.

474.	Cole, L.E., et al., *Macrophage Proinflammatory Response to Francisella tularensis Live Vaccine Strain Requires Coordination of Multiple Signaling Pathways.* The Journal of Immunology, 2008. 180(10): p. 6885-6891.

475.	Henry, T., et al., *Type I interferon signaling is required for activation of the inflammasome during Francisella infection.* The Journal of Experimental Medicine, 2007. 204(5): p. 987-994.

476.	Bauernfeind, F.G., et al., *Cutting Edge: NF-κB Activating Pattern Recognition and Cytokine Receptors License NLRP3 Inflammasome Activation by Regulating NLRP3 Expression.* The Journal of Immunology, 2009. 183(2): p. 787-791.

477.	Liu, F., et al., *Priming and Activation of Inflammasome by Canarypox Virus Vector ALVAC via the cGAS/IFI16–STING–Type I IFN Pathway and AIM2 Sensor.* The Journal of Immunology, 2017. 199(9): p. 3293-3305.

478.	Webster, S.J., et al., *Detection of a microbial metabolite by STING regulates inflammasome activation in response to Chlamydia trachomatis infection.* PLOS Pathogens, 2017. 13(6): p. e1006383.

479.	Molnarfi, N., et al., *The Production of IL-1 Receptor Antagonist in IFN-β-Stimulated Human Monocytes Depends on the Activation of Phosphatidylinositol 3-Kinase but Not of STAT1.* The Journal of Immunology, 2005. 174(5): p. 2974-2980.

480.	Guarda, G., et al., *Type I Interferon Inhibits Interleukin-1 Production and Inflammasome Activation.* Immunity, 2011. 34(2): p. 213-223.

481.   Hernandez-Cuellar, E., et al., *Cutting Edge: Nitric Oxide Inhibits the NLRP3 Inflammasome.* The Journal of Immunology, 2012. 189(11): p. 5113-5117.

482.   Mishra, B.B., et al., *Nitric oxide controls the immunopathology of tuberculosis by inhibiting NLRP3 inflammasome–dependent processing of IL-1β.* Nature immunology, 2012. 14: p. 52.

483.   Swanson, K.V., et al., *A noncanonical function of cGAMP in inflammasome priming and activation.* The Journal of Experimental Medicine, 2017.

484.   Kovalenko, A., et al., *Caspase-8 deficiency in epidermal keratinocytes triggers an inflammatory skin disease.* The Journal of Experimental Medicine, 2009. 206(10): p. 2161-2177.

485.   Cuda, C.M., et al., *Caspase-8 Acts as a Molecular Rheostat to Limit RIPK1- and MyD88-Mediated Dendritic Cell Activation().* Journal of immunology (Baltimore, Md. : 1950), 2014. 192(12): p. 5548-5560.

486.   Rajput, A., et al., *RIG-I RNA Helicase Activation of IRF3 Transcription Factor Is Negatively Regulated by Caspase-8-Mediated Cleavage of the RIP1 Protein.* Immunity, 2011. 34(3): p. 340-351.

487.   Andrade, W.A., et al., *Group B Streptococcus degrades cyclic-di-AMP to modulate STING-dependent type I interferon production.* Cell host & microbe, 2016. 20(1): p. 49-59.

488.   Dey, R.J., et al., *Inhibition of innate immune cytosolic surveillance by an M. tuberculosis phosphodiesterase.* Nature Chemical Biology, 2016. 13: p. 210.

489.   Dobbs, N., et al., *STING activation by translocation from the ER is associated with infection and autoinflammatory disease.* Cell host & microbe, 2015. 18(2): p. 157-168.

490.   Christensen, M.H., et al., *HSV-1 ICP27 targets the TBK1-activated STING signalsome to inhibit virus-induced type I IFN expression.* The EMBO Journal, 2016. 35(13): p. 1385-1399.

491.   Lau, L., et al., *DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway.* Science, 2015. 350(6260): p. 568-571.

492.   Lam, E. and E. Falck-Pedersen, *Unabated Adenovirus Replication following Activation of the cGAS/STING-Dependent Antiviral Response in Human Cells.* Journal of Virology, 2014. 88(24): p. 14426-14439.

493.   Look, D.C., et al., *Direct Suppression of Stat1 Function during Adenoviral Infection.* Immunity, 1998. 9(6): p. 871-880.

494.   Liu, Y., et al., *Hepatitis B Virus Polymerase Disrupts K63-Linked Ubiquitination of STING To Block Innate Cytosolic DNA-Sensing Pathways.* Journal of Virology, 2015. 89(4): p. 2287-2300.

495.   Ding, Q., et al., *Hepatitis C virus NS4B blocks the interaction of STING and TBK1 to evade host innate immunity.* Journal of Hepatology, 2013. 59(1): p. 52-58.

496.   Nitta, S., et al., *Hepatitis C virus NS4B protein targets STING and abrogates RIG-I–mediated type I interferon-dependent innate immunity.* Hepatology, 2013. 57(1): p. 46-58.

497.   Rodriguez-Madoz, J.R., et al., *Inhibition of the Type I Interferon Response in Human Dendritic Cells by Dengue Virus Infection Requires a Catalytically Active NS2B3 Complex.* Journal of Virology, 2010. 84(19): p. 9760-9774.

498.   Aguirre, S., et al., *DENV Inhibits Type I IFN Production in Infected Cells by Cleaving Human STING.* PLOS Pathogens, 2012. 8(10): p. e1002934.

499.   Yu, C.-Y., et al., *Dengue Virus Targets the Adaptor Protein MITA to Subvert Host Innate Immunity.* PLOS Pathogens, 2012. 8(6): p. e1002780.

500.   Sun, L., et al., *Coronavirus Papain-like Proteases Negatively Regulate Antiviral Innate Immune Response through Disruption of STING-Mediated Signaling.* PLoS ONE, 2012. 7(2): p. e30802.

501.   Clementz, M.A., et al., *Deubiquitinating and Interferon Antagonism Activities of Coronavirus Papain-Like Proteases.* Journal of Virology, 2010. 84(9): p. 4619-4629.

502.   Xing, Y., et al., *The papain-like protease of porcine epidemic diarrhea virus negatively regulates type I interferon pathway by acting as a viral deubiquitinase.* The Journal of General Virology, 2013. 94(Pt 7): p. 1554-1567.

503.   Chen, X., et al., *SARS coronavirus papain-like protease inhibits the type I interferon signaling pathway through interaction with the STING-TRAF3-TBK1 complex.* Protein & Cell, 2014. 5(5): p. 369-381.

504.   Wu, J.-j., et al., *Inhibition of cGAS DNA Sensing by a Herpesvirus Virion Protein.* Cell host & microbe, 2015. 18(3): p. 333-344.

505.   Zhang, G., et al., *Cytoplasmic isoforms of Kaposi sarcoma herpesvirus LANA recruit and antagonize the innate immune DNA sensor cGAS.* Proceedings of the National Academy of Sciences of the United States of America, 2016. 113(8): p. E1034-E1043.

506.   Sun, C., et al., *Evasion of innate cytosolic DNA sensing by a gammaherpesvirus facilitates establishment of latent infection().* Journal of immunology (Baltimore, Md. : 1950), 2015. 194(4): p. 1819-1831.

507.   Lee, K., et al., *Flexible Use of Nuclear Import Pathways by HIV-1.* Cell host & microbe, 2010. 7(3): p. 221-233.

508.   Schaller, T., et al., *HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency.* PLOS Pathogens, 2011. 7(12): p. e1002439.

509.   Rasaiyaah, J., et al., *HIV-1 evades innate immune recognition through specific cofactor recruitment.* Nature, 2013. 503: p. 402.

510.   Yan, N., et al., *The cytosolic exonuclease TREX1 inhibits the innate immune response to HIV-1.* Nature immunology, 2010. 11(11): p. 1005-1013.

511.   Manel, N., et al., *A cryptic sensor for HIV-1 activates antiviral innate immunity in dendritic cells.* Nature, 2010. 467: p. 214.

512.   Lahaye, X., et al., *The Capsids of HIV-1 and HIV-2 Determine Immune Detection of the Viral cDNA by the Innate Sensor cGAS in Dendritic Cells.* Immunity, 2013. 39(6): p. 1132-1142.

513.   Kalamvoki, M. and B. Roizman, *HSV-1 degrades, stabilizes, requires, or is stung by STING depending on ICP0, the US3 protein kinase, and cell derivation.* Proceedings of the National Academy of Sciences, 2014. 111(5): p. E611-E617.

514.   Archer, K.A., J. Durack, and D.A. Portnoy, *STING-Dependent Type I IFN Production Inhibits Cell-Mediated Immunity to Listeria monocytogenes.* PLOS Pathogens, 2014. 10(1): p. e1003861.

515.   Osborne, S.E., et al., *Type I interferon promotes cell-to-cell spread of Listeria monocytogenes.* Cellular Microbiology, 2017. 19(3): p. e12660-n/a.

516.   Shimada, K., et al., *Oxidized Mitochondrial DNA Activates the NLRP3 Inflammasome during Apoptosis.* Immunity, 2012. 36(3): p. 401-414.

517.   Nakahira, K., et al., *Autophagy proteins regulate innate immune response by inhibiting NALP3 inflammasome-mediated mitochondrial DNA release.* Nature immunology, 2011. 12(3): p. 222-230.

518.   Carlos, D., et al., *Mitochondrial DNA Activates the NLRP3 Inflammasome and Predisposes to Type 1 Diabetes in Murine Model.* Frontiers in Immunology, 2017. 8: p. 164.

519.   Liu, J., et al., *Circulating Cell Free Mitochondrial DNA is a Biomarker in the Development of Coronary Heart Disease in the Patients with Type 2 Diabetes.* Clin Lab, 2015. 61(7): p. 661-7.

520.   Jönsen, A., et al., *Mitochondrial DNA polymorphisms are associated with susceptibility and phenotype of systemic lupus erythematosus.* Lupus, 2009. 18(4): p. 309-312.

521.   Perl, A., R. Hanczko, and E. Doherty, *Assessment of Mitochondrial Dysfunction in Lymphocytes of Patients with Systemic Lupus Erythematosus*, in *Autoimmunity: Methods and Protocols*, A. Perl, Editor. 2012, Humana Press: Totowa, NJ. p. 61-89.

522.   Dall'Era, M., et al., *Type I interferon correlates with serological and clinical manifestations of SLE.* Annals of the Rheumatic Diseases, 2005. 64(12): p. 1692-1697.

523.   Bauer, J.W., et al., *Elevated Serum Levels of Interferon-Regulated Chemokines Are Biomarkers for Active Human Systemic Lupus Erythematosus.* PLoS Medicine, 2006. 3(12): p. e491.

524.   Wang, H., et al., *Neutrophil Extracellular Trap Mitochondrial DNA and Its Autoantibody in Systemic Lupus Erythematosus and a Proof-of-Concept Trial of Metformin.* Arthritis & Rheumatology, 2015. 67(12): p. 3190-3200.

525.   Caielli, S., et al., *Oxidized mitochondrial nucleoids released by neutrophils drive type I interferon production in human lupus.* The Journal of Experimental Medicine, 2016. 213(5): p. 697-713.

526.   Reijns, Martin A.M., et al., *Enzymatic Removal of Ribonucleotides from DNA Is Essential for Mammalian Genome Integrity and Development.* Cell, 2012. 149(5): p. 1008-1022.

527.   Bartsch, K., et al., *Absence of RNase H2 triggers generation of immunogenic micronuclei removed by autophagy.* Human Molecular Genetics, 2017. 26(20): p. 3960-3972.

528.   Mackenzie, K.J., et al., *Ribonuclease H2 mutations induce a cGAS/STING-dependent innate immune response.* The EMBO Journal, 2016. 35(8): p. 831-844.

529.   Pokatayev, V., et al., *RNase H2 catalytic core Aicardi-Goutières syndrome–related mutant invokes cGAS–STING innate immune-sensing pathway in mice.* The Journal of Experimental Medicine, 2016. 213(3): p. 329-336.

530.   Crow, Y.J., et al., *Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and mimic congenital viral brain infection.* Nature Genetics, 2006. 38: p. 910.

531.   Günther, C., et al., *Defective removal of ribonucleotides from DNA promotes systemic autoimmunity.* The Journal of Clinical Investigation, 2015. 125(1): p. 413-424.

532.   Crow, Y.J., et al., *Characterization of Human Disease Phenotypes Associated with Mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and IFIH1.* American journal of medical genetics. Part A, 2015. 0(2): p. 296-312.

533.	Yang, Y.-G., T. Lindahl, and D.E. Barnes, *Trex1 Exonuclease Degrades ssDNA to Prevent Chronic Checkpoint Activation and Autoimmune Disease.* Cell, 2007. 131(5): p. 873-886.

534.	Ablasser, A., et al., *TREX1 Deficiency Triggers Cell-Autonomous Immunity in a cGAS-Dependent Manner.* The Journal of Immunology, 2014. 192(12): p. 5993-5997.

535.	Gray, E.E., et al., *cGAS is required for lethal autoimmune disease in the Trex1-deficient mouse model of Aicardi-Goutieres Syndrome().* Journal of immunology (Baltimore, Md. : 1950), 2015. 195(5): p. 1939-1943.

536.	Crow, Y.J., et al., *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutières syndrome at the AGS1 locus.* Nature Genetics, 2006. 38: p. 917.

537.	Richards, A., et al., *C-terminal truncations in human 3'-5' DNA exonuclease TREX1 cause autosomal dominant retinal vasculopathy with cerebral leukodystrophy.* Nature Genetics, 2007. 39: p. 1068.

538.	Rice, G., et al., *Heterozygous Mutations in TREX1 Cause Familial Chilblain Lupus and Dominant Aicardi-Goutières Syndrome.* American Journal of Human Genetics, 2007. 80(4): p. 811-815.

539.	Lee-Kirsch, M.A., et al., *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus.* Nature Genetics, 2007. 39: p. 1065.

540.	Lindahl, T., et al., *Biochemical properties of mammalian TREX1 and its association with DNA replication and inherited inflammatory disease.* Biochemical Society Transactions, 2009. 37(3): p. 535-538.

541.	Mannherz, H.G., et al., *A New Function for an Old Enzyme: The Role of DNase I in Apoptosis*, in *Pathways for Cytolysis*, G.M. Griffiths and J. Tschopp, Editors. 1995, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 161-174.

542.	Napirei, M., S. Wulf, and H.G. Mannherz, *Chromatin breakdown during necrosis by serum Dnase1 and the plasminogen system.* Arthritis & Rheumatism, 2004. 50(6): p. 1873-1883.

543.	Skiljevic, D., et al., *Serum DNase I activity in systemic lupus erythematosus: correlation with immunoserological markers, the disease activity and organ involvement*, in *Clinical Chemistry and Laboratory Medicine*2013. p. 1083.

544.	Yasutomo, K., et al., *Mutation of DNASE1 in people with systemic lupus erythematosus.* Nature Genetics, 2001. 28: p. 313.

545.	Bodaño, A., et al., *Study of DNASE I gene polymorphisms in systemic lupus erythematosus susceptibility.* Annals of the Rheumatic Diseases, 2007. 66(4): p. 560-561.

546.	Bodaño, A., et al., *Association of a non-synonymous single-nucleotide polymorphism of DNASEI with SLE susceptibility.* Rheumatology, 2006. 45(7): p. 819-823.

547.	Kawane, K., et al., *Requirement of DNase II for Definitive Erythropoiesis in the Mouse Fetal Liver.* Science, 2001. 292(5521): p. 1546-1549.

548.	Krieser, R.J., et al., *Deoxyribonuclease IIα is required during the phagocytic phase of apoptosis and its loss causes perinatal lethality.* Cell Death And Differentiation, 2002. 9: p. 956.

549.	Ahn, J., et al., *STING manifests self DNA-dependent inflammatory disease.* Proceedings of the National Academy of Sciences of the United States of America, 2012. 109(47): p. 19386-19391.

550.   Laguette, N., et al., *SAMHD1 is the dendritic– and myeloid–cell–specific HIV–1 restriction factor counteracted by Vpx.* Nature, 2011. 474(7353): p. 654-657.

551.   Hrecka, K., et al., *Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein.* Nature, 2011. 474(7353): p. 658-661.

552.   Kretschmer, S., et al., *SAMHD1 prevents autoimmunity by maintaining genome stability.* Annals of the Rheumatic Diseases, 2015. 74(3): p. e17-e17.

553.   Maelfait, J., et al., *Restriction by SAMHD1 Limits cGAS/STING-Dependent Innate and Adaptive Immune Responses to HIV-1.* Cell Reports, 2016. 16(6): p. 1492-1501.

554.   George, C.X., L. John, and C.E. Samuel, *An RNA Editor, Adenosine Deaminase Acting on Double-Stranded RNA (ADAR1).* Journal of Interferon & Cytokine Research, 2014. 34(6): p. 437-446.

555.   Neeman, Y., et al., *RNA editing level in the mouse is determined by the genomic repeat repertoire.* RNA, 2006. 12(10): p. 1802-1809.

556.   Liddicoat, B.J., et al., *RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself.* Science, 2015. 349(6252): p. 1115-1120.

557.   Pestal, K., et al., *Isoforms of RNA-Editing Enzyme ADAR1 Independently Control Nucleic Acid Sensor MDA5-Driven Autoimmunity and Multi-organ Development.* Immunity, 2015. 43(5): p. 933-944.

558.   Jeremiah, N., et al., *Inherited STING-activating mutation underlies a familial inflammatory syndrome with lupus-like manifestations.* The Journal of Clinical Investigation, 2014. 124(12): p. 5516-5520.

559.   Liu , Y., et al., *Activated STING in a Vascular and Pulmonary Syndrome.* New England Journal of Medicine, 2014. 371(6): p. 507-518.

560.   Tang, E.D. and C.-Y. Wang, *Single Amino Acid Change in STING Leads to Constitutive Active Signaling.* PLoS ONE, 2015. 10(3): p. e0120090.

561.   König, N., et al., *Familial chilblain lupus due to a gain-of-function mutation in STING.* Annals of the Rheumatic Diseases, 2017. 76(2): p. 468-472.

562.   Dunn, G.P., et al., *A critical function for type I interferons in cancer immunoediting.* Nature immunology, 2005. 6: p. 722.

563.   Woo, S.-R., L. Corrales, and T.F. Gajewski, *The STING pathway and the T cell-inflamed tumor microenvironment.* Trends in Immunology, 2015. 36(4): p. 250-256.

564.   Yang, X., et al., *Targeting the Tumor Microenvironment with Interferon-β Bridges Innate and Adaptive Immune Responses.* Cancer Cell, 2014. 25(1): p. 37-48.

565.   Woo, S.-R., et al., *STING-Dependent Cytosolic DNA Sensing Mediates Innate Immune Recognition of Immunogenic Tumors.* Immunity, 2014. 41(5): p. 830-842.

566.   Deng, L., et al., *STING-Dependent Cytosolic DNA Sensing Promotes Radiation-Induced Type I Interferon-Dependent Antitumor Immunity in Immunogenic Tumors.* Immunity, 2014. 41(5): p. 843-852.

567.   Liu, X., et al., *CD47 blockade triggers T cell–mediated destruction of immunogenic tumors.* Nature Medicine, 2015. 21: p. 1209.

568.	Klarquist, J., et al., *STING-mediated DNA sensing promotes antitumor and autoimmune responses to dying cells.* Journal of immunology (Baltimore, Md. : 1950), 2014. 193(12): p. 6124-6134.

569.	Diamond, M.S., et al., *Type I interferon is selectively required by dendritic cells for immune rejection of tumors.* The Journal of Experimental Medicine, 2011. 208(10): p. 1989-2003.

570.	Ho, Samantha S.W., et al., *The DNA Structure-Specific Endonuclease MUS81 Mediates DNA Sensor STING-Dependent Host Rejection of Prostate Cancer Cells.* Immunity, 2016. 44(5): p. 1177-1189.

571.	Ahn, J., et al., *Inflammation-driven carcinogenesis is mediated through STING.* Nature Communications, 2014. 5: p. 5166.

572.	Carroll, E.C., et al., *The Vaccine Adjuvant Chitosan Promotes Cellular Immunity via DNA Sensor cGAS-STING-Dependent Induction of Type I Interferons.* Immunity, 2016. 44(3): p. 597-608.

573.	Wang, H., et al., *cGAS is essential for the antitumor effect of immune checkpoint blockade.* Proceedings of the National Academy of Sciences of the United States of America, 2017. 114(7): p. 1637-1642.

574.	Alsaab, H.O., et al., *PD-1 and PD-L1 Checkpoint Signaling Inhibition for Cancer Immunotherapy: Mechanism, Combinations, and Clinical Outcome.* Frontiers in Pharmacology, 2017. 8: p. 561.

575.	Fu, J., et al., *STING agonist formulated cancer vaccines can cure established tumors resistant to PD-1 blockade.* Science translational medicine, 2015. 7(283): p. 283ra52-283ra52.

576.	Corrales, L., et al., *Direct Activation of STING in the Tumor Microenvironment Leads to Potent and Systemic Tumor Regression and Immunity.* Cell Reports, 2015. 11(7): p. 1018-1030.

577.	Li, T., et al., *Antitumor Activity of cGAMP via Stimulation of cGAS-cGAMP-STING-IRF3 Mediated Innate Immune Response.* Scientific Reports, 2016. 6: p. 19049.

578.	Demaria, O., et al., *STING activation of tumor endothelial cells initiates spontaneous and therapeutic antitumor immunity.* Proceedings of the National Academy of Sciences of the United States of America, 2015. 112(50): p. 15408-15413.

579.	Ohkuri, T., et al., *STING contributes to anti-glioma immunity via triggering type-I IFN signals in the tumor microenvironment.* Cancer immunology research, 2014. 2(12): p. 1199-1208.

580.	Tang, C.-H.A., et al., *Agonist-mediated activation of STING induces apoptosis in malignant B cells.* Cancer research, 2016. 76(8): p. 2137-2152.

581.	Curran, E., et al., *STING Pathway Activation Stimulates Potent Immunity against Acute Myeloid Leukemia.* Cell Reports, 2016. 15(11): p. 2357-2366.

582.	Xia, T., et al., *Deregulation of STING Signaling in Colorectal Carcinoma Constrains DNA Damage Responses and Correlates With Tumorigenesis.* Cell Reports, 2016. 14(2): p. 282-297.

583.	Xia, T., H. Konno, and G.N. Barber, *Recurrent Loss of STING Signaling in Melanoma Correlates with Susceptibility to Viral Oncolysis.* Cancer research, 2016. 76(22): p. 6747-6759.

584.	Dilley, R.L. and R.A. Greenberg, *ALTernative Telomere Maintenance and Cancer.* Trends in Cancer, 2015. 1(2): p. 145-156.

585.	Henson, J.D., et al., *DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity.* Nature Biotechnology, 2009. 27: p. 1181.

586.	Bryan, T.M., et al., *Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines.* Nature Medicine, 1997. 3: p. 1271.

587.	Chen, Y.-A., et al., *Extrachromosomal telomere repeat DNA is linked to ALT development via cGAS-STING DNA sensing pathway.* Nature Structural &Amp; Molecular Biology, 2017.

588.	Swann, J.B., et al., *Demonstration of inflammation-induced cancer and cancer immunoediting during primary tumorigenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2008. 105(2): p. 652-656.

589.	Song, S., et al., *Decreased expression of STING predicts poor prognosis in patients with gastric cancer.* Scientific Reports, 2017. 7: p. 39858.

590.	Gaston, J., et al., *Intracellular STING inactivation sensitizes breast cancer cells to genotoxic agents.* Oncotarget, 2016. 7(47): p. 77205-77224.

591.	Huang, L., et al., *DNA sensing via the Stimulator of Interferon Genes (STING) adaptor in myeloid dendritic cells induces potent tolerogenic responses().* Journal of immunology (Baltimore, Md. : 1950), 2013. 191(7): p. 3509-3513.

592.	Munn, D.H. and A.L. Mellor, *IDO in the Tumor Microenvironment: Inflammation, Counter-Regulation, and Tolerance.* Trends in Immunology, 2016. 37(3): p. 193-207.

593.	Lemos, H., et al., *STING promotes the growth of tumors characterized by low antigenicity via IDO activation.* Cancer research, 2016. 76(8): p. 2076-2081.

594.	Hall, J., et al., *Discovery of PF-06928215 as a high affinity inhibitor of cGAS enabled by a novel fluorescence polarization assay.* PLoS ONE, 2017. 12(9): p. e0184843.

595.	Kellenberger, C.A. and M.C. Hammond, *Chapter Eight - In Vitro Analysis of Riboswitch–Spinach Aptamer Fusions as Metabolite-Sensing Fluorescent Biosensors*, in *Methods in Enzymology*, D.H. Burke-Aguero, Editor. 2015, Academic Press. p. 147-172.

596.	Kellenberger, C.A., et al., *RNA-Based Fluorescent Biosensors for Live Cell Imaging of Second Messengers Cyclic di-GMP and Cyclic AMP-GMP.* Journal of the American Chemical Society, 2013. 135(13): p. 4906-4909.

597.	Bose, D., et al., *An RNA-Based Fluorescent Biosensor for High-Throughput Analysis of the cGAS-cGAMP-STING Pathway.* Cell Chemical Biology, 2016. 23(12): p. 1539-1549.

598.	Ren, A., et al., *Structural Basis for Molecular Discrimination by a 3',3'-cGAMP Sensing Riboswitch.* Cell Reports, 2015. 11(1): p. 1-12.

599.	Roembke, B.T., et al., *A cyclic dinucleotide containing 2-aminopurine is a general fluorescent sensor for c-di-GMP and 3[prime or minute],3[prime or minute]-cGAMP.* Molecular BioSystems, 2014. 10(6): p. 1568-1575.

600.	Andreeva, L., et al., *cGAS senses long and HMGB/TFAM-bound U-turn DNA by forming protein–DNA ladders.* Nature, 2017. 549(7672): p. 394-398.

601.	Luecke, S., et al., *cGAS is activated by DNA in a length-dependent manner.* EMBO Reports, 2017. 18(10): p. 1707-1715.

602.   Zhang, Y., et al., *The DNA sensor, cyclic GMP-AMP synthase (cGAS) is essential for induction of IFN beta during Chlamydia trachomatis infection( ).* Journal of immunology (Baltimore, Md. : 1950), 2014. 193(5): p. 2394-2404.

603.   Lee, A., et al., *The N terminus of cGAS de-oligomerizes the cGAS:DNA complex and lifts the DNA size restriction of core-cGAS activity.* FEBS Letters, 2017. 591(6): p. 954-961.

604.   Sun, B., et al., *Dengue virus activates cGAS through the release of mitochondrial DNA.* Scientific Reports, 2017. 7(1): p. 3594.

605.   Dejean, L.M., et al., *Oligomeric Bax Is a Component of the Putative Cytochrome c Release Channel MAC, Mitochondrial Apoptosis-induced Channel.* Molecular Biology of the Cell, 2005. 16(5): p. 2424-2432.

606.   Martinez-Caballero, S., et al., *Assembly of the Mitochondrial Apoptosis-induced Channel, MAC.* Journal of Biological Chemistry, 2009. 284(18): p. 12235-12245.

607.   Torbet, J. and E. DiCapua, *Supercoiled DNA is interwound in liquid crystalline solutions.* The EMBO Journal, 1989. 8(13): p. 4351-4356.

608.   Yang, H., et al., *cGAS is essential for cellular senescence.* Proceedings of the National Academy of Sciences, 2017. 114(23): p. E4612-E4620.

609.   Shumilov, A., et al., *Epstein–Barr virus particles induce centrosome amplification and chromosomal instability.* Nature Communications, 2017. 8: p. 14257.

610.   Si, H. and E.S. Robertson, *Kaposi's Sarcoma-Associated Herpesvirus-Encoded Latency-Associated Nuclear Antigen Induces Chromosomal Instability through Inhibition of p53 Function.* Journal of Virology, 2006. 80(2): p. 697-709.

611.   Gruhne, B., et al., *The Epstein–Barr virus nuclear antigen-1 promotes genomic instability via induction of reactive oxygen species.* Proceedings of the National Academy of Sciences, 2009. 106(7): p. 2313-2318.

612.   Ivanov, S., et al., *A novel role for HMGB1 in TLR9-mediated inflammatory responses to CpG-DNA.* Blood, 2007. 110(6): p. 1970-1981.

613.   Dumitriu, I.E., et al., *Requirement of HMGB1 and RAGE for the maturation of human plasmacytoid dendritic cells.* European Journal of Immunology, 2005. 35(7): p. 2184-2190.

614.   Ngo, H.B., J.T. Kaiser, and D.C. Chan, *The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA.* Nature Structural &Amp; Molecular Biology, 2011. 18: p. 1290.

615.   Swinger, K.K., et al., *Flexible DNA bending in HU–DNA cocrystal structures.* The EMBO Journal, 2003. 22(14): p. 3749-3760.

616.   Hammel, M., et al., *HU multimerization shift controls nucleoid compaction.* Science Advances, 2016. 2(7): p. e1600650.

617.   Zhang, L., et al., *Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization.* Science (New York, N.Y.), 2015. 350(6259): p. 404-409.

618.   Peisley, A., et al., *Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments.* Proc Natl Acad Sci U S A, 2012. 109.

619.	Ngo, H.B., et al., *Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation.* Nature Communications, 2014. 5: p. 3077-3077.

620.	Lood, C., et al., *Neutrophil extracellular traps enriched in oxidized mitochondrial DNA are interferogenic and contribute to lupus-like disease.* Nature Medicine, 2016. 22: p. 146.

621.	Bonaldi, T., et al., *Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion.* The EMBO Journal, 2003. 22(20): p. 5551-5560.

622.	Xu, J., et al., *Identification of a Natural Viral RNA Motif That Optimizes Sensing of Viral RNA by RIG-I.* mBio, 2015. 6(5).

623.	Chen, Y.G., et al., *Sensing Self and Foreign Circular RNAs by Intron Identity.* Molecular Cell, 2017. 67(2): p. 228-238.e5.

624.	Lasda, E. and R. Parker, *Circular RNAs Co-Precipitate with Extracellular Vesicles: A Possible Mechanism for circRNA Clearance.* PLoS ONE, 2016. 11(2): p. e0148407.

625.	Sze, A., et al., *Host Restriction Factor SAMHD1 Limits Human T Cell Leukemia Virus Type 1 Infection of Monocytes via STING-Mediated Apoptosis.* Cell host & microbe, 2013. 14(4): p. 422-434.

626.	Herzner, A.-M., et al., *Sequence-specific activation of the DNA sensor cGAS by Y-form DNA structures as found in primary HIV-1 cDNA.* Nature immunology, 2015. 16: p. 1025.

627.	Dewannieux, M. and T. Heidmann, *Endogenous retroviruses: acquisition, amplification and taming of genome invaders.* Current Opinion in Virology, 2013. 3(6): p. 646-656.

628.	Molaro, A. and H.S. Malik, *Hide and seek: how chromatin-based pathways silence retroelements in the mammalian germline.* Current opinion in genetics & development, 2016. 37: p. 51-58.

629.	Goering, W., T. Ribarska, and W.A. Schulz, *Selective changes of retroelement expression in human prostate cancer.* Carcinogenesis, 2011. 32(10): p. 1484-1492.

630.	Kudo-Saito, C., et al., *Induction of immunoregulatory CD271+ cells by metastatic tumor cells that express human endogenous retrovirus H.* Cancer Res, 2014. 74(5): p. 1361-70.

631.	Szpakowski, S., et al., *Loss of epigenetic silencing in tumors preferentially affects primate-specific retroelements.* Gene, 2009. 448(2): p. 151-167.

632.	Hung, T., et al., *The Ro60 autoantigen binds endogenous retroelements and regulates inflammatory gene expression.* Science, 2015. 350(6259): p. 455-459.

633.	Chiappinelli, Katherine B., et al., *Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses.* Cell, 2016. 164(5): p. 1073.

634.	Roulois, D., et al., *DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts.* Cell, 2015. 162(5): p. 961-973.

635.	Karijolich, J., E. Abernathy, and B.A. Glaunsinger, *Infection-Induced Retrotransposon-Derived Noncoding RNAs Enhance Herpesviral Gene Expression via the NF-κB Pathway.* PLOS Pathogens, 2015. 11(11): p. e1005260.

636.	Zhao, K., et al., *Modulation of LINE-1 and Alu/SVA Retrotransposition by Aicardi-Goutières Syndrome-Related SAMHD1.* Cell Reports, 2013. 4(6): p. 1108-1115.

637.   Gasior, S.L., et al., *The Human LINE-1 Retrotransposon Creates DNA Double-strand Breaks.* Journal of molecular biology, 2006. 357(5): p. 1383-1393.

638.   Maxwell, P.H., W.C. Burhans, and M.J. Curcio, *Retrotransposition is associated with genome instability during chronological aging.* Proceedings of the National Academy of Sciences, 2011. 108(51): p. 20376-20381.

639.   Mankan, A.K., et al., *Cytosolic RNA:DNA hybrids activate the cGAS–STING axis.* The EMBO Journal, 2014. 33(24): p. 2937-2946.

640.   Rennekamp, A.J. and P.M. Lieberman, *Initiation of Epstein-Barr Virus Lytic Replication Requires Transcription and the Formation of a Stable RNA-DNA Hybrid Molecule at OriLyt.* Journal of Virology, 2011. 85(6): p. 2837-2850.

641.   Prichard, M.N., et al., *Identification of Persistent RNA-DNA Hybrid Structures within the Origin of Replication of Human Cytomegalovirus.* Journal of Virology, 1998. 72(9): p. 6997-7004.

642.   Yasukawa, T., et al., *Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand.* The EMBO Journal, 2006. 25(22): p. 5358-5371.

643.   Granier, S. and B. Kobilka, *A new era of GPCR structural and chemical biology.* Nature Chemical Biology, 2012. 8: p. 670.

644.   Endres, N.F., et al., *Regulation of the catalytic activity of the EGF receptor.* Current Opinion in Structural Biology, 2011. 21(6): p. 777-784.

645.   Salahpour, A., S. Angers, and M. Bouvier, *Functional Significance of Oligomerization of G-protein-coupled Receptors.* Trends in Endocrinology & Metabolism, 2000. 11(5): p. 163-168.

646.   Chabre, M. and M. le Maire, *Monomeric G-Protein-Coupled Receptor as a Functional Unit.* Biochemistry, 2005. 44(27): p. 9395-9403.

647.   Wu, H., *Higher-Order Assemblies in a New Paradigm of Signal Transduction.* Cell, 2013. 153(2): p. 287-292.

648.   Medzhitov, R., et al., *MyD88 Is an Adaptor Protein in the hToll/IL-1 Receptor Family Signaling Pathways.* Molecular Cell, 1998. 2(2): p. 253-258.

649.   Lin, S.-C., Y.-C. Lo, and H. Wu, *Helical assembly in the MyD88–IRAK4–IRAK2 complex in TLR/IL-1R signalling.* Nature, 2010. 465: p. 885.

650.   Ve, T., et al., *Structural basis of TIR-domain-assembly formation in MAL- and MyD88-dependent TLR4 signaling.* Nature Structural &Amp; Molecular Biology, 2017. 24: p. 743.

651.   Ferrao, R., et al., *IRAK4 Dimerization and trans-Autophosphorylation Are Induced by Myddosome Assembly.* Molecular Cell, 2014. 55(6): p. 891-903.

652.   Wang, L., et al., *Crystal structure of human IRAK1.* Proceedings of the National Academy of Sciences, 2017. 114(51): p. 13507-13512.

653.   Park, Y.C., et al., *Structural basis for self-association and receptor recognition of human TRAF2.* Nature, 1999. 398: p. 533.

654.   Ye, H., et al., *Distinct molecular mechanism for initiating TRAF6 signalling.* Nature, 2002. 418: p. 443.

655.   Yin, Q., et al., *E2 interaction and dimerization in the crystal structure of TRAF6.* Nature Structural &Amp; Molecular Biology, 2009. 16: p. 658.

656.   Ferrao, R., et al., *Structural Insights into the Assembly of Large Oligomeric Signalosomes in the Toll-Like Receptor–Interleukin-1 Receptor Superfamily.* Science Signaling, 2012. 5(226): p. re3-re3.

657.   Lu, A., et al., *Plasticity in PYD assembly revealed by cryo-EM structure of the PYD filament of AIM2.* Cell Discovery, 2015. 1: p. 15013.

658.   Vajjhala, P.R., et al., *The Inflammasome Adaptor ASC Induces Procaspase-8 Death Effector Domain Filaments.* Journal of Biological Chemistry, 2015. 290(49): p. 29217-29230.

659.   Lu, A., et al., *Molecular basis of caspase-1 polymerization and its inhibition by a new capping mechanism.* Nature Structural &Amp; Molecular Biology, 2016. 23: p. 416.

660.   Xu, H., et al., *Structural basis for the prion-like MAVS filaments in antiviral innate immunity.* eLife, 2014. 3: p. e01489.

661.   Baril, M., et al., *MAVS Dimer Is a Crucial Signaling Component of Innate Immunity and the Target of Hepatitis C Virus NS3/4A Protease.* Journal of Virology, 2009. 83(3): p. 1299-1311.

662.   Tsukamoto, H., et al., *Lipopolysaccharide-binding protein-mediated Toll-like receptor 4 dimerization enables rapid signal transduction against lipopolysaccharide stimulation on membrane-associated CD14-expressing cells.* International Immunology, 2010. 22(4): p. 271-280.

663.   Latz, E., et al., *Ligand-induced conformational changes allosterically activate Toll-like receptor 9.* Nature immunology, 2007. 8: p. 772.

664.   Motshwene, P.G., et al., *An Oligomeric Signaling Platform Formed by the Toll-like Receptor Signal Transducers MyD88 and IRAK-4.* Journal of Biological Chemistry, 2009. 284(37): p. 25404-25411.

665.   Jiang, X., et al., *Ubiquitin-induced oligomerization of the RNA sensors RIG-I and MDA5 activates antiviral innate immune response.* Immunity., 2012. 36.

666.   Jin, T., et al., *Structure of the Absent in Melanoma 2 (AIM2) Pyrin Domain Provides Insights into the Mechanisms of AIM2 Autoinhibition and Inflammasome Assembly.* The Journal of Biological Chemistry, 2013. 288(19): p. 13225-13235.

667.   del Toro Duany, Y., B. Wu, and S. Hur, *MDA5—filament, dynamics and disease.* Current Opinion in Virology, 2015. 12: p. 20-25.

668.   Fu, T.-M., et al., *Cryo-EM Structure of Caspase-8 Tandem DED Filament Reveals Assembly and Regulation Mechanisms of the Death-Inducing Signaling Complex.* Molecular Cell, 2016. 64(2): p. 236-250.

669.   Yin, Q., et al., *Molecular Mechanism for p202-Mediated Specific Inhibition of AIM2 Inflammasome Activation.* Cell Reports, 2013. 4(2): p. 327-339.

Acknowledgements

First of all I would like to thank my supervisor and mentor Prof. Dr. Karl-Peter Hopfner for a chance to work on such challenging and exciting projects, his vision and ideas that made this work possible. I deeply appreciate his sincere involvement and trust that helped me during my whole PhD and the joy of discovery that we shared together. Under his supervision I learned a lot as a specialist and a person and his support always encouraged me to face challenges without hesitation and to develop my own view and ideas.

I also want to express my sincere gratitude to Prof. Dr. Veit Hornung and Prof. Dr. Heinrich Leonhardt for their interest in my work, timely support and intellectual input, which were absolutely crucial for a successful accomplishment of my PhD project. I enjoyed and learned a lot from our discussions!

My deep appreciation is also addressed to Prof. Dr. Martin Schlee and Anna Herzner, who first introduced me to an exciting world of structured DNA species. Thank you for your great ideas and motivating enthusiasm!

Furthermore, I thank my colleagues Björn Hiller and Andreas Maiser for their optimism, productiveness and enormous time that they spent for the cell-based and microscopy experiments, despite their busyness with their own projects.

I highly appreciate the tremendous help from Dirk Kostrewa and want to thank him not only for building the cGAS structure from difficult diffraction data under the time pressure, but also for his patience and kindness in answering my questions and correcting my mistakes.

I would also like to thank my talented and motivated colleagues Charlotte Lässig, Carina Mann and David Drexler, who put so much effort, time and engagement to help me with my research. Thank you for staying optimistic and interested through the stressful times and especially for your patience and your kind understanding of my emotional involvement!

My special thanks are also addressed to Agata Butryn who offered her help in need, supported me in the first structure evaluation and speeded up my PhD. Thank you for your professionalism and sympathy!

Additionally, I want to thank Siret Somarokov and Gabriele Stöhr for their interest and creativity. I also highly appreciate the input from Gregor Witte and Katja Lammens that helped me with crystallization trials, as well as the support from Claudia Isakaj and Olga Fetscher in urgent and time-consuming experiments. I thank very much Moritz Gaidt and Stefan Bauernfried for their amazing cell lines, which were absolutely indispensable for my work.

Moreover, my deepest gratitude is addressed to Tobias Deimling who took care of me at the beginning of my PhD, taught me all the basics in the lab, supported my curiosity and answered my countless questions. I also sincerely thank Martina Müller for her kind support in the lab and during the synchrotron trips.

I also express my admiration and thankfulness to the whole Hopfner lab members for their help in the lab, fruitful scientific discussions, as well as for emotional support and involvement.

I would also like to thank my dearest friends and family that are always with me, despite the long distances between us. Your love, trust and faith are keeping me alive and are the biggest treasure I was blessed to have.