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

Recognition of cytosolic nucleotides by the innate immune system: Structural view on RIG-I and the STING pathway

Tobias Sebastian Deimling

aus

Starnberg, Deutschland

2014

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung 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 14.08.2014

Tobias Sebastian Deimling

Dissertation eingereicht am	14.08.2014
1. Gutachter:	Prof. Dr. Karl-Peter Hopfner
2. Gutachter:	Prof. Dr. Karl-Klaus Conzelmann
Mündliche Prüfung am	02.10.2014

This thesis has been prepared from November 2010 to October 2014 in the laboratory of Professor Dr. Karl-Peter Hopfner at the Gene Center of the Ludwig-Maximilians-University of Munich (LMU).

It is a cumulative thesis based on following publications:

Civril F, Bennett M, Moldt M, Deimling T, Witte G, Schiesser S, Carell T, Hopfner KP (2011) The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling. *EMBO Rep* **12**(11):1127-34

Cavlar T*, Deimling T*, Ablasser A, Hopfner KP, Hornung V (2013) Species-specific detection of the antiviral small-molecule compound CMA by STING. *EMBO J* **32**(10):1440-50

Ablasser A, Goldeck M, Cavlar T, Deimling T, Witte G, Röhl I, Hopfner KP, Ludwig J, Hornung V (2013) cGAS produces a 2'-5'-linked cyclic dinucleotide second messenger that activates STING. *Nature* **498**(7454):380-4

Civril F*, Deimling T*, de Oliveira Mann C, Ablasser A, Moldt M, Witte G, Hornung V, Hopfner KP (2013) Structural mechanism of cytosolic DNA sensing by cGAS. *Nature* **498**(7454):332-7

Deimling T, Cui S, Lammens K, Hopfner KP, Witte G (2014) Crystal and solution structure of the human RIG-I SF2 domain. *Acta Cryst*. F**70**(8):1027-1031

*: equal contribution

Table of contents

1. Introduction

1.1. Sensing by the innate immune system

1.1.1. Pattern Recognition Receptors (PRRs)

After a pathogen managed to bypass the physical and chemical barriers of the body the innate immune system is the next level of defense. It is an evolutionary ancient system found in all animal and plant species that is composed of germline encoded genes responsible for fighting infections. These innate immune genes form the complete immune system in most species except for vertebrates. The vertebrate immune system comprises the germline encoded factors of the innate immune system and additional factors belonging to the adaptive immune system. The latter evolves with the infection leading to highly pathogen specific defense mechanisms, e.g. formation of antibodies. There is a smooth transition between the anti-infectious actions of the innate and the adaptive immune system. Both, the innate and the adaptive immune system, act in concert and complement each other in clearing the host from the pathogen. Thereby, the innate immune system reacts very fast upon infection but in a rather generic way while the adaptive immune system is activated through the innate immune system and reacts relatively late after infection but very specific. In addition, the adaptive immune system is responsible for the generation of a long term immunological memory [1, 2].

The basis for triggering an appropriate immune response lies in the ability to recognize the presence of a pathogen. The innate immune system is responsible for this recognition process. It provides germline encoded receptor proteins, termed pattern recognition receptors (PRRs) that evolved to sense special pathogen derived and host signature molecules during infection. Each PRR is specialized to bind to one specific signature molecule or class of signature molecules termed microbial/pathogen associated molecular pattern (MAMP/PAMP) or damage associated molecular pattern (DAMP). After MAMP/DAMP binding, PRRs activate signaling cascades that induce the immune response [3, 4].

MAMPs comprise a variety of different kinds of pathogen derived molecules, which are ether structurally distinct from any host molecule, e.g. lipopolysaccharide (LPS, bacterial cell wall component) or appear in abnormal locations, e.g. cytosolic DNA originating from DNA viruses or bacterial genomes. Since their structures usually are essential for the survival of the pathogen, MAMPs show just a very small degree of structural variance. This links the observed high evolutionary conservation of each MAMP and its presence in a broad range of pathogens with the circumstance that many different infections can be sensed by just a small set of PRRs. A prominent example for this observation is the conserved lipid A, which is a component of the outer cell wall of Gram-negative bacteria. The PRR Toll-like receptor (TLR) 4 recognizes lipid A and thereby is able to sense the presence of any Gram-negative bacteria infecting the host [5]. The evolutionary selected specificity of PRRs for MAMPs furthermore contributes to the ability of the immune system to distinguish between self and non-self structures. This is important to avoid an immune response against the own body. In fact, several autoimmune diseases are linked to malfunctioning PRRs, e.g. TLR2/4 play a role in type 1 diabetes [6] or the NOD2 role in Crohn’s disease [7-9].

Like MAMPs, DAMPs are also recognized by PRRs. They also comprise a variety of different kinds of molecules but in this case originating from the host. Generally, these molecules are released from stressed or damaged cells, thereby losing their normal function in the context of a “healthy” cell and contributing to the induction of inflammation as a danger signal. A prominent example is the High-mobility group protein B1 (HMGB1). It is a nuclear protein involved in chromatin organization and transcription regulation. Upon endotoxic shock it is released by hematopoietic cells and can be sensed e.g. by the PRR RAGE (Receptor for Advanced Glycation Endproducts) leading to the induction of an inflammatory response [10-12].

Several protein families act as PRRs and even though they can be structurally very different they often share a common concept of initiating signaling. After binding to the corresponding MAMP or DAMP, PRR signaling domains interact with downstream adaptor proteins resulting in signaling cascades that activate terminal effector proteins like caspases and transcription factors. This leads to the expression, maturation and secretion of proinflammatory cytokines, chemokines and type I interferons (IFNs). These molecules in turn trigger a second wave of autocrine and paracrine signaling resulting in the induction of the immune response.

1.1.2. Toll-like receptors (TLRs)

The best studied PRRs belong to the mammalian Toll-like receptor (TLR) family. They were first identified in the 1990s based on the homology to the *Drosophila* Toll protein which plays a role in establishing the embryonic dorsal-ventral pattern [13, 14]. The 10 TLRs known in humans (12 in mice) are mainly expressed in dendric cells (DCs) and monocytes/macrophages. All TLRs are type I transmembrane proteins which are found either in the plasmamembrane (TLR1, TLR2, TLR4, TLR5, TLR6 and TLR11) or the endosomes (TLR3, TLR7, TLR8, TLR9 and TLR10). Both of these groups sense a broad range of different MAMPs and DAMPs. The plasmamembrane bound TLRs mainly sense extracellular bacterial and viral lipids and Lipoproteins while the endosomal TLRs recognize luminal appearing nucleic acids. Upon MAMP/DAMP binding via their extracellular/endosome lumenal leucin-rich repeats (LRRs) TLRs undergo homo- and heterodimerization depending on the bound substrate. The signal is then transduced to the inside of the cell by the cytoplasmatic TIR (Toll/Interleukin 1 Receptor) domains that interact with TIR domains of downstream adaptors. This activates the NF-κB (Nuclear Factor Kappa-Light-Chain-Enhancer of Activated B Cells), MAP (Mitogen-Activated Protein) kinase and IRF3 (Interferon Regulatory Factor 3) pathways, respectively, involving complex signaling cascades including ubiquitinylation, proteolysis and phosphorylation. The resulting production and release of proinflammatory cytokines and type I interferons induces the immune response [15-17].

1.1.3. NOD-like receptors (NLRs)

TLRs are considered to be the most important PRRs for recognizing infections. But because of their localization in the plasmamembrane and endosomes TLR sensing is restricted to extracellular pathogens. Since many pathogenic processes such as viral replication take place inside the cell, intracellular sensing mechanisms are also essential for triggering an immune response. Indeed, several different cytosolic PRR families crucially contribute to pathogen recognition inside the cell. An example is the NOD-like receptor (NLR) family that comprises at least 23 members in humans.

They are more diverse in structure but typically they possess either N-terminal CARDs (Caspase Activation and Recruitment Domain) or pyrin domains (PYD) followed by a central nucleotide-binding and oligomerization domain (NOD; also known as NACHT domain) and C-terminal leucin-rich repeats (LRRs). Similar to TLRs, the NLR LRRs are considered to be involved in MAMP/DAMP sensing. The CARDs and PYDs are responsible for transducing the signal via protein-protein interactions while the central NOD/NACHT domain seems to play a role in ATP-dependent oligomerization. Well described examples of NLR family members are NOD2 (Nucleotide-Binding and Oligomerization Domain Containing 2) and NLRC3 (NOD-Like Receptor Family CARD Containing 3). NOD2 is involved in the sensing of cytosolic peptidoglycans (PGs) which are bacterial cell wall components. Binding to PGs is suggested to induce oligomerization and interaction with downstream adaptors, eventually leading to the activation of the NF-κB and MAP kinase pathways. NLRC3 is activated by various different stimuli like potassium efflux and the generation of reactive oxygen species (ROS) but the exact sensing mechanism remains to be clarified. Furthermore, NLRC3 activation does not stimulate the transcription of cytokines but leads to the formation of a high molecular weight multi-protein signaling complex called inflammasome. Inflammasomes are composed of higher order oligomers of the corresponding NLRs, thereby forming a platform for the recruitment and activation of caspase 1 that by proteolytic cleavage contributes to the maturation of cytokines like interleukin 1β (IL1β) and interleukin 18 (IL18). Many other NLRs such as NLRP1 or NLRC4 also form inflammasomes [18-20].

1.2. The RIG-I signaling pathway senses cytosolic RNAs

It has been known for more than 50 years that exogenous added nucleic acids such as DNA and RNA have immunostimulatory effects on cells [21, 22]. In the case of RNA, several TLRs were identified as important PRRs. An example is TLR3, which was found to be involved in the sensing of long dsRNAs by stimulating cells with the synthetic analog polyinosinic-polycytidylic acid (poly(I:C)) [23]. Additionally, TLR7 and TLR8 were shown to play roles in the sensing of ssRNA in a species specific manner [17, 24]. However, due to their endosomal localization TLR3/7/8 are restricted to recognize only extracellular RNAs. Furthermore, TLR expression is limited to just a few cell types of the immune system like dendric cells (DCs). These features, in combination with the observation that RNA viruses induce type I IFN production in cell lines lacking TLR3/7/8, suggested the existence of a cytosolic RNA recognition pathway independent of TLR signaling [25]. Indeed, the discovery of the RIG-I-like receptors (RLR) confirmed this assumption. The RLR family comprises three members: The name giving RIG-I (Retinoic Acid Inducible Gene-I), MDA5 (Melanoma Differentiation Associated Gene 5) and LGP2 (Laboratory of Genetics and Physiology 2). RIG-I and MDA5 are cytosolic RNA sensors responsible for recognizing RNA virus infections and transmitting the signal to downstream factors of the RLR pathway, whereas the role of LGP2 is not fully resolved yet. It is supposed to act as modulator of the RLR signaling dependent on the infection [26].

1.2.1. The RIG-I-like receptor (RLR) family

RIG-I and MDA5 were both identified in cancer cells. RIG-I was first described in 1997 as a protein that is expressed after stimulating acute promyelocytic leukemia cells with retinoic acid [27] while MDA5 was originally described as an IFN inducible protein involved in differentiation and growth of melanoma cells [28]. In 2004, both proteins were connected to the innate immune system as RNA helicases that are essential for the induction of IFN production during infection with RNA viruses [29, 30]. Various following studies described RIG-I and MDA5 as direct sensors for viral RNAs in the cytosol [15, 16, 24, 26, 31-33]. Reports about the last member of the

RLR family, LGP2, are inconsistent. LGP2 was identified as negative regulator of RIG-I-dependent IFN production [34], however, later studies described it as positive regulator of RLR signaling [35, 36].

All three proteins are closely related and form a subfamily within the superfamily 2 (SF2) of type DExD/H-box ATPases. The central SF2 type DECH-box domain is composed of two RecA (Recombinase A) folds (Hel1 and Hel2) and a unique insertion domain (Hel2i) in the second RecA fold. The SF2 domain is involved in RNA binding and responsible for the ATPase activity. At the C-terminus, the bridging domain links the SF2 domain to the regulatory domain (RD). The RD is a small flat domain which confers part of the ligand specificity. Major differences between the RLR members in respect to the domain architecture are found at the N-terminus. Both, RIG-I and MDA5 contain two CARDs at the N-terminus which mediate downstream signaling. However, the LGP2 protein, which is thought to play a regulatory role lacks these two CARDs (Fig.1.2.1).

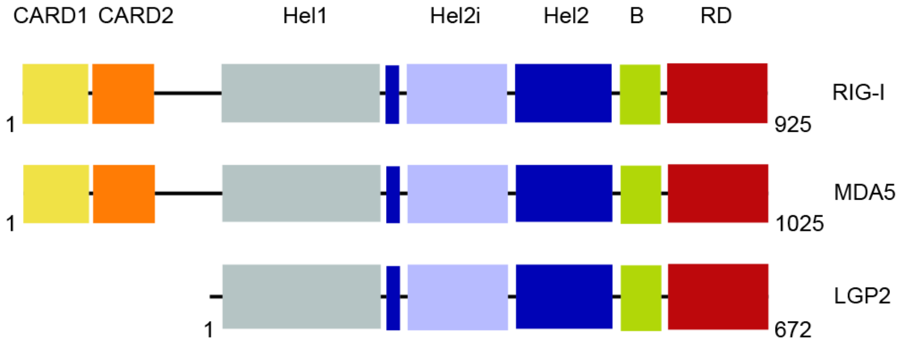


Fig.1.2.1. Domain architecture of RIG-I, MDA5 and LGP2. The SF2 domain is composed of Hel1, Hel2, Hel2i, and the bridging domain (B). At the C-terminus, the bridging domain links the RD to the SF2 domain. RIG-I and MDA5 contain two additional N-terminal CARDs.

1.2.2. RIG-I and MDA5 sensing and signaling

RIG-I and MDA5 recognize mainly distinct viruses. RIG-I senses primarily infections of negative ssRNA viruses such as influenza A virus (*Orthomyxoviridae*), vesicular stomatitis virus (VSV; *Rhabdoviridae*), and rift valley virus (*Bunyaviridae*) but also of some positive-stranded viruses like japanese encephalitis virus (JEV; *Flaviviridae*).

MDA5 on the other hand senses mainly positive ssRNA- and dsRNA viruses such as encephalomyocarditis virus (EMCV; *Picornaviridae*) and norovirus (*Caliciviridae*). However, in some infections, especially of viruses from the *Paramyxoviridae* (e.g. sendai virus (SeV)), *Flaviviridae* (e.g. west nile virus) and *Togaviridae* (e.g. semliki forest virus) families, sensing of both receptors overlaps [31, 33].

The explanation for the different specificities of both receptors for different viruses is reflected by the different RNA features both receptors recognize. RIG-I and MDA5 bind to dsRNA structures but with different length specificities. Experiments with different poly(I:C) constructs showed that RIG-I-dependent production of IFNs is stimulated with short chains (~300 bp) while MDA5 can only be activated with longer chains (>4 kb). These results were further confirmed in context of viral infected cells: VSV, for example, produces short dsRNAs activating only RIG-I-dependent IFN production. Conversely, EMCV infection, which produces long dsRNAs, activates only MDA5-dependent signaling [37].

A further feature that is important for RIG-I signaling is the presence of a 5’-triphosphate end with a short stretch of double-stranded RNA. This structure is directly recognized by the RD of RIG-I. The 5’-triphosphate dsRNA end provides an elegant mechanism of discrimination between self and non-self RNA since it is not found in normal host mRNAs. Nevertheless, the origin of these RNA structures is still under extensive research. Some genomes of negative ssRNA viruses and viral transcripts can fold back into a panhandle structure forming such a 5’-triphosphate dsRNA end. Defective-interfering (DI) RNAs are hairpin-like RNA structures and byproducts of viral replication and transcription which also provide this feature for RIG-I recognition [38-45]. Furthermore, there are host processed self RNAs which can be sensed by RIG-I. RNaseL is an RNA virus infection activated endonuclease that cleaves viral RNAs as well as host RNAs. The cleavage creates short RNAs (~200 nt) with 5’-OH- and 3’-monophosphate ends which activate RIG-I signaling [46, 47]. This expands the range of physiological RIG-I antagonists to viral genomes and transcripts, intermediates and byproducts of viral replication as well as RNaseL cleavage products.

Apart from the tendency to sense long RNAs, there is little information about MDA5 ligands. In fact, long dsRNAs are not sufficient to induce MDA5-dependent signaling. A study showed that rather RNA webs are needed for MDA5 activation. These RNA webs are complex high molecular weight structures containing stretches of ss- and dsRNA suggested to occur during viral replication and transcription [48].

After ligand binding, RIG-I and MDA5 form oligomers, which is a hallmark for the activation of both receptors. In the case of RIG-I the oligomerization depends on its ATPase activity. It is suggested that RIG-I scans along RNA utilizing the SF2 domain as ATP-dependent motor [49]. Binding to the 5’-triphosphate dsRNA end via the RD triggers loading of further RIG-I molecules and the formation of a filament along the RNA [50]. However, filament formation is not just dependent on ATPase activity but also on ubiquitinylation. RIG-I ubiquitinylation is partially mediated by the RING finger E3 ubiquitin ligases TRIM25 (Tripartite Motif-Containing Protein 25) and Riplet and involves covalent linkage as well as non-covalent interactions of K63 linked polyubiquitin chains with the RIG-I CARDs and RD, respectively [51-54]. It is suggested that the polyubiquitin chains provide a mechanism for scaffold formation supporting RIG-I oligomerization even with RNAs that are too short for a proper filament platform [55].

Generally, the activation mechanism of MDA5 is less well characterized than that of RIG-I. Similar to RIG-I, MDA5 forms filaments on dsRNA. In these filaments MDA5 is arranged in a head-to-tail pattern. This MDA5 oligomerization is not dependent on its ATPase activity which implicates a somewhat different activation mechanism compared to RIG-I. Indeed, it is suggested that ATP-hydolysis is involved in the proper alignment of the MDA5 CARDs along the filament for downstream signaling instead of scanning the RNA. Additionally, the MDA5 RD is not involved in the RNA-end sensing but is thought to contribute to MDA5-RNA filament formation by cooperative dsRNA recognition [56-60]. Furthermore, the involvement of ubiquitinylation is still discussed in the field. It has been shown that K63 polyubiquitin chains non-covalently interact with MDA5 CARDs and thereby lead to MDA5 filament formation and activation of signaling [52], whereas other studies failed to show an effect of ubiquitinylation in MDA5 activation [56].

Once activated, the CARDs of both receptor proteins are responsible for transmitting the signal to the downstream factors. For RIG-I, ligand binding is thought to release the CARDs from an autoinhibited state (see below) which contributes to polyubiquitin mediated oligomerization. The CARDs of unstimulated MDA5 are not trapped in an autoinhibibited state, however, their proper arrangement during filament formation is needed for downstream signaling. These signaling competent CARDs in RIG-I and

MDA5 filaments interact with the downstream adaptor MAVS (Mitochondrial Antiviral Signaling; also known as: CARDIF, VISA, IPS1) [61-64]. MAVS is a transmembrane protein predominantly anchored in the outer mitochondrial membrane. It comprises an N-terminal CARD followed by a proline-rich domain and a C-terminal transmembrane domain. The importance of MAVS as component of the RLR pathway was confirmed by knockdown and knockout experiments in several cell-lines and in mice, respectively [65, 66]. The cytosolic CARD domain of MAVS is responsible for homotypic interaction with RIG-I and MDA5 CARDs. Thereby, the oligomeric RIG-I and MDA5 CARDs are thought to function as “seeds” for the oligomerization of MAVS. After induction, MAVS oligomerization proceeds independently of further RIG-I/MDA5 binding leading to the formation of prion-like aggregates on the surface of mitochondria. This way, MAVS provides an activation platform for downstream signaling over the TBK1 and NF-κB axes (Fig.1.2.2.) [67]. This aggregation based mechanism offers a fast and strong amplification of the RLR signal allowing for a rapid response of the immune system [15, 16, 24, 26, 31-33].

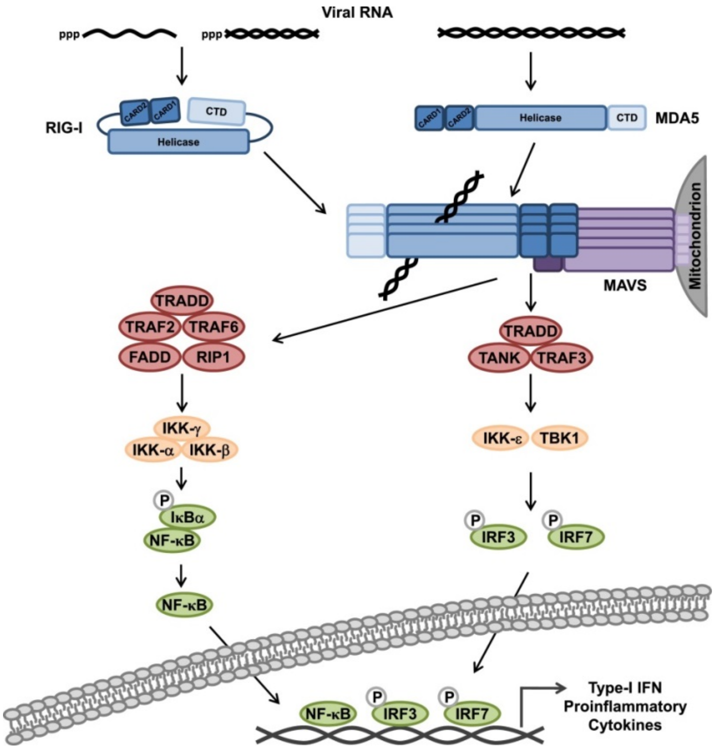


Fig.1.2.2. Schematic overview of the RLR signaling pathway. Binding of viral RNAs to RIG-I and/or MDA5 leads to RLR induced oligomerization of MAVS on the surface of mitochondria. This activates the NF-κB and TBK1 pathways, respectively, and leads to the expression of type I IFNs and proinflammatory cytokines (Figure taken from [68]).

1.2.3. Structural mechanism of RNA sensing by RIG-I and MDA5

A recent series of structural reports analysed the mechanism of RNA sensing and signaling via RIG-I. The first structural information was obtained of the RD in isolation and in complex with 5’-triphosphate- and blunt ended dsRNA, respectively. The RD is a small flat domain comprising a core of antiparallel β-sheets that is stabilized by a conserved C-4-type zinc finger. The 5’-end of dsRNA is bound via several highly conserved residues forming a positively charged and shallow binding groove [69-73]. The C-terminal RD is connected to Hel2 of the SF2 domain via the so called bridging- or pincer domain. It is an elbow-like structure composed of two α-helices interacting with Hel1 and Hel2. This nanomechanical structure transmits RNA binding of the RD to the SF2 domain, thereby connecting ligand recognition and motor function. The SF2 domain of RIG-I has a “C-shaped” structure. As mentioned before, it contains the two RecA-like folds Hel1 and Hel2 as well as the unique insertion Hel2i. Compared to other SF2 proteins RIG-I-Hel1 and -Hel2 are spaced rather far apart forming a relatively “open” ATP-binding cleft [74, 75]. Hel2i is a bundle of six α-helices inserted in the Hel2 fold. Together with Hel1 and Hel2 it shares the RNA binding site in the center of the “C-shaped” SF2 domain [76]. At the N-terminus the two CARDs are connected to Hel1 via a flexible linker.

In the RNA free state RIG-I exists as a monomer in a rather open and extended conformation [77]. It is suggested that the RD is moving freely sensing for ligands while the CARDs are trapped in an autoinhibitory state via direct interactions between the second CARD and Hel2i [78]. Upon ligand recognition by the RD, a dramatic conformational change is triggered mediated by the bridging domain. It involves rearrangements of the whole SF2 domain and a relocalization of the RD resulting in a closed O-ring structure formed around the dsRNA. Thereby, the RIG-I SF2 domain interacts with the RNA only via the sugar-phophate-backbone providing a sequence independent binding mode. In addition, ATPase activity of the SF2 domain leads to different degrees of further ring-compaction. Thereby, binding of ATP results in the highest compaction which decreases with different stages of hydrolysis [79-81]. This ring formation and compaction eventually interrupts the CARD-Hel2i interactions which releases the CARDs making them accessible for ubiquitinylation and competent for oligomerization and downstream signaling (Fig.1.2.3.1.). Even though, the exact role of ATPase activity is not fully resolved yet, this mechanism explains how RIG-I signaling is repressed during the absence of adequate ligands and turned on during viral infection [82].

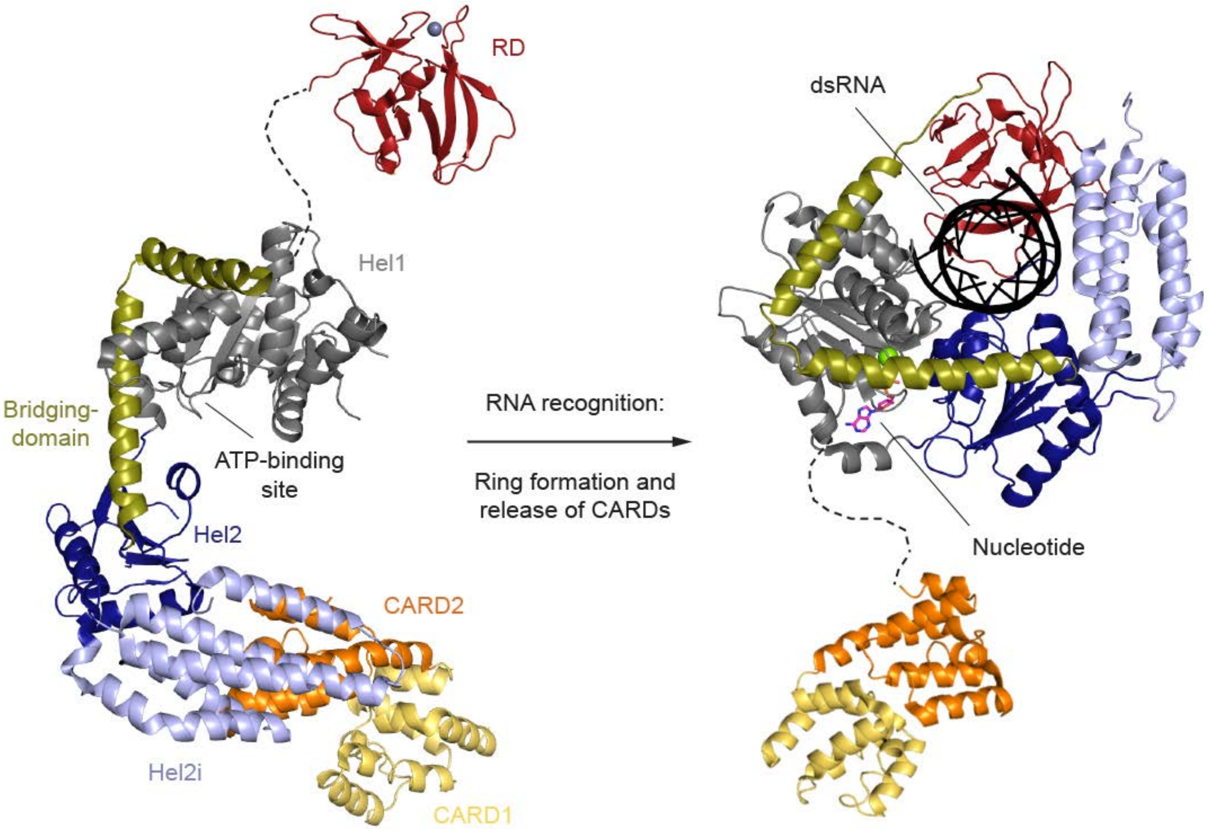


Fig.1.2.3.1. Current model for the activation of RIG-I. Binding of 5'-triphosphate dsRNA to the free moving RD triggers ring formation. This disturbs the autoinhibiting CARD2-Hel2i interactions and releases the CARDs for oligomerization and downstream signaling (RD structure: 2QFB [69]; RIG-I full length and CARDs structure: 4A2W [78]; RIG-I-RNA structure: 3TMI [77]).

In contrast to RIG-I, the mechanism of MDA5 activation is not completely resolved, yet. The crystal structure of the MDA5 SF2-RD in complex with RNA shows an overall similar structure with subtle differences compared to RIG-I. The MDA5-SF2 domain adopts a comparable conformation like in RIG-I but due to a shortened helix (highly conserved) in the MDA5 Hel2i the RD is positioned differently. It does not cap the 5'-end of the RNA like the RIG-I-RD but binds in a slightly different orientation recognizing the dsRNA stem. This leads to an overall more open "C-shaped" complex around the RNA in contrast to the closed O-ring observed with RIG-I (Fig.1.2.3.2.). No crystallographic information about the effect on MDA5-CARDs positioning upon RNA binding is currently available, but electron microscopy data and modeling approaches suggest that the CARDs form patches of signaling-competent oligomers along the head-to-tail filaments of MDA5 [56, 57]. Nevertheless, structural details of filament formation and signaling repression in the absence of MDA5 ligands are still unclear. In addition, the exact role of ATP-hydrolysis in positioning the CARDs for signaling needs to be resolved.

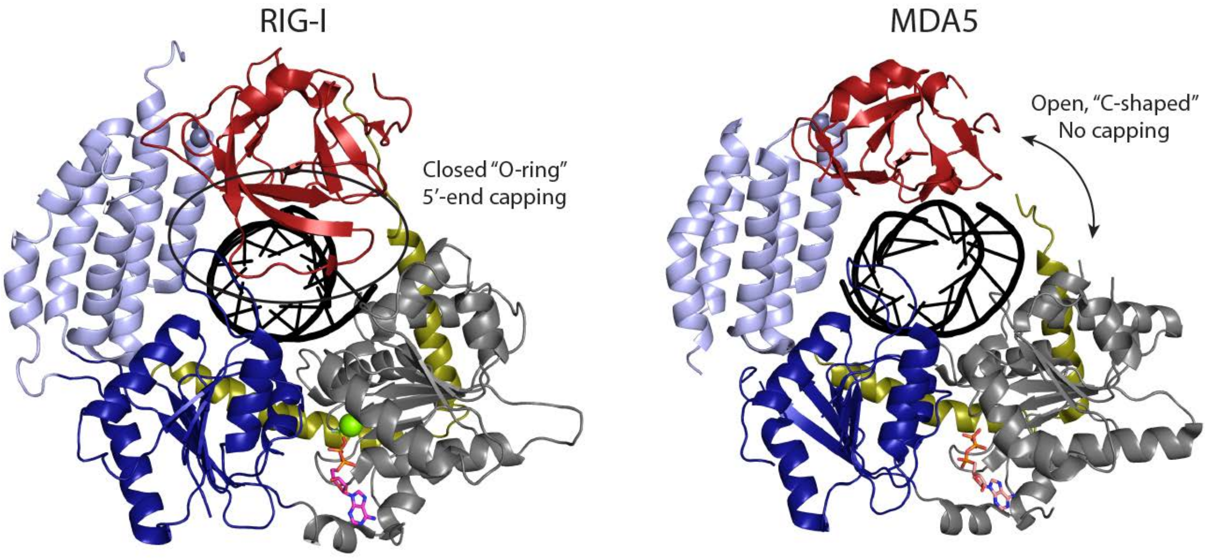


Fig.1.2.3.2. Structural comparison of RIG-I and MDA5 in RNA bound state. RIG-I forms a closed ring around RNA. Additionally, the RIG-I RD caps the 5'-end (RIG-I structure: 3TMI [77]). MDA5 exhibits an open "C-shaped" structure. The RD binds the RNA without capping (MDA5 structure: 4GL2 [56]).

In spite of the high similarities of RIG-I and MDA5, the existing structural data suggest a different activation mechanism for both receptors. This is also supported by different ligand specificities and their involvement in the recognition of mainly distinct viral infections [83].

1.2.4. Further regulation of RLR signaling

Signaling pathways are generally tightly regulated on different levels involving regulation strategies such as activator/inhibitor binding, posttranslational modifications and degradation. This allows for fine tuning, flexibility and adjustment of the signal resulting in an appropriate response corresponding to the stimulus.

Apart from the already mentioned essential polyubiquitinylation, which is involved in oligomerization of RIG-I, other polyubiquitin-dependent regulatory mechanisms contribute to activation of both receptors [84]. The E3 ubiquitin ligases c-Cbl (Casitas B-Lineage Lymphoma Proto-Oncogene) and RNF125 (RING-Finger Protein 125) provide a negative feedback mechanism to avoid overstimulation of the RLR-pathway. Upon viral infection they mark RIG-I and MDA5, respectively, with K48-linked polyubiquitin chains, which targets them for proteasomal degradation [85, 86]. The deubiquitinase USP4 (Ubiquitin-Specific Protease 4), on the other hand, removes K48-linked polyubiquitin chains from RIG-I. Thereby, it acts as positive regulator of RIG-I signaling by preventing its degradation by the proteasome [87]. Another deubiquitinase involved in the regulation of RIG-I is CYLD (Cylindromatosis). This protein was shown to be able to remove K63-linked ubiquitin chains, thereby inhibiting RIG-I activation [88].

Except for ubiquitinylation other posttranslational modifications play a role in the regulation of both receptors. RIG-I CARDs are phosphorylated at S8 and T170 by PKC-α (Protein Kinase C α) or PKC-β (Protein Kinase C β) while MDA5 CARDs are phosphorylated at S88. These phosphorylations inhibit CARD-CARD interactions with downstream adaptor MAVS thereby silencing RLR signaling. However, phosphorylation can be reversed by the two phosphatases PP1α (protein phosphatase 1 α) and PP1γ (protein phosphatase 1 γ), which renders both receptors competent for signaling again [89-91].

Another mechanism of regulation apart from covalent modifications and degradation is the interaction with endogenous activators or inhibitors. An example is the shortened isoform of poly(ADP-ribose) polymerase 13 called ZAPS. This protein was shown to associate with RIG-I and to stimulate its ATPase activity and oligomerization [92]. A further activator of RIG-I signaling is PACT (Protein Kinase, Interferon-Inducible Double-Stranded RNA-Dependent Activator), a dsRNA binding protein that interacts with the RD and promotes RIG-I signaling [93]. An example for an activator of MDA5-dependent signaling is the protein RAVER1 (Ribonucleoprotein PTB-Binding 1). It was shown to bind to MDA5 and enhance its affinity to dsRNA, thereby supporting MDA5 activation [94]. In contrast, DAK (Dihydroaceton Kinase), interacts with MDA5 and prevents binding to dsRNA or downstream factors serving as an inhibitor for MDA5-dependent stimulation of type I IFN production [95].

All these factors display a complex regulatory network that allows for adjustment, fine tuning and flexibility of the RLR signaling.

1.2.5. Viral evasion on the level of RLRs

The evolved immune systems of the diverse species are generally very effective in fighting infections. Nevertheless, some viruses have developed evasion mechanisms that counteract the host immune system and ensure their proliferation. These evasion mechanisms often directly target components of the host immune system including the RLR recognition and signaling pathway.

An example for a mechanism that conceals the viral genome from RIG-I recognition is the processing of the 5'-triphosphate dsRNA end by some viruses (e.g. crimean-congo haemorrhagic fever virus (CCHFV), borna disease virus (BDV), all *Bunjaviridae*). These viruses utilize virus-encoded phosphatases to hydrolyze the 5'-triphosphate to create a 5'-monophosphate end or endonucleases to resect the 5'-strand of the dsRNA end to generate a 3'-overhang, respectively. These structures cannot be recognized by RIG-I and therefore do not trigger RIG-I-dependent IFN responses during infections with those viruses [96-98]. Another example for a concealing mechanism is the capping of the 5'-triphosphate dsRNA end. Members of the *Picornaviridea* and *Caliciviridea* encode the protein VPg, which is covalently linked to the 5'-end of the genome thereby circumventing recognition by RIG-I [99].

Several viruses manipulate posttranslational modifications to suppress immune responses. One example was reported for kaposi's sarcoma-associated herpesvirus (KSHV). Beside others, its genome encodes the protein ORF64. This protein exhibits deubiquitinase activity and is able to decrease the level of K63-linked polyubiquitin chains covalently bound to RIG-I which dampens its signaling [100, 101]. A further example is the V protein from viruses belonging to the family of *Paramyxoviridea* (e.g. measles virus (MV), nipah virus (NiV)). It was shown that the MV- and NiV-V proteins can interact with PP1α/γ and inhibit the dephosphorylation of MDA5 CARDs thereby rendering MDA5 inactive [102].

Interestingly, the V proteins of the *Paramyxoviridea* family suppress the immune response utilizing multiple mechanisms that can differ from virus to virus [103]. Apart from the described inhibition of dephosphorylation the V protein is also an example for a viral factor that can inhibit the ATPase activity of its PRR. A structural study could show how the V protein of parainfluenza virus 5 (PIV5) incorporates itself into the Hel2 fold of MDA5. This binding disrupts the SF2 domain and abolishes ATPase activity providing a further mechanism that prevents MDA5-dependent signaling [104].

Many viruses encode proteases, which are involved in their replication cycle. In addition, several of these proteases also play roles in immune evasion. For example, the 3C^{pro} protease of some members of the *Picornaviridea* family (e.g. poliovirus, rhinovirus, echovirus, Encephalomyocarditisvirus (EMCV)) cleaves and inactivates RIG-I during infection [105]. Another example for a viral protease that is involved in silencing RLR signaling is 2A^{pro}, encoded by enterovirus 71 (EV71). Upon infection it cleaves MDA5 thus inactivating it [106, 107]. However, RLR cleavage does not always rely on direct involvement of viral proteases. Poliovirus infection, for example, was shown to induce apoptosis leading to proteasomal and caspase mediated degradation of MDA5 which contributes to the suppression of an immune response [108, 109].

Finally, many viruses use sequestration mechanism to avoid RLR-dependent signaling. ORF3b and ORF6 are proteins encoded by the severe acute respiratory syndrome coronavirus (SARS-CoV). Both proteins localize at the outer membrane of mitochondria and block the RIG-I-MAVS interaction by binding to either MAVS or RIG-I [110, 111]. In infections with respiratory syncytial virus (RSV) the formation of inclusion bodies (IB) inside the cell is observed. This formation is dependent on the viral proteins N and P. During infection relocalization of RIG-I, MDA5 and MAVS to these IB is accompanied by a significant decrease of IFN production [112].

During evolution viruses have developed various strategies to evade the host immune system. In the case of RLR recognition and signaling this involves mechanisms like concealing the viral genome and transcripts, manipulation of posttranslational modifications, inhibition of the ATPase activity, proteolytic cleavage and sequestration as displayed by the presented examples [31, 68].

1.2.6. Further cytosolic RNA sensors

Apart from RLRs, other proteins act as cytosolic RNA sensors. PKR (dsRNA-Dependent Protein Kinase R) was first described in 1976 [113]. It is a serine/threonine kinase that is activated by binding to cytosolic dsRNAs. Activated PKR phosphorylates eIF2 (Eukaryotic Translation Factor 2), which leads to suppression of translation. During viral infection this mechanism inhibits viral replication by preventing the production of essential viral proteins [114]. In addition, PKR was also implicated in the induction of type I IFN production but the exact mechanism is not fully understood yet [25, 115].

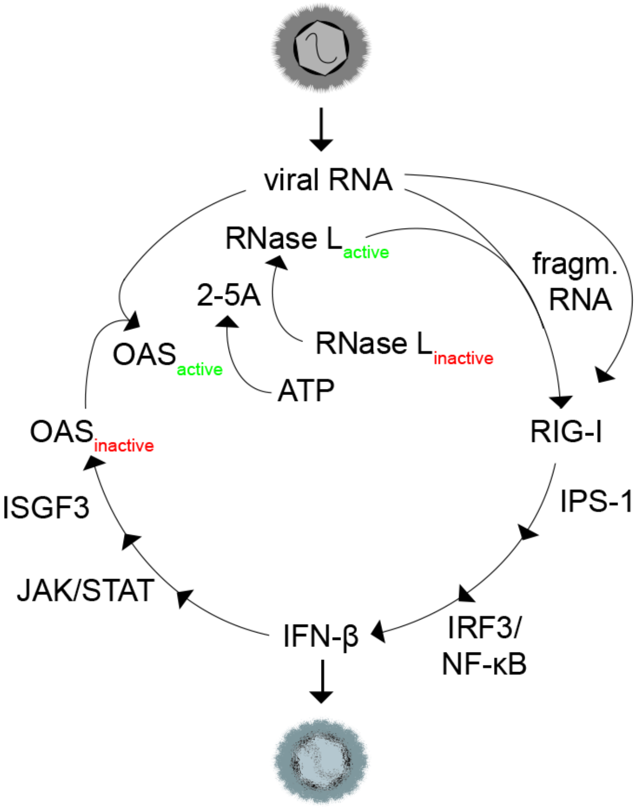


Fig.1.2.6. Positive feedback loop for RIG-I-OAS-RNaseL activation. Cytosolic viral RNAs activate the RIG-I pathway leading to the production of type I IFNs. The IFN signal is recognized and transduced by the JAK/STAT system and the expression of members of the OAS family is induced. Following the activation by cytosolic viral RNAs OAS members produce 2’-5’ oligoadenylates. These second messengers activate the nuclease activity of RNaseL and the production of further RIG-I ligands (Figure taken with minor changes from [47]).

Another protein family involved in sensing of cytosolic dsRNAs is the OAS (2’-5’-Oligoadenylate Synthase) family. It consists of four members (OAS1, OAS2, OAS3 and OASL) and forms an own subfamily in the superfamily of NTases [116]. With the exception of OASL all members act as cytosolic dsRNA sensors. Upon viral infection production of OAS1/2/3 is induced over the JAK/STAT (Januskinase/Signal Transducers and Activators of Transcription) pathway in an IFN-dependent manner. In the cytosol, binding to dsRNA triggers activation of OAS1/2/3 and the production of 2’-5’ oligoadenylates from ATP with lengths of two to over ten nucleotides. The adenosines in these chains are linked via unique 2’-5’-diphosphoester bonds. 2’-5’ oligoadenylates act as second messenger molecules and activate the downstream target RNaseL. As mentioned before, RNaseL cleaves ssRNAs, thereby producing ligands for RIG-I. This mechanism connects RLR signaling with OAS sensing and establishes a positive feedback loop that leads to a rapid and robust IFN signal (Fig.1.2.6.) [46, 47, 117-120].

1.3. The STING signaling pathway senses cytosolic DNA and cyclic dinucleotides

DNA normally resides in the nucleus of eukaryotic cells but during infections pathogen and/or host DNAs can also appear in other compartments, in the cytosol or the extracellular environment via diverse mechanisms (e.g. injection of genome by DNA virus, DNA release by necrotic cells). These abnormally localized DNAs are sensed as danger signals by PRRs which eventually induce the production and release of proinflammatory cytokines and type I IFNs.

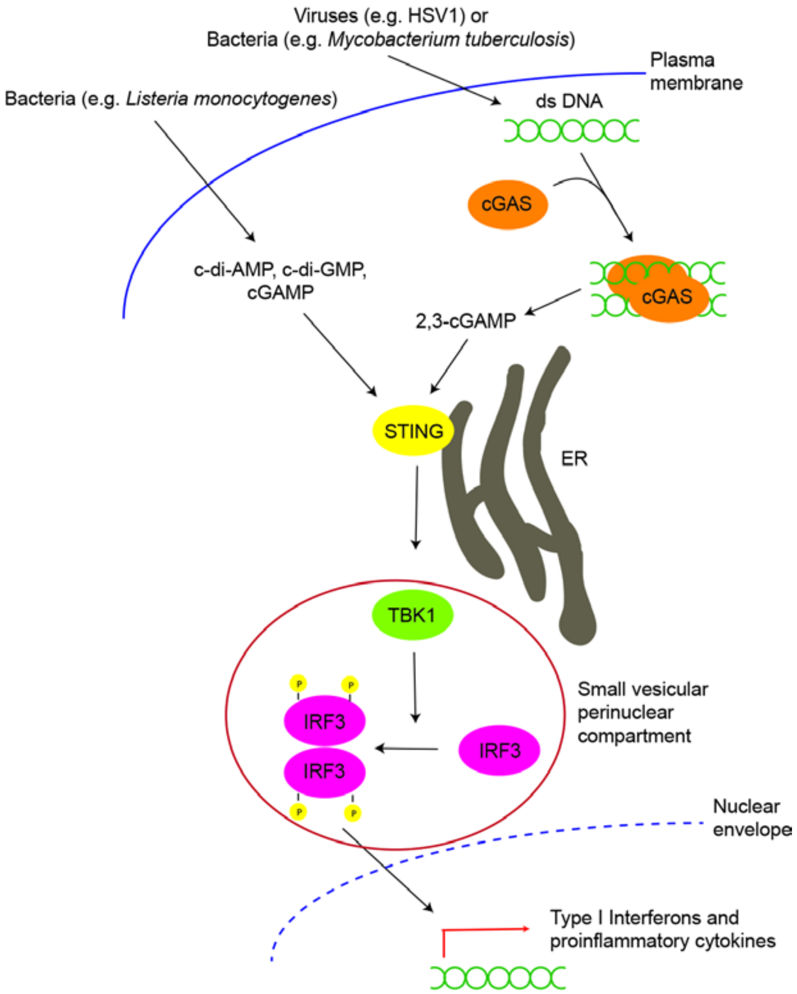


Fig.1.3. STING pathway. Cyclic-dinucleotide sensing branch: STING acts as PRR for bacterial second messenger molecules c-di-AMP, c-di-GMP and cGAMP. DNA sensing branch: cGAS is the PRR for cytosolic dsDNA. Binding activates cGAS leading to the production of the eukaryotic second messenger 2,3-cGAMP. 2,3-cGAMP is recognized by STING. Activated STING translocates to small vesicular perinuclear compartments forming an activation platform for TBK1. Hereupon, TBK1 phosphorylates IRF3, which dimerizes and relocates into the nucleus acting as transcription factor for type I IFNs and proinflammatory cytokines.

The first PRR responsible for inducing type I IFN production upon DNA sensing discovered was TLR9. It particularly senses unmethylated CpG rich DNA sequences that are frequent in bacterial and viral genomes [121]. Exactly like the RNA sensing by TLR3/7/8, DNA sensing by TLR9 is restricted to extracellular DNAs and plasmacytoid dendric cells due to its cell type specific expression and endosomal localization. Experiments with other cell types and TLR9 depleted cells with DNA injected into the cytosol also showed robust type I IFN production demonstrating the existence of a further TLR9-independent cytosolic DNA sensing mechanism [122, 123]. The following years of research finally culminated in the discovery of the STING pathway. It is named after the central adaptor protein STING (Stimulator of Interferon Genes) that forms the DNA sensing branch of this pathway together with the sensor protein cGAS (Cyclic GMP-AMP Synthase) [26, 124] (Fig.1.3.).

1.3.1. STING is the central adaptor protein for the cytosolic DNA signaling pathway

Human STING (ERIS, MITA, TMEM173) is a 42 kDa (379 amino acids) protein consisting of an N-terminal transmembrane domain (TMD; 1-135) followed by a ligand binding domain (LBD; 150-340) and a C-terminal tail (CTT; 340-379) (Fig.1.3.1.). Bioinformatical analysis predicts an unstructured CTT and three to four transmembrane helices in the TMD. The structure of the LBD was solved by several groups and will be discussed below.

Fig.1.3.1. Domain architecture of STING. It consists of a N-terminal transmembrane domain (TMD) followed by the cytosolic ligand binding domain (LBD) and the C-terminal tail (CTT).

On amino acid level, human STING is 81 % similar and 68 % identical to the mouse ortholog (42 kDa, 378 amino acids). In fact, STING orthologs/homologs are found in diverse vertebrate (e.g. *Gallus gallus*, *Danio rerio*, *Xenopus tropicalis*) and invertebrate species (e.g. *Drosophila melanogaster*, *Nasonia vitripennis*, *Hydra magnipapillata*).

In 2008, STING was first identified as a molecule interacting with the major histocompatibility complex II, though the relevance of this discovery remains to be clarified [125]. At the same time, by using cDNA overexpression screens combined with IFN-β promoter activation readout, three other groups independently described STING as a key innate immune signaling protein involved in the sensing of cytosolic nucleic-acids [126-128]. Following studies confirmed the importance of STING for the induction of a type I IFN response. Experiments in HEK293T cells, which do not express endogenous STING, continuously produce IFN-β upon STING overexpression [126]. Vice versa, STING knockdown experiments in cell types endogenously expressing STING (e.g. DCs, bone marrow derived macrophages (BMDM), murine embryonic fibroblasts (MEFs)) showed abolished IFN-β production upon infection with DNA viruses (e.g. herpes simplex virus 1 (HSV-1)) or stimulation with dsDNAs (e.g. interferon stimulatory DNA (ISD)). The high susceptibility to HSV-1 virus of STING knockout mice furthermore demonstrated the *in vivo* importance of STING in fighting DNA virus infections [129].

STING is particularly expressed in tissues connected with the immune system like lung, peripheral leukocytes, placenta, spleen and thymus. Low expression is observed in the brain, colon, kidney, liver, skeletal muscles and small intestines. In expressing cells, it is anchored via its TMD in the endoplasmic reticulum (ER) localized partially in regions associated with mitochondrial membranes [126, 127, 129]. These structures are referred to as mitochondria-associated ER membranes (MAMs) and are known to play an important role for the oxidative metabolism and the Ca^{2+} transmission of mitochondria [130]. Some studies reported an interaction between STING and components of the RIG-I pathway (MAVS, RIG-I) in these structures and an involvement of STING in the signaling of cytosolic RNA [126-129]. However, later studies could not confirm a role of STING in the cytosolic RNA induced type I IFN production [131, 132].

Inactive STING has been described to occur in different oligomeric states. One study demonstrated that inactive STING resides as monomer in the ER. Dimerization via the cytosolic domain of STING eventually triggers interferon expression [127]. In contrast, other studies showed that the cytosolic domain of STING is already dimeric in solution suggesting a dimeric inactive state [133-139].

While the inactive form of STING is still unclear, several studies describe the activation of STING. Upon virus infection or stimulation with dsDNAs STING relocates from the ER over the Golgi apparatus to ER connected punctuate perinuclear vesicular compartments. The exact nature of these structures is still vague but confocal fluorescence microscopy assays suggest the formation of higher order STING oligomers. Interestingly, STING colocalizes with the TANK binding kinase 1 (TBK1) which plays an important role in innate immune signaling [129, 140]. Additionally, co-immunoprecipitation data show an interaction between STING and TBK1 after immunostimulation. Furthermore, the recruitment and activation of transcription factor IRF3 (Interferon Regulatory Factor 3) is observed. In combination, these data establish the model in which STING forms an activation platform for TBK1 and IRF3 leading to interferon expression via TBK1/IRF3 [141].

In addition to this, STING-dependent activation of TBK1 leads to activation of STAT6 (Signal Transducer and Activator of Transcription 6) and following expression of certain STAT6-dependent cytokines. This virus infection induced STAT6 activation differs from the interleukin-dependent canonical activation and suggests the existence of a new virus infection signaling pathway [132].

1.3.2. STING is a direct sensor for cyclic dinucleotides

Cyclic dinucleotides, e.g. cyclic-di-AMP, cyclic-di-GMP or cyclic-AMP-GMP (Fig.1.3.2.) are bacterial second messenger molecules. They are found in a wide range of diverse bacteria where they are involved in several different cellular processes like bacterial motility, biofilm formation, DNA integrity sensing and virulence [142-144]. In eukaryotic cells the presence of cyclic dinucleotides triggers the production of type I IFNs [145-147]. Because of their essential roles in bacteria and their wide distribution among the diverse bacterial species, cyclic dinucleotides are the ideal MAMPs for the recognition of bacterial infections by the eukaryotic immune system.

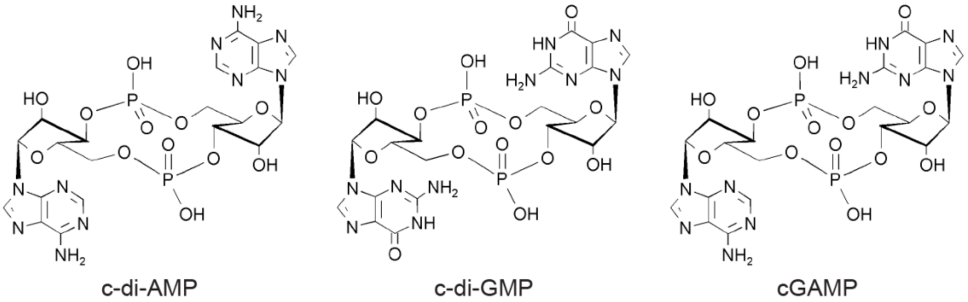


Fig.1.2.3. Cyclic dinucleotides. Examples for known bacterial second messengers: cyclic-di-AMP (c-di-AMP), cyclic-di-GMP (c-di-GMP) and cyclic-GMP-AMP (cGAMP).

Recent research has shown the crucial involvement of the STING pathway in sensing cyclic dinucleotides and the resulting induction of type I IFN expression [131, 148]. Biochemical studies were able to show direct binding of these dinucleotides to STING by using UV-crosslinking assays [149]. Dissociation constants of c-di-GMP bound to STING determined by isothermal titration calorimetry (ITC) are in the range of K_d = 2.5-5 µM and several co-crystal structures of the STING-LBD with bound c-di-GMP and cGAMP could be solved (see below) [133-138, 150].

1.3.3. The DNA sensor cGAS produces 2,3-cGAMP

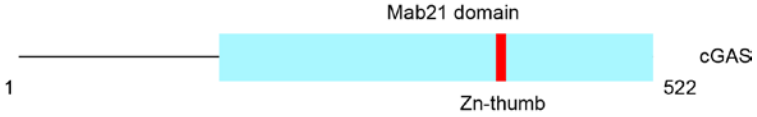


Fig.1.3.3.1. Domain architecture of cGAS. The N-terminus is predicted to be unstructured. The C-terminus consists of a Mab21 domain carrying an unique Zn-finger insertion called Zn-thumb.

Human cGAS (MB21D1, c6orf150) is a 59 kDa (522 amino acids) protein. It consists of an unstructured and poorly conserved N-terminus (1-160) followed by a Mab21 domain (161-522) carrying a unique zinc-finger insertion (Zn-thumb; 390-405; Fig.1.3.3.1.). The Mab21 domain belongs to the family of nucleotidyltransferases (NTases) including the cytosolic RNA sensor protein OAS1 (2,5-Oligoadenylate Synthase 1) and pre-mRNA modifying enzyme PAP (Polyadenylate Polymerase) [116]. In vertebrates the cGAS Mab21 domain is well conserved (e.g. *sus scrofa*, *mus musculus*) but in invertebrates it is more homologous to the Mab21 domain from the human MB21L1 protein.

cGAS was discovered in 2011 as a protein that suppresses viral replication upon overexpression [151]. The exact role of cGAS was unclear until the end of the year 2012. In two publications the Chen group describes elegant *in vivo* and *in vitro* experiments including elaborated transfection assays and biochemical purification procedures identifying cGAS and its role as sensor in the STING-dependent cytosolic DNA signaling pathway. The authors of the two studies could show that upon direct DNA binding cGAS utilizes ATP and GTP to produce a non-canonical cyclic dinucleotide, which is bound by STING thereby activating IFN-β expression [152, 153]. This discovery of a new second messenger synthesized by a cytosolic DNA sensor connects DNA sensing and the recognition of bacterial cyclic dinucleotides by the STING pathway. The exact nature of the cyclic dinucleotide was resolved by several following studies using ribonuclease digestion assays combined with small molecule NMR and mass spectrometry or protein crystallography. It could be shown that the activated cGAS protein synthesizes a completely new kind of eukaryotic second messenger, namely cyclic 2,5-GMP-3,5-AMP (2,3-cGAMP). This cyclic dinucleotide differs from the canonical bacterial cGAMP at the phosphodiester linkage between the 2'-position of GMP and the 5'-position of AMP. It is probably formed in a two-step mechanism which comprises the initial formation of a linear intermediate pppGMP-2'-p-5'-AMP and a final cyclization reaction to yield the product. Each step is accompanied by pyrophosphate release (Fig.1.3.3.2.) [154-156].

Other studies support the importance of cGAS and 2,3-cGAMP for the STING pathway. Overexpression of cGAS activates STING-dependent type I IFN production while knockdown and knockout of cGAS abolishes STING-dependent type I IFN expression upon HSV-1 infection or dsDNA treatment in diverse cells types. cGAS knockout mice furthermore were highly susceptible to HSV-1 infections compared to wild type mice [157].

Noteworthy, 2,3-cGAMP has been shown to spread from producing cells via gap junctions to neighboring cells inducing STING-dependent type I IFN production in these cells. This small-molecule mediated communication between neighboring cells is ideal for initiating a further level of defense against infections and is able to prime neighboring cells [158].

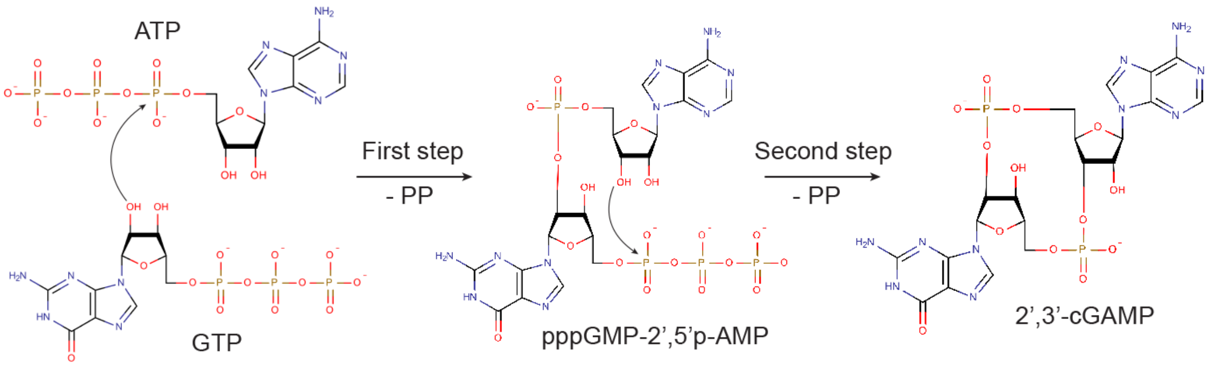


Fig.1.3.3.2. Reaction mechanism of 2,3-cGAMP production. In the first step, the nucleophilic attack of the 2’-OH of GTP against the α-phosphate of ATP is catalyzed. Pyrophosphate released yields the linear intermediate pppGMP-2’-p-5’-. In a second step, the 3’-OH of the adenosine moiety attacks the α-phosphate of the guanosine to complete the cyclization reaction to 2,3-cGAMP.

1.3.4. Structural studies of cGAS

The poorly conserved unstructured N-terminus of cGAS has been shown to be not involved in cytosolic DNA signaling. In contrast, the Mab21 domain harboring the DNA binding and the catalytic site is essential [152]. The fact that cGAS has been identified to be the cytosolic DNA sensor in combination with the Mab21 domain synthesizing a new second messenger molecule led to several publications describing crystal structures of this domain with and without DNA/nucleotides immediately after [156, 159-163].

The DNA and nucleotide free (apo-) form comprises two lobes which are separated by the active site cleft (Fig.1.3.4.). The N-terminal lobe contains the NTase fold (two-leaved highly twisted β-sheet flanked by two helices) while the C-terminal lobe is composed of a bundle of four α-helices. Opposing the active site cleft the two lobes are connected by a long α-helix (spine-helix) which flanks the slightly concave highly positively charged surface (platform). The other side of the surface is flanked by a protruding Zn-binding loop (Zn-thumb).

The crystal structure of the cGAS-Mab21 domain bound to dsDNA shows a 2:2 protein:DNA complex (Fig.1.3.4.). The DNA is positioned along two highly positively charged binding sites and is stabilized by side chain and main chain contacts to the sugar-phosphate-backbone and intrusion of arginines into the minor groove. DNA binding furthermore induces several conformational changes including a nicking of the spine helix leading to reordering and slight closure of the active site. Apart from DNA, several structures contained either the nucleotide substrates ATP and GTP, different intermediates or the end product 2,3-cGAMP in the active site.

The co-crystal structures of cGAS with different nucleotides mechanistically support the two step reaction mechanism suggested for the formation of 2,3-cGAMP. However, the complete reaction cycle is not yet fully resolved, as from the sterical point of view the linear reaction intermedieates would have to bind, reorient and rebind in order to be able to obtain the 2,3-cGAMP product [164, 165]. The observed DNA binding mode is sequence-independent and explains the role of cGAS as a broad range DNA sensor.

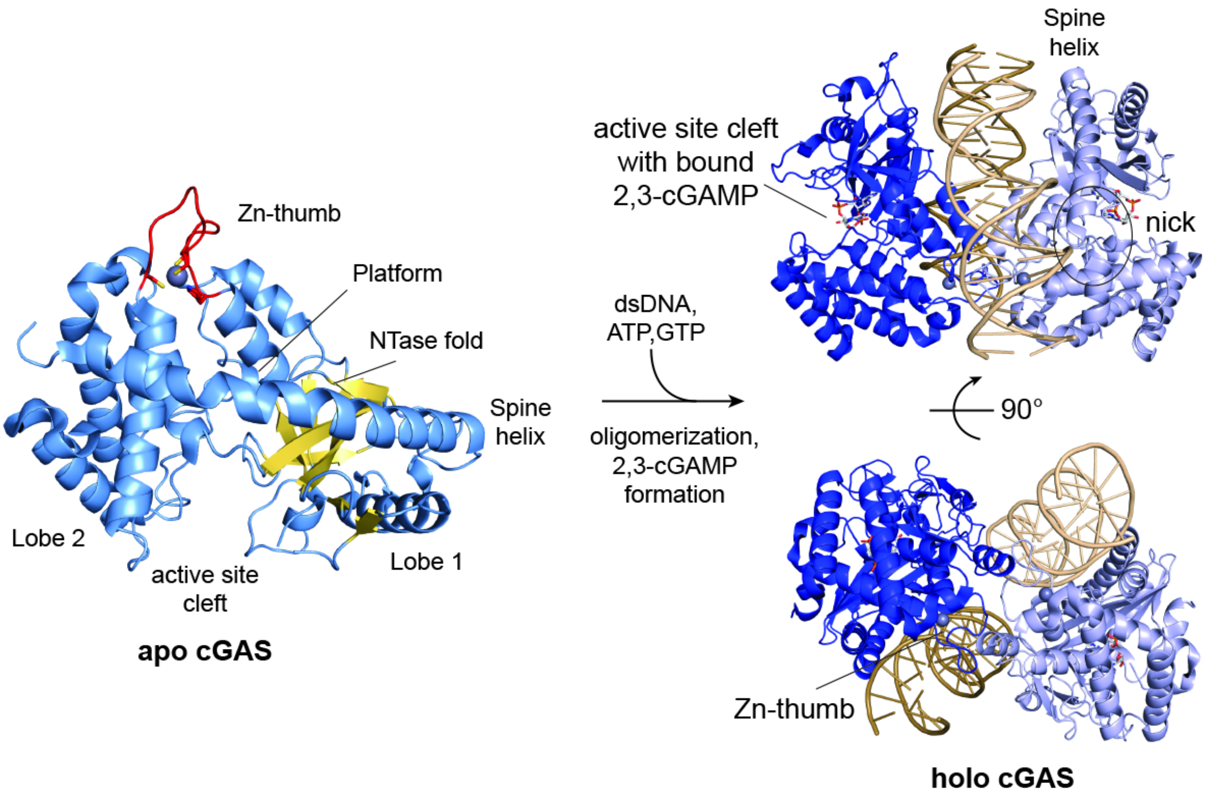


Fig.1.3.4. Activation of cGAS. Apo cGAS consists of lobe 1 and lobe 2. Lobe 1 contains the NTase fold (yellow). Both lobes are separated by the active site cleft. Opposing the active site cleft the platform is positioned, flanked by the spine helix on one side and by the Zn-thumb on the other side. Upon substrate binding cGAS forms a 2:2 protein:DNA complex. This leads to the nicking of the spine helix and a rearrangement of the active site inducing the formation of 2,3-cGAMP from ATP and GTP (apo cGAS structure: 4JLX [159], holo cGAS structure: 4LEZ [162]).

1.3.5. Structural studies of STING

Several groups were able to solve the structure of the STING LBD which is involved in dimerization and responsible for cyclic dinucleotid binding [133-139, 150, 166]. Ligand free (apo-) LBD crystallized as a homodimer forming a V-shaped structure. Each protomer possesses a unique α/β-fold containing a central twisted five stranded β-sheet, flanked by three α-helices at the dimer interface side and two α-helices at the opposing surface side (Fig.1.3.5.). In the holoprotein, 2,3-cGAMP is bound in a cavity at the bottom of the dimer interface stabilized by π-π stacking interactions and hydrogen bonds. The binding of 2,3-cGAMP mediates the formation of a lid which is composed of a four stranded antiparallel β-sheet (two strands of each protomer) that closes the ligand binding site. Apart from the lid formation, 2,3-cGAMP binding leads to a general compaction of approximately ~20 Å compared to the apoprotein (Fig.1.3.5.) [150, 166]. This compaction and the formation of the lid are suggested to contribute to the activation of STING but the exact mechanism is still unclear.

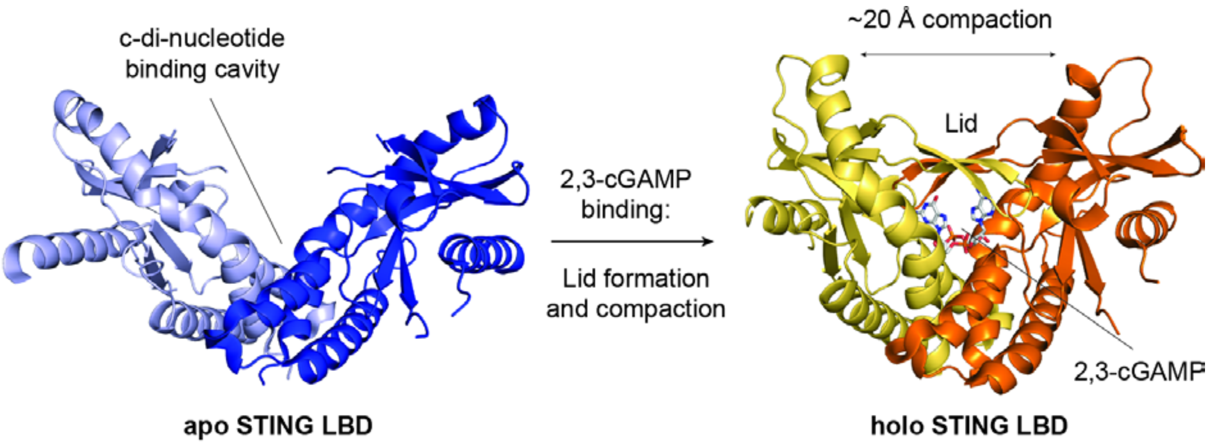


Fig.1.3.5. Ligand binding mechanism of STING. Apo STING LBD is a V-shaped dimer. 2,3-cGAMP is bound in the cavity at the bottom of the V. Binding induces the formation of a lid that closes the binding cavity and a further compaction of the dimer (apo STING LBD structure: 4EF5 [134], holo STING LBD structure: 4KSY [166]).

Human LBD structures with bacterial c-di-GMP bound differ from the 2,3-cGAMP bound structures. The c-di-GMP is positioned in the same ligand binding site but less deep and less coordinated than the endogenous ligand. The compaction is less distinct and the lid structure is not resolved [134, 136, 137] or in a different conformation [135]. Just one human structure [133] shows a well ordered lid. A possible explanation could be the use of the rare human H232 isoform [167] for crystallization that has been reported to be less responsive to c-di-GMP [166]. The human structure with the ordered lid contains the abundant R232 variant. This residue is located at the end of the outer strands of the lid. It contacts c-di-GMP via an Mg^{2+}-ion contributing to a proper lid formation. Thus, the H232 isoform could be compromised in proper lid formation which is observed as disordering of the lid region in the crystal structures and at the same time explains its functional impairment in sensing c-diGMP.

In general, the structures demonstrate that STING works as a dimer. Additionally, they support the model of STING being at the same time a PRR for bacterial cyclic dinucleotides and a central receptor for the cGAS-dependent cytosolic DNA sensing. Nevertheless, the structures still cannot explain how the signal is carried on in downstream signaling. Lid formation and compaction are considered to be part of STING activation but how this contributes to TBK1/IRF3 activation is yet unclear. The closed lid could provide an interaction surface for downstream factors but the CTT which was shown to be essential for TBK1 activation is not resolved in any structure [141].

1.3.6. Further regulation of the cGAS-STING axis

Several studies have identified mainly STING as a target for several different regulatory modifications.

The two RING finger E3 ubiquitin ligases tripartite motif-containing protein 56 and 32 (TRIM56, TRIM32) were shown to be responsible for K63-linked ubiquitinylation of STING at K150 after virus infection. This modification is suggested to contribute to STING activation by facilitating its oligomeriztion [168, 169]. Interestingly, the E3 ubiquitin ligase RING finger protein 5 (RNF5) was reported to also ubiquitinylate STING at the same position but as a K48-linked modification in a virus infection dependent manner. This was shown to negatively regulate STING by inducing its proteasomal degradation providing a regulatory negative feedback loop [170].

Another negative feedback loop for STING activation was reported to rely on the presence of cytosolic 2,3-cGAMP. Upon DNA sensing cGAS produces 2,3-cGAMP. Via an unknown mechanism 2,3-cGAMP inhibits AMPK (AMP-Activated Protein Kinase) which leads to the activation of ULK1 (UNC-51-Like Kinase 1). Activated ULK1 in turn phosphorylates STING at S366 (located in the CTT). This is suggested to render STING inactive, eventually leading to its degradation [171]. However, this finding is in contrast to another report where the phosphorylation of S366 has been shown to be essential for the STING-dependent activation of IRF3 by TBK1 [141]. The mutations S366D and S366A render STING inactive emphasizing the critical role of this residue but at the same time not clarifying if S366 phosphorylation regulates STING activity positively or negatively.

A further phosphorylation site is S358 which is also located in the CTT. It was shown that S358 phosphorylation is mediated by TBK1 and essential for downstream signaling after virus infection [128].

Besides of posttranslational modifications, STING-dependent signaling is also regulated by protein-protein interactions. NLRC3 was shown to bind to STING thereby disturbing the STING-TBK1 interaction, reducing the type I IFN production upon stimulation with DNA and c-di-GMP and HSV-1 infection [172]. A similar mechanism was observed for MRP (MITA-Related Protein). MRP is a splice variant of STING lacking the CTT. It is suggested to form heterodimers with STING also disturbing STING-TBK1 interaction and thereby inhibiting IFN-β expression [173].

The different regulatory modifications and inhibitors observed display a complex system which enables a fine tuning of the STING signaling. This is particularly important because major defects in the regulation of signal transduction are often causes of severe diseases.

1.3.7. STING pathway and autophagy

Autophagy is a digestion mechanism for various cytosolic components. It is mainly controlled by the autophagy related proteins (ATGs) ULK1/2 (UNC-51-Like Kinase 1/2), Beclin1, VPS34 (Vacuole Protein Sorting 34) and ATG14L (Autophagy Related 14 Like). Hallmark of autophagy is the formation of the autophagosome, a double membrane compartment around the component that has to be digested. After formation, the autophagosome fuses with a lysosome leading to the degradation of the inner membrane and the autophagosomal content. Autophagy is the key mechanism in several cellular processes. It is responsible for the degradation of potentially toxic protein aggregates that cannot be processed by the proteasomal pathway and plays a role in the organelle quality control by digesting defective organelles (e.g. leaky mitochondria). In times of starvation, autophagy is able to provide the cell with energy and other resources like amino acids from degraded proteins. Additionally, it plays a role as regulator and effector of innate and adaptive immunity (e.g. digestion of invading pathogens) [174, 175].

It is known that autophagy can be triggered by innate immune signalling pathways. Indeed, several studies have connected cGAS and STING to autophagy. Infections with DNA viruses HSV-1 and HCMV (human cytomegalovirus) induce autophagy in a STING-dependent manner [176, 177]. The same is observed during infections with *Mycobacterium tuberculosis*. The bacterium releases DNA into the cytosol of the host cell triggering autophagy via the STING pathway [178, 179].

After stimulation with DNA, STING colocalizes in the punctuate foci with Atg9a, p62 and LC3 which are all essential factors in autophagy. In fact, Atg9a depleted cells are not just impaired in autophagy but they also show an enhanced type I IFN production upon DNA stimulation suggesting that STING is negatively regulated by Atg9a [140].

Contradicting data exists about the role of autophagy factor Atg7. Atg7 deficient cells (defective in autophagy) were shown to be less effective in type I IFN expression in one study [176] while another study reported that knockdown of Atg7 does not alter cytokine production and STING translocation after DNA stimulation [140].

The possible role of the autophagy controlling factor ULK1/2 for the regulation of the STING pathway has been discussed before. Upon DNA sensing via cGAS and 2,3-cGAMP production ULK1 is activated and induces autophagy. At the same time, it negatively regulates STING probably via direct phosphorylation [171].

A further autophagy induction mechanism also relies on cGAS. After stimulation, DNA bound cGAS interacts with autophagy controlling protein Beclin-1. This interaction suppresses 2,3-cGAMP production by cGAS and releases Beclin-1 it from its inhibitor Rubicon. Hereupon, Beclin-1 gets activated and triggers autophagy while at the same time type I IFN production is reduced [180].

PRR signalling is important for the induction of an immune response and at the same time excessive signalling must be avoided to prevent uncontrolled and self-damaging immune reactions. The connection of the STING pathway to autophagy provides both: The activation of autophagy to clear the host from cytosolic pathogenic compounds and negative feedback mechanisms to reduce cytokine production thereby avoiding deleterious effects.

1.3.8. STING pathway and disease

The STING pathway is connected to several diseases. It was shown to play a role in virus and bacterial infections as well as in diseases caused by protozoan parasites (Table.1.3.8.1.) [181]. In most of these infections STING pathway activation contributes essentially to IFN production.

Even though the STING pathway is very effective in the induction of an immune response, some pathogens successfully developed mechanism to shut down this pathway. The NS2B3 protease complex of Dengue virus (DENV) cleaves human STING and thereby inactivates it [182, 183]. *Coronaviruses* reduce IFN production by interrupting STING dimerization via their papain-like protease NSP3 [184]. Furthermore, hepatitis C virus (HCV) protein NS4B interrupts STING interaction with components of the signalling complex e.g. TBK1 [185, 186]. These evasion mechanisms connect STING to RNA virus infections since DENV, HCV and *Coronaviruses* are all RNA viruses. However, it is suggested that these evasion mechanisms act on a second, indirect level of infection, which suppresses the induction of an immune response upon sensing released DNAs from already destroyed host cells via the STING pathway.

Apart from diseases caused by pathogenic infections the STING pathway is also involved in autoimmune diseases. An example is the Aicardi-Goutières syndrome (AGS) which is a hereditary disease affecting the skin and the brain. The disease manifests itself as a calcifying vasculitis leading to an early onset progressive encephalopathy. Mutations in the gene encoding for the protein TREX1 (Three Prime Repair Exonuclease 1) are the cause for this disease in Aicardi-Goutières syndrome group 1 (AGS1) patients [187]. TREX1 is an ER associated protein that degrades

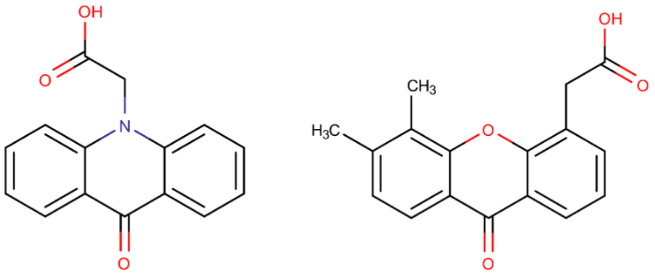
Pathogen name	**Pathogen class**	**Publication**
Vesicular stomatitis virus (VSV)	(-)ssRNA virus	[126, 127, 129]
Sendai virus (SeV)	(-)ssRNA virus	[127, 128]
Newcastle disease virus (NDV)	(-)ssRNA virus	[127]
Encephalomyocarditisvirus (EMCV)	(+)ssRNA virus	[129]
Japanese Encephalitis virus (JEV)	(+)ssRNA virus	[192]
Dengue virus (DENV)	(+)ssRNA virus	[182, 183]
Hepatitis C virus (HCV)	(+)ssRNA virus	[185, 186]
Severe acute respiratory syndrome coronavirus (SARS-CoV)	(+)ssRNA virus	[184, 193]
Human T cell leukemia virus type 1 (HTLV-1)	ssRNA-RT virus	[194]
Human immunodeficiency virus (HIV)	ssRNA-RT virus	[195, 196]
Herpes simplex virus 1 (HSV-1)	dsDNA virus	[129, 176, 177, 197, 198]
Nucleopolyhedrovirus (NPV)	dsDNA virus	[129]
Human cytomegalovirus (HCMV)	dsDNA virus	[129]
Modified vaccinia virus ankara (MVA)	dsDNA virus	[199]
Human papillomavirus (HPV)	dsDNA virus	[200]
Human serotype5 adenovirus (Ad5)	dsDNA	[201]
Hepatitis B virus (HBV)	dsDNA-RT virus	[202]
Listeria monocytogenes	Gram-positive bacterium, intracellular pathogen	[131, 203, 204]
Streptococcus pneumoniae	Gram-positive bacterium, extracellular pathogen	[205, 206]
Brucella abortus	Gram-negative bacterium, intracellular pathogen	[207]
Chlamydia trachomatis	Gram-negative bacterium, intracellular pathogen	[208]
Chlamydia muridarum	Gram-negative bacterium, intracellular pathogen	[209]
Francisella tularensis	Gram-negative bacterium, intracellular pathogen	[210]
Legionella pneumophila	Gram-negative bacterium, intracellular pathogen	[211]
Mycobacterium tuberculosis	Bacterium, no Gram staining, intracellular pathogen	[178, 179]
Plasmodium falciparum	Protozoan parasite	[212]

Table 1.3.8.1.: Viruses, bacteria and protozoan pathogens that have been shown to effect the STING pathway.

endogenous retroelements and byproducts of replication. TREX1 deficiency causes the accumulation of these DNAs and triggers a persistent type I IFN production which eventually activates factors of the adaptive immune system leading to autoimmune inflammation and fibrosis. Experiments in mice and in cell culture showed that the additional lack of components of the STING pathway (e.g. cGAS, STING, IRF3) leads to the loss of the persistent IFN signalling. Indeed, mice were completely rescued from the deleterious symptoms of the TREX1 deficiency phenotype which supports the critical role of the STING pathway in AGS [188-191].

1.3.9. STING pathway as a drug target

Since the connection between type I IFNs and antiviral and antibacterial responses of the Immune system were known researchers have tried to find activators for type I IFN production to fight infections. In this respect, sensing pathways of the innate immune system provide promising targets for the development of antiviral and antibacterial compounds. In fact, recent research on two molecules, DMXAA (5,6-Dimethylxanthenone-4-acetic acid) and CMA (10-carboxymethyl-9-acridanone) (Fig.1.3.9.), revealed their potential as activator of type I IFN production via the STING pathway.



CMA DMXAA

Fig.1.3.9.: Chemical structure of CMA (10-carboxymethyl-9-acridanone) and DMXAA (5,6-Dimethylxanthenone-4-acetic acid).

DMXAA was described as a vascular disrupting agent with potential use in cancer treatment [213], but failed in phase III clinical trials for application as a chemotherapeutic [214]. Nevertheless, it was shown to be able to induce type I IFNs and cytokines production in a STING-dependent manner in the mouse system providing a potential use as antiviral- and antibacterial drug [215, 216]. Biochemical assays and crystal structures of DMXAA bound to mouse STING provided evidence that DMXAA activates STING by direct interaction. However, these results did not translate into the human system. DMXAA could neither interact directly with human STING nor could it stimulate IFN production in human cells [150, 217].

The same was observed with CMA. CMA has been already described in 1976 as an antiviral compound [218]. The antiviral properties of CMA could be traced back to its ability to stimulate type I IFN production in the murine system [219]. Biochemical and crystallographic studies could show that CMA directly binds and activates mouse STING but fails to do so in the human system [220].

Both compounds are bound by mouse STING in the nucleotide binding site in a similar way. Binding leads to compaction and lid formation nearly identical to the active conformation of human STING. Strikingly, the nucleotide binding site is completely conserved in both species and only small structural differences between human and mouse STING probably contribute to the different affinities for CMA and DMXAA. This is also supported by the finding that a nucleotide binding site mutant (S162A) of human STING was able to bind DMXAA in an ITC assay [150].

Although both components could not bind to human STING these studies still can provide the basis for structure driven development of activators of type I IFN production.

1.3.10. Further proposed DNA sensors

Additional cytosolic DNA sensors were proposed in several studies before the identification of cGAS.

RNA polymerase III is able to transcribe transfected poly(dA:dT) into a 5'-triphosphate RNA that can fold into a double-stranded RNA ligand activating RIG-I. Even though, restricted to AT-rich DNA sequences, this mechanism probably allows for the utilisation of a RNA signalling pathway to defend against some DNA virus- and bacterial infections [221, 222].

Another cytosolic DNA sensing mechanism relies on DNA sensor AIM2 (absent in melanoma 2) and the adaptor protein ASC (apoptosis-associated speck-like protein containing a CARD). AIM2 recognizes cytosolic dsDNAs and recruits adaptor protein ASC to form an inflammasome. Like the NLR-dependent inflammasomes the AIM2 inflammasome does not trigger expression of type I IFNs but it is responsible for the maturation of IL1β and IL18 by activating caspase-1 [223-226].

Further putative DNA sensors are DAI (DNA-Dependent Activator of IRFs; also known as: ZBP1, DLM1) [227], the AIM2-like protein IFI16 (Gamma-Interferon-Inducible Protein 16) [228], the DExD/H-box helicase DDX41 [229] and two complexes involved in DNA damage repair: DNA-PK (DNA-Dependent Protein Kinase: composed of Ku70, Ku80 and DNA-PKc) [230, 231] and MRE11 (Meiotic Recombination 11 Homolog A)-Rad50-complex [232]. However, most of the studies connecting these proteins to DNA sensing rely on immunoprecipitation, immunofluorescence and RNAi knockdown experiments, which may measure indirect or nonspecific effects. The lack of direct experimental evidence and the existence of contradicting data might result from the fact that these proteins and complexes function as DNA sensors and contribute to the STING pathway induced IFN response just in certain cell types and/or certain infections [26, 31, 124].

2. Publications

2.1. The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signaling

2.1.1 Summary

Understanding the mechanisms of viral invasion and the reactions of the immune system is important for fighting infections. In this publication we present the structure of the mouse RIG-I SF2 domain in complex with the ATP analogue AMP-PNP (adenylyl-β,γ-imidotriphosphate). This structure enabled us to give insights into the mechanism of antiviral signalling of RIG-I.

The selenomethionine-labelled RIG-I SF2 domain was crystallized using an in-drop proteolysis approach. The structure displays an overall open “C-shaped” architecture composed of the three sub-folds 1A, 2A and 2B (denoted as Hel1, Hel2 and Heli in other publications; see Introduction). The RecA-like fold 1A is positioned at the N-terminus followed by the second RecA-like fold 2A. Fold 2B consists of a bundle of six α-helices and is inserted into fold 2A. The C-terminus is formed by two peculiar bend α-helices resembling an arm and elbow (denoted as bridging domain or pincer domain in other publications; see Introduction). In the nucleotide binding pocked located in fold 1A the ATP analogue AMP-PNP is bound.

In addition to the crystallization trials, we collected SAXS (Small-angle X-ray scattering) data of the nucleotide free human SF2 domain (signal-off state). The scattering amplitudes of nucleotide free human SF2 domain resembled very well the calculated amplitudes of the mouse crystal structure. Furthermore, the mouse crystal structure fitted well into the SAXS envelope of the human SF2 domain. This led to the assumption that the observed conformation of the mouse crystal structure is highly similar to the solution structure and represents an open signal-off state of the SF2 domain.

Obviously, the binding mode of AMP-PNP in the crystal structure describes the nucleotide binding mode in the signal-off state. In this mode, the hydrolysis of the bound nucleotide cannot be stimulated by the SF2 domain due to the misalignment of ATPase motives in fold 1A and 2A. Comparison with other SF2 enzymes revealed that a substantial conformational change is necessary to adopt a signal-on state that possesses ATPase activity. In this state the SF2 domain probably folds around the bound RNA which leads to the proper alignment of the ATPase motives and a general compaction. To validate the possible signal-on state mutational analysis in combination with ATPase assays was carried out. The results supported the model of a conformational switch between the off- and the on state of RIG-I signalling.

In the signal-off state the RIG-I SF2 adopts an open conformation while the CARDs rest in a signalling incompetent state. Binding of 5’-triphosphate dsRNA by the RIG-I RD triggers a conformational change that results in compaction of the SF2 domain around the RNA. This signal-on state stimulates ATPase activity and is accompanied by the release of the CARD inhibition.

With this publication we contributed to the understanding of the activation mechanism of RIG-I signalling, which is important for deciphering the complex proceedings during viral infections.

2.1.2. Contribution

The author of this thesis contributed to this publication by purifying the human RIG-I SF2 domain. Additionally, the author collected the SAXS data of the human RIG-I SF2 domain and helped with the analysis of this data.

2.1.3. Paper

The following paper was published 2011 in EMBO reports.

RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling

Filiz Civril, Matthew Bennett, Manuela Moldt, Tobias Deimling, Gregor Witte, Stefan Schiesser, Thomas Carell & Karl-Peter Hopfner

EMBO rep. 2011 Oct 28; **12**(11): 1127 - 34

EMBO
reports

scientific report

The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling

Filiz Civril[1], *Matthew Bennett*[1], *Manuela Moldt*[1], *Tobias Deimling*[1], *Gregor Witte*[1,2], *Stefan Schiesser*[3], *Thomas Carell*[3,4] *&* *Karl-Peter Hopfner*[1,2,4+]

[1]Department of Biochemistry at the Gene Center, [2]Munich Center for Advanced Photonics, [3]Department of Chemistry, and [4]Center for Integrated Protein Science, Ludwig-Maximilians-University Munich, Munich, Germany

RIG-I detects cytosolic viral dsRNA with 5′ triphosphates (5′-ppp-dsRNA), thereby initiating an antiviral innate immune response. Here we report the crystal structure of superfamily 2 (SF2) ATPase domain of RIG-I in complex with a nucleotide analogue. RIG-I SF2 comprises two RecA-like domains 1A and 2A and a helical insertion domain 2B, which together form a 'C'-shaped structure. Domains 1A and 2A are maintained in a 'signal-off' state with an inactive ATP hydrolysis site by an intriguing helical arm. By mutational analysis, we show surface motifs that are critical for dsRNA-stimulated ATPase activity, indicating that dsRNA induces a structural movement that brings domains 1A and 2A/B together to form an active ATPase site. The structure also indicates that the regulatory domain is close to the end of the helical arm, where it is well positioned to recruit 5′-ppp-dsRNA to the SF2 domain. Overall, our results indicate that the activation of RIG-I occurs through an RNA- and ATP-driven structural switch in the SF2 domain.

Keywords: ATPase; crystal structure; innate immunity; RIG-I; viral RNA

EMBO *reports* (2011) **12**, 1127–1134. doi:10.1038/embor.2011.190

INTRODUCTION

The innate immune system is the first line of defence against infections by pathogens. In the innate immune system, pattern recognition receptors (PRRs) distinguish self versus non-self by binding molecular patterns that are present on pathogen-associated molecules (pathogen-associated molecular patterns (PAMPs)) but typically not on host molecules (Takeuchi & Akira, 2010). In mammals, the formation of PRR–PAMP complexes start signalling cascades that activate the transcription factors nuclear factor-κB and interferon regulatory factors, triggering host defense mechanisms such as the activation of interferon-regulated genes and inflammatory responses.

RIG-I-like receptors (RLRs) sense cytoplasmic viral RNA and comprise RIG-I, MDA5 (melanoma differentiation-associated protein 5) and LGP2 (laboratory of genetics and physiology 2; Kang *et al*, 2002; Yoneyama *et al*, 2004; Rothenfusser *et al*, 2005). RIG-I senses infections from viruses such as hepatitis C virus, Sendai virus, influenza virus, vesicular stomatitis virus, rabies virus and Japanese encephalitis virus (Kato *et al*, 2006).

After sensing RNA ligands, RIG-I associates with interferon-β promoter stimulator 1 (IPS-1), a process that involves RIG-I ubiquitination and/or interaction with ubiquitin chains (Gack *et al*, 2007; Zeng *et al*, 2010).

The optimal ligand for RIG-I has been found to be base-paired or double-stranded RNA (dsRNA) molecules containing a 5′ triphosphate (5′-ppp-dsRNA; Hornung *et al*, 2006; Pichlmair *et al*, 2006; Schlee *et al*, 2009). RIG-I contains two N-terminal caspase activation and recruitment domains (CARDs), which are required for interaction with IPS-1 (Yoneyama *et al*, 2004), a superfamily 2 helicase/translocase/ATPase (SF2) domain and a C-terminal regulatory/repressor domain (RD; Saito *et al*, 2007). The RD of RIG-I is required to activate the SF2 domain by binding 5′-ppp-RNA (Cui *et al*, 2008; Takahasi *et al*, 2008). Recent structural data show that RLR RDs are base-paired end-recognition modules with a preference for 5′-ppp-dsRNA in the case of RIG-I (Lu *et al*, 2010; Wang *et al*, 2010). However, the functional role of the SF2 domain of the RIG-I is unclear. RLRs possess the typical seven core motifs (I, Ia, II–VI) that are involved in the recognition of ATP and nucleic acid substrates (Singleton *et al*, 2007). In contrast to bona fide helicases that unwind duplex nucleic acids, RIG-I translocates on dsRNA in a 5′-ppp-dependent manner (Myong *et al*, 2009). It is unclear how dsRNA recognition by the SF2 domain and 5′-ppp recognition are linked to RIG-I activation, but current models indicate conformational changes that expose CARDs for downstream signalling (Yoneyama & Fujita, 2009).

To help reveal the mechanism of ATP-dependent RIG-I activation, we have determined the crystal structure of the mouse RIG-I SF2 domain with adenosine 5′-(β,γ-imido)triphosphate (AMP-PNP) bound.

---

[1]Department of Biochemistry at the Gene Center,
[2]Munich Center for Advanced Photonics,
[3]Department of Chemistry, and
[4]Center for Integrated Protein Science, Ludwig-Maximilians-University Munich, Feodor-Lynen-Strasse 25, Munich 81377, Germany
[+]Corresponding author. Tel: +49 (0) 89 2180 76953; Fax: +49 (0) 89 2180 76999; E-mail: hopfner@genzentrum.lmu.de

Received 8 September 2011; accepted 8 September 2011;

Fig 1 | Structure of the mouse RIG-I SF2 domain. (**A**) Front view of RIG-I^{SF2} along the nucleotide-binding cleft. RIG-I^{SF2} consists of three domains: SF2 domain 1A (yellow) and SF2 domain 2A (green) are connected by a short linker helix (grey) and form the conserved ATP-binding and hydrolysis core. The helical bundle domain 2B (red) is a specific feature of RIG-I/FANCM/Hef helicases, indicating that it is involved in double-stranded nucleic acid binding or translocation. An unusual arm (orange), unique to RIG-I-like receptors, reaches from domain 2 across domain 1 and stabilizes the observed 'open' conformation. (**B**) Top view of RIG-I, coloured as in (**A**). (**C**) Close-up view of the helical arm and its elbow (orange), embracing the helical protrusion from domain 1A, with a hydrophobic interface. (**D**) Structure-based sequence alignment of selected RIG-I-like receptors with highlighted conserved residues and annotated motifs. The secondary structural elements are shown on top of the alignment. AMP-PNP, adenosine 5′-(β,γ-imido)triphosphate; RIG-I, retinoic acid inducible gene I; SF2, superfamily 2.

►

RESULTS

Crystal structure of the mouse RIG-I helicase domain

The SF2 domain of mouse RIG-I (mmRIG-I) was crystallized using an in-drop proteolysis approach from selenomethionine-labelled protein, and its structure determined to a resolution of 2.2 Å (which we have denoted as RIG-I^{SF2}). The structure has been deposited at the Protein Data Bank with ID 3TBK. The SF2 domain of RIG-I is a 'C'-shaped particle consisting of three structural domains denoted 1A, 2A and 2B in analogy to other SF1 and 2 enzymes (Fig 1A,B). 'RecA-like' domain 1A extends from residues 244 to 445 and carries motifs Q, I (Walker A), Ia, Ib, Ic, II (Walker B), IIa and III. It is connected to domain 2A by a short, α-helical hinge. 'RecA-like' domain 2A comprises residues 456–467 and 610–746 and carries motifs IV, V and VI (Fig 1C). 'Helical insertion' domain 2B is a compact bundle of six α-helices (α10–α15) inserted into domain 2A at the loop between β8 and α16. It is a specific feature of RLRs and the related Hef/FANCM DNA repair ATPases (supplementary Fig S3 online). Domain 2B is situated 'on top' of domain 2A, at the site that typically binds to nucleic acids in SF2 enzymes, indicating that it could be an important element of the dsRNA recognition or translocase function of RIG-I.

The three domains are connected to each other by small, mainly hydrophobic interfaces, which seem to be conserved in RIG-I, opening the possibility for conformational changes that might be important in the context of PAMP recognition and signalling. However, an unusual 'helical arm' (α18 and α19) reaches from domain 2A back to and across domain 1A. Bound to the inner side of the 'elbow' is a helical protrusion in domain 1 (α8) that interacts with both arm helices by an extended aromatic/hydrophobic interface (Fig 2A). The arm therefore seems to maintain stabilization between domains 1A and 2A in the absence of RNA in the observed conformation (which we have denoted as 'open' as described below).

Solution structure of the RIG-I helicase

To confirm that the observed 'open' state resembles the solution structure of RIG-I^{SF2}, we used small-angle X-ray scattering (SAXS) on nucleotide-free human RIG-I^{SF2}.

The scattering amplitudes measured for human RIG-I^{SF2} match those calculated from the crystal structure model of mouse RIG-I^{SF2} very well (Fig 2B). The crystal structure of mouse RIG-I^{SF2} also docks well into the averaged SAXS envelope of human RIG-I^{SF2} (Fig 2C). This shows that human RIG-I^{SF2} adopts a similar conformation in solution to mouse RIG-I^{SF2} in the crystal. We therefore propose that this conformation represents a 'signal-off' state of the RIG-I^{SF2} domain, found in the absence of an RNA ligand.

Helicase motifs and AMP-PNP recognition

The seven helicase motifs are located in loops on the surfaces of domains 1A and 2A that sandwich ATP during the ATP-hydrolysis cycle (Fig 3A). The AMP-PNP ligand is clearly defined in the electron density, bound to motif I (Fig 3B; supplementary Fig S1B online). The adenine moiety is situated in a partially polar, partially hydrophobic pocket. It is recognized by two hydrogen bonds from Q248 (Q-motif) to adenine N_6 and N_7 and by a hydrogen bond from the carbonyl oxygen of K243 to N_6. The walls of the pocket are formed by the side chains of R245, L242 and F273, providing extra stacking and hydrophobic interactions.

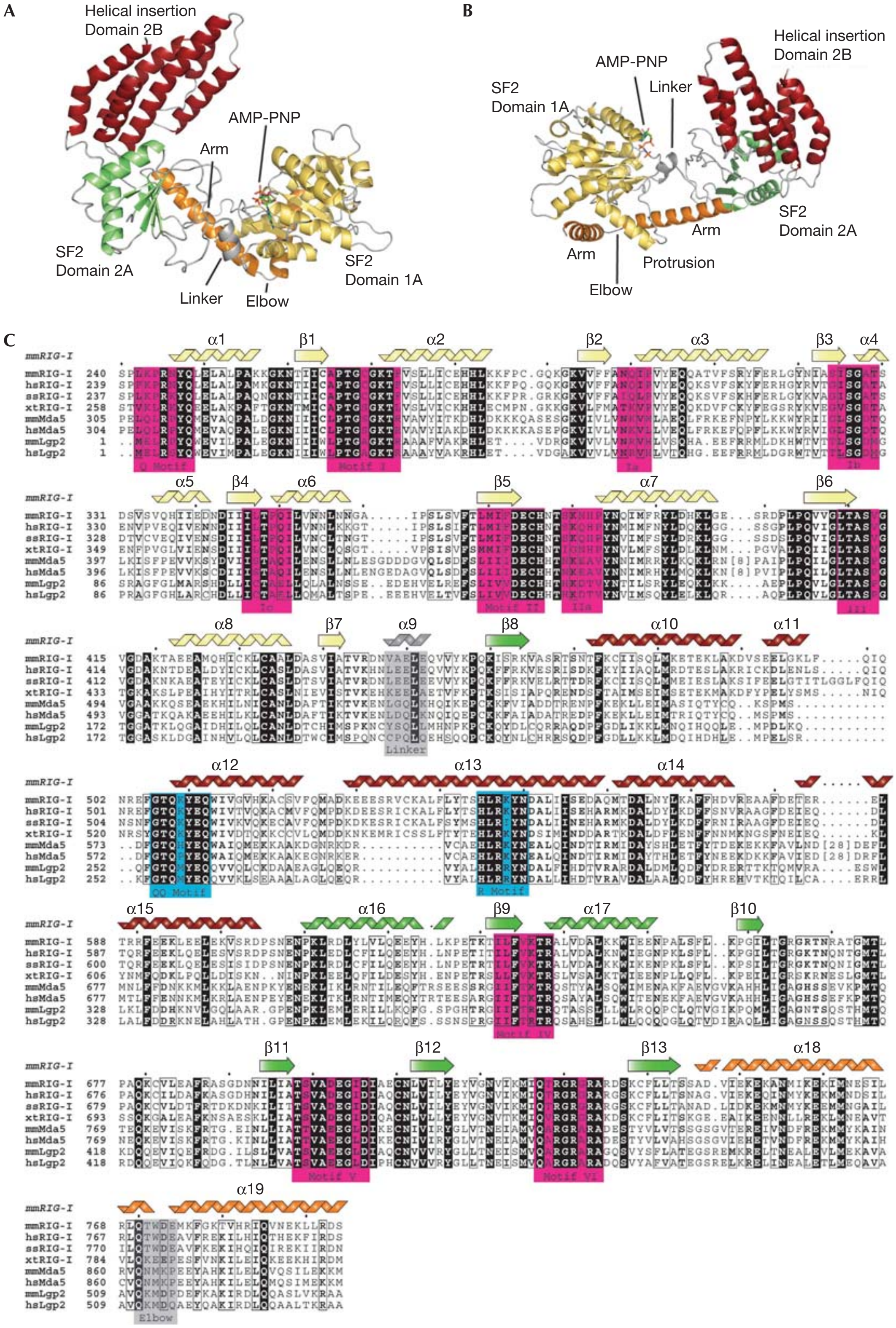
The β- and γ-phosphates are bound to motif I, whereas the α-phosphate lacks direct contacts (Fig 3B). The triphosphate chain appears slipped by one phosphate compared with the canonical binding; the γ-phosphate is in the position that is typically occupied by the β-phosphate. It is unclear whether this is a specific feature of RIG-I or whether conformational movements occur to allow canonical binding of the triphosphate chain.

Motif III is located in the loop that connects β6 to the α8 protrusion that binds to the helical arm. In contrast to mutations in other motifs, which abolish RIG-I activation by viral ligands, mutation of motif III was found to render RIG-I constitutively active (Bamming & Horvath, 2009), suggesting that perturbation of the protrusion–helical arm interaction might allow RIG-I activation in the absence of proper ligands.

Functional motifs are not aligned in the 'signal-off' state

Analysis of the relative location of motifs I–III and IV–VI with respect to each other shows that domain 2A needs to rotate substantially to engage domain 1A to form a functional ATPase site. For instance, motif VI residue R731 typically interacts with the triphosphate backbone, yet in the crystal structure it is $\sim$25 Å from the γ-phosphate bound to motif I. Thus, the lack of ATPase activity in the signal-off state of RLRs can be explained by the stabilization of domains 1A and 2A in the observed 'open' orientation with misaligned helicase motifs. As we crystallized RIG-I SF2 in the presence of AMP-PNP and yet observe an 'open' structure, ATP alone is probably not sufficient to orientate the SF2 domains of RIG-I properly. In SF2 enzymes, nucleic acids are typically required to appropriately position domains 1A and 2A for ATP hydrolysis to occur. In addition, specific recognition of proper RNA ligands requires insertion domains, such as domain 2B in RIG-I^{SF2}.

To learn more about the conformational change and model how RIG-I^{SF2} might bind to RNA in an activated form, we independently superimposed RIG-I^{SF2} domains 1A and domains 2A–2B with equivalent domains 1A and 2A of the activated RNA-bound form of VASA, an SF2 DEAD box helicase with similarity to RIG-I^{SF2} (PDB 2DB3; Fig 4A; Fairman-Williams *et al*, 2010).



This shows that a rotation of domain 2A–2B of ~90° would properly align all motifs for activation. Notably, in the modelled 'signal-on' conformation, the helical domain 2B is situated opposite domain 1A forming a channel that is shaped to grip dsRNA (Fig 4B). The superposed conformation is also consistent with the presence of the helical arm. Although the

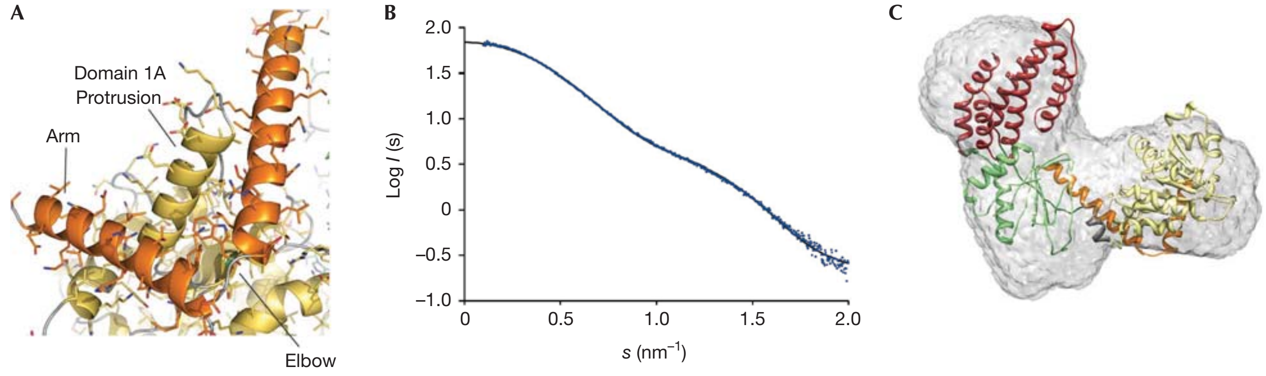


Fig 2 | Human RIG-I^{SF2} solution structure. (**A**) Experimental X-ray scattering intensities of human RIG-I^{SF2} (blue dots) and calculated scattering intensities from the crystal structure of mouse RIG-I^{SF2} (black line), indicating that the mouse RIG-I^{SF2} crystal structure is a good model for the human RIG-I^{SF2} solution structure. (**B**) Averaged small-angle X-ray scattering envelope obtained from 16 independent *ab initio* models from human RIG-I^{SF2} (grey) shows the three-lobed 'C' shape, with the model of the crystal structure of mouse RIG-I^{SF2}. Figures and docking were generated using the Situs package (Wriggers, 2010) and University of California, San Francisco Chimera (Pettersen *et al*, 2004). RIG-I, retinoic acid inducible gene I.

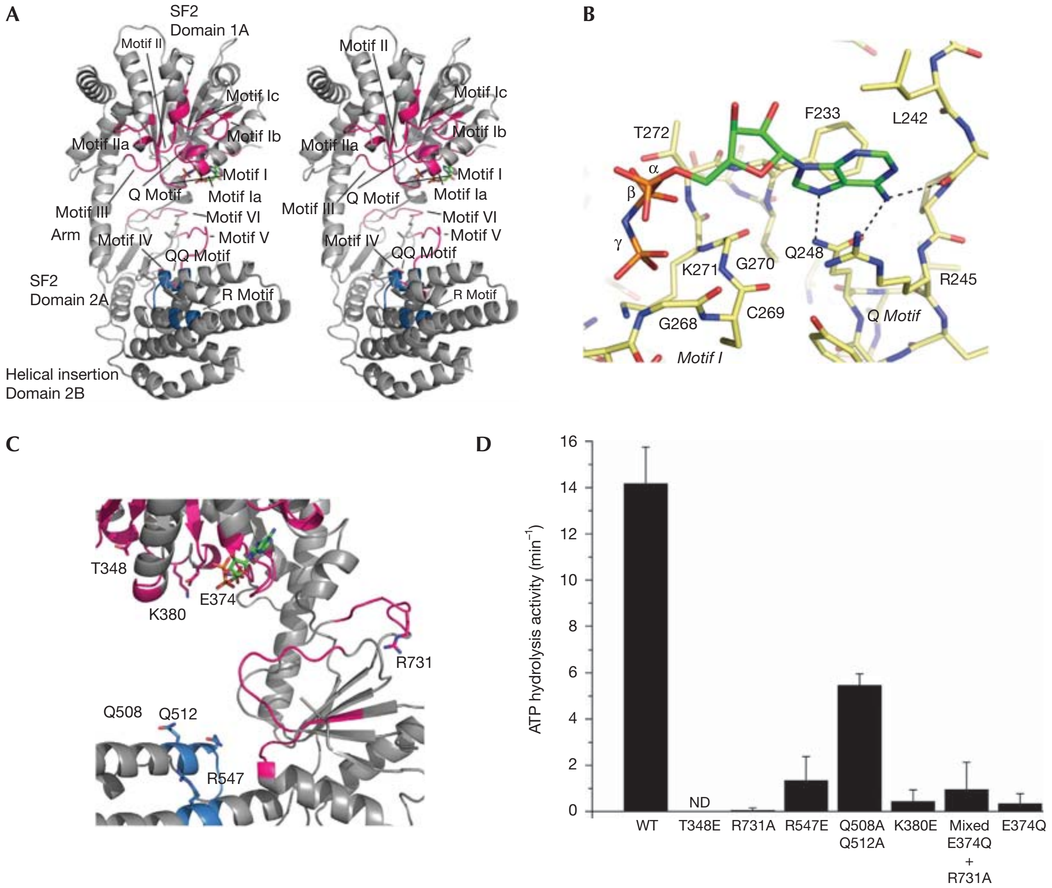


Fig 3 | Conserved motifs and mutational analysis. (**A**) Stereo view of the structure of RIG-I^{SF2} with conserved functional motifs shown. (**B**) Close-up view of the adenosine 5′-(β,γ-imido)triphosphate (AMP-PNP; colour-coded sticks: green, carbon; red, oxygen; blue, nitrogen; orange, phosphorus)-binding moiety with highlighted hydrogen bonds of the adenine recognition site. The protein is shown as colour-coded sticks with yellow carbons. (**C**) Close-up view of the motifs mutated in this study. Mutated side chains are annotated and shown as sticks. (**D**) ATP hydrolysis activity of RIG-I SF2 domain mutants. Mixed E374Q + R731A represents a 1:1 mixture of single mutants E374Q and R731A. Plotted bars: mean ± s.d. ($n = 3$–6). ND, not detectable; RIG-I, retinoic acid inducible gene I; SF2, superfamily 2; WT, wild type.

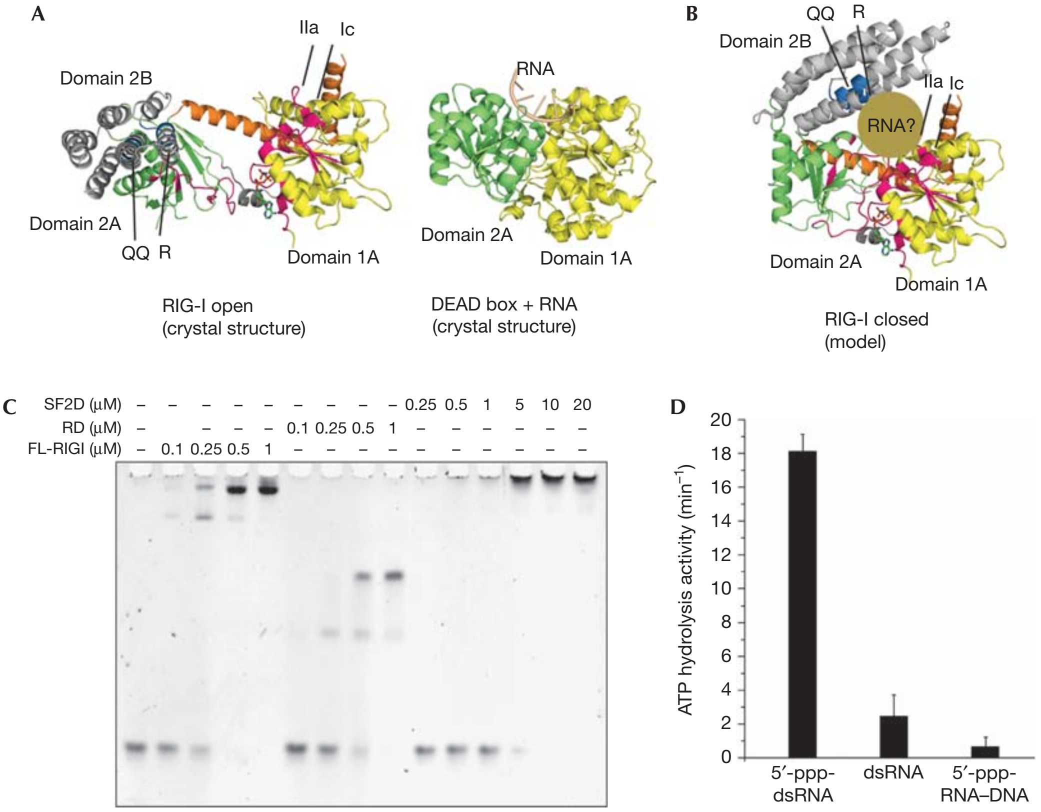


Fig 4 | Formation of an ATP hydrolysis site requires a structural switch. (**A**) Comparison of RIG-I^{SF2} (open conformation) with DEAD-box enzyme VASA (closed, RNA-bound conformation). Domains 1A are superimposed, showing the strongly differing orientation of domain 2 in RIG-I, stabilized by the helical arm. (**B**) Model of the RIG-I^{SF2} closed conformation, generated by superimposing domain 2A on the structure of VASA domain 2A. All helicase motifs align properly, and this conformation positions QQ and R motifs opposite motifs Ic and IIa, where they are ideally placed to grip dsRNA (brown). (**C**) Electrophoretic mobility shift analysis of full-length (FL) RIG-I, RD and SF2 domain (SF2D) with 5′-ppp-dsRNA. The RNA-binding affinities of RD and FL are comparable and higher than that of SF2 domain, showing that RD provides the main binding affinity of RIG-I. (**D**) ATP hydrolysis activity of RIG-I (200 nM) activated with 5′-ppp-dsRNA (25mer), dsRNA (25mer) or 5′-ppp-RNA–DNA hybrid. Plotted bars: mean ± s.d. ($n = 7$–10). dsRNA, double-stranded RNA; 5′-ppp, 5′ triphosphates; RD, repressor domain; RIG-I, retinoic acid inducible gene I; SF2, superfamily 2.

portion of domain 2A that connects to the helical arm rotates significantly, it only shifts its position by 11 Å and only modest conformational changes are required in the helical arm to accommodate this shift.

RIG-I activation requires conformational change

To understand better how RNA binding might cause a transition between the off and on states, we performed mutational analyses to locate residues important for dsRNA-stimulated ATPase activity in full-length RIG-I.

Most of the dsRNA-binding activity of RIG-I is contributed by RD (Fig 4C), which thus masks effects of mutations in SF2 in a direct RNA-binding assay (supplementary Fig S2A online). Unfortunately, without RD, SF2 alone is mostly inactive because it needs RD to present RNA for binding. However, the ATPase activity of full-length RIG-I is very sensitive to RNA-induced conformations in SF2 and can therefore help map RNA-binding motifs.

Point mutations in domain 1A motif II (E374Q) and domain 2A motif VI (R731A; Fig 3C) practically abolish 5′-ppp-dsRNA-induced ATP hydrolysis (Fig 3D), consistent with our model that these motifs must come together to form an active site. It has also been proposed that RIG-I could form multimers, in principle allowing domains 1A and 2A from different RIG-I molecules to function *in trans*, without necessitating a conformational change. However, when E374Q and R731A mutants are combined in the same reaction, ATPase activity remains diminished (Fig 3D). This lack of complementation indicates that domains 1A and 2A from the same RIG-I molecule need to come together by a conformation change to hydrolyse ATP.

Two highly conserved motifs (named QQ-motif and R-motif) form a positively charged face on the helical bundle on domain 2B (Fig 1C). In the crystal structure, these motifs are not positioned to bind RNA bound at domain 1A. However, in the model for the closed conformation (Fig 4B), the QQ- and R-motifs are repositioned opposite domain 1A in the proposed RNA-binding

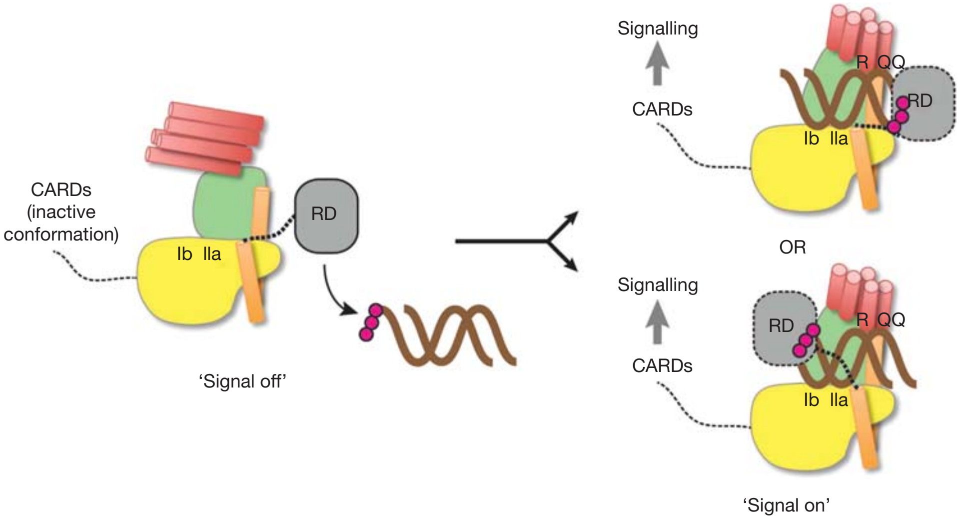


Fig 5 | Proposed model for RIG-I activation by a conformational switch in the SF2 domain. RD binds to dsRNA with 5′ triphopshates (5′-ppp; magenta spheres) and might recruit it to the SF2 domain. RNA and ATP binding switches SF2 into signal-on conformation by gripping RNA between motifs Ic/IIa on domain IA and R/QQ on domain IIB. The position of the helical arm with the short linker to RD might allow RD to bind 5′-ppp-RNA ends cooperatively with SF2. The precise position of RD, which from our structure might bind on either side of the RNA duplex (both possibilities are shown), and the position of activation and recruitment domains (CARDs) in signal-off and -on states, remain to be determined. dsRNA, double-stranded RNA; RD, repressor domain; RIG-I, retinoic acid inducible gene I; SF2, superfamily 2.

cleft. To test the relevance of these motifs, we mutated R 547 to glutamate and Q508/Q512 to alanine, and found ATPase activity reduced to ~10% and 40%, respectively (Fig 3D). The mutants retained RNA-binding activity (supplementary Fig S2B online), although this is probably due to the high affinity of the RD domain for RNA ligands (Fig 4C).

Swi/Snf2 helicases, in addition to sharing significant homology with RLRs (Fairman-Williams *et al*, 2010), are also stimulated by double-stranded nucleic acids and translocate on dsDNA by ATP-dependent tracking of the minor groove. Binding to dsDNA occurs through motifs Ia, Ib and Ic, recognizing the 3′→5′ strand, and IIa (recognizing the 5′→3′ strand; Durr *et al*, 2005).

This functional and sequence similarity between RIG-I and Swi2/Snf2 enabled us to analyse the recognition of nucleic acid duplexes by motifs Ic and motif IIa (T 348 and K 380, respectively; Fig 3C). To test the functional relevance of these motifs, we mutated T 348 and K 380 to glutamate. Both mutations abolished 5′-ppp-RNA-stimulated ATPase activity of full-length mmRIG-I (Fig 3D). As these residues are involved in nucleic acid binding in Swi/Snf2, it seems likely that motifs Ic and IIa are similarly involved in RNA duplex binding in RIG-I. A 5′-ppp-RNA–DNA hybrid binds to RIG-I similarly to 5′-ppp-dsRNA, but lacks the ability to induce ATP hydrolysis (Fig 4D; supplementary Fig S2C online), showing that a proper ATPase site is only formed when both strands of the ligand are RNA.

A model for RIG-I activation by RNA binding

Our structure, together with mutational analyses, shows how the SF2 domain of RIG-I could function as an RNA ligand-induced activation switch (Fig 5).

Interestingly, as there are only nine residues between the last structurally defined residue in RIG-I^{SF2} and the first structurally defined residues in RIG-I RD (missing in our structure), we can also roughly position RD with respect to an RNA ligand (Fig 5).

Our structure therefore provides a high-resolution framework for the signal-off state of the SF2 domain and a plausible mechanistic model for a signal-off to signal-on switch. The conservation of all motifs and the helical arm in MDA5 and LGP2 indicates that the structural model presented here is also a good framework to understand and analyse these other RLRs.

METHODS

Crystallization and structure determination. For protein purification and crystallization conditions see supplementary information online. Selenomethionine-labelled RIG-I^{SF2} was crystallized from ΔRD–RIG-I in complex with ANP-PNP using an in-drop proteolysis approach with subtilisin, which generated an SF2 helicase core with two internal loops removed (supplementary Fig S1A online). X-ray diffraction data were collected at Swiss Light Source (SLS) at the X06SA beamline at 100 K and a wavelength of 0.97972 Å (Se peak). Phenix.autosol (Terwilliger *et al*, 2009) was used to locate Se sites and produce a solvent-flattened map (supplementary Fig S1B online). A model was built using Phenix.autobuild (Adams *et al*, 2010). The model was manually improved using Coot (Emsley *et al*, 2010) and refined using Phenix.refine to an R_{work}/R_{free} of 19.1%/23.3%. Data and model statistics are shown in Table 1. The structure has been deposited at the Protein Data Bank with ID 3TBK.

Small-angle X-ray scattering. Small-angle X-ray scattering experiments were conducted at the The European Molecular Biology Laboratory/Deutsches Electronen-Synchotron X33 beamline. All proteins used were purified by size-exclusion chromatography before SAXS measurements, with running buffer used as reference. Data were processed and analysed using the ATSAS package

Table 1 | Data collection and refinement statistics

Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	45.98 86.21 152.77
α, β, γ (deg)	90, 90, 90
Resolution (Å)	50.0–2.14 (2.27–2.14)*
R_{sym}	10.2% (46.3%)
$I/\sigma I$	16.09 (3.35)
Completeness	99.4% (96.5%)
Redundancy	5.5 (5.2)
Phasing	
Figure of merit	0.47
Correlation coefficient	11.9
Refinement	
Resolution (Å)	43.89–2.14 (2.2–2.14)
No. of reflections	34,468
R_{work}/R_{free}*	19.1/23.3 (20.9/27.9)
No. of atoms	
Protein	4,182
Ligand/ion	35
Water	311
B-factors	
Protein	32.74
Ligand/ion	28.53
Water	35.28
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (deg)	0.912
Ramachandran values	
Favoured	98.6%
Allowed	1.4%
Outliers	0%
Protein Data Bank accession code	3TBK

*Values in parentheses are for highest-resolution shell.

(Konarev *et al*, 2006) as described in Putnam *et al* (2007). Theoretical scattering curves were calculated using CRYSOL (Svergun *et al*, 1995). Sets of independent *ab initio* models were calculated using GASBOR, and then averaged and aligned using DAMAVER (Volkov & Svergun, 2003). Figures including docking were generated using the Situs-Package (Wriggers, 2010) and UCSF Chimera (Pettersen *et al*, 2004).

RNA and DNA oligonucleotides. HPLC-grade RNA and DNA oligonucleotides were purchased from Biomers. The forward sequence used in this study is 5′-ACCAAACAAGAGAAGAAACAUGUAC-3′. 5′-ppp-RNA was synthesized as previously described (Ludwig & Eckstein, 1989; Paul *et al*, 2006) and purified by C_{18} reverse-phase chromatography at 94% purity confirmed by capillary electrophoresis.

ATPase assay. The reactions were performed in 100 mM Tris–HCl (pH 8.0), 150 mM NaCl, 10 μM $ZnCl_2$, 5 mM $MgCl_2$, 2% glycerol and 2 mM dithiothreitol (DTT). Increasing concentrations (25 nM–0.5 μM) of proteins were incubated with 1 μM *in vitro*-synthesized 5′-ppp-dsRNA, unless otherwise indicated, and 100 μM ATP including 10 nM of [γ-^{32}P]ATP for 30 min at 37 °C. Free phosphate was separated by thin-layer chromatography. Images were analysed with ImageJ. Only the linear part of the concentration curve was used for calculations.

Electrophoretic mobility shift assay. A measure of 200 nM RNA was incubated with the indicated amount of purified protein for 30 min on ice in 20 mM Tris–HCl (pH 8.0), 150 mM NaCl, 10 μM $ZnCl_2$ and 2 mM DTT reaction buffer. Samples were separated by native PAGE and stained with Gel-Red (Biotium). Gel images were analysed by ImageJ.

Supplementary information is available at EMBO *reports* online (http://www.emboreports.org).

ACKNOWLEDGEMENTS
We thank Sheng Cui for discussions and valuable support during the early phase of these studies. We thank the Max-Planck-Crystallization Facility, Martinsried, for initial screening for crystallization conditions and the staff of the Swiss Light Source (SLS; Villigen, Switzerland), The European Molecular Biology Laboratory/Deutsches Electronen-Synchotron (EMBL/DESY; Hamburg, Germany) and European Synchrotron Radiation Facility (ESRF; Grenoble, France) for excellent support. This work was funded by the National Institutes of Health grant U19AI083025 and grants from the Deutsche Forschungsgemeinschaft (DFG HO2489/3 and SFB455) to K.-P.H. K.-P.H. and T.C. acknowledge financial support by the Center for Integrated Protein Science Munich. We also acknowledge generous grants of beam time by DESY, SLS and ESRF.

Author contributions: F.C. purified and crystallized RIG-I^{SF2}, participated in structure determination and performed biochemical analyses. M.B. determined the structure and built the atomic model. M.M. helped with protein purification and crystal setups. T.D. purified human RIG-I^{SF2} and G.W. performed the SAXS analysis. S.S. synthesized 5′-ppp-RNA and T.C. designed and supervised the chemical synthesis. K.-P.H. designed and supervised the research on RIG-I and wrote the manuscript.

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

REFERENCES

Adams PD *et al* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**(Part 2)**:** 213–221

Bamming D, Horvath CM (2009) Regulation of signal transduction by enzymatically inactive antiviral RNA helicase proteins MDA5, RIG-I and LGP2. *J Biol Chem* **284:** 9700–9712

Cui S, Eisenacher K, Kirchhofer A, Brzozka K, Lammens A, Lammens K, Fujita T, Conzelmann KK, Krug A, Hopfner KP (2008) The C-terminal regulatory domain is the RNA 5′-triphosphate sensor of RIG-I. *Mol Cell* **29:** 169–179

Durr H, Korner C, Muller M, Hickmann V, Hopfner KP (2005) X-ray structures of the *Sulfolobus solfataricus* SWI2/SNF2 ATPase core and its complex with DNA. *Cell* **121:** 363–373

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**(Part 4)**:** 486–501

Fairman-Williams ME, Guenther UP, Jankowsky E (2010) SF1 and SF2 helicases: family matters. *Cur Opin Struct Biol* **20:** 313–324

Gack MU *et al* (2007) TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* **446:** 916–920

Hornung V *et al* (2006) 5′-Triphosphate RNA is the ligand for RIG-I. *Science* **314:** 994–997

Kang DC, Gopalkrishnan RV, Wu Q, Jankowsky E, Pyle AM, Fisher PB (2002) mda-5: an interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties. *Proc Natl Acad Sci USA* **99:** 637–642

Kato H *et al* (2006) Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441:** 101–105

Konarev PV, Petoukhov MV, Volkov VV, Svergun DI (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. *J Appl Crystallogr* **39:** 277–286

Lu C, Xu H, Ranjith-Kumar CT, Brooks MT, Hou TY, Hu F, Herr AB, Strong RK, Kao CC, Li P (2010) The structural basis of 5′ triphosphate double-stranded RNA recognition by RIG-I C-terminal domain. *Structure* **18:** 1032–1043

Ludwig J, Eckstein F (1989) Rapid and efficient synthesis of nucleoside 5′-O-(1-thiotriphosphates), 5′-triphosphates and 2′,3′-cyclophosphorothioates using 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one. *J Org Chem* **54:** 631–635

Myong S, Cui S, Cornish PV, Kirchhofer A, Gack MU, Jung JU, Hopfner KP, Ha T (2009) Cytosolic viral sensor RIG-I is a 5′-triphosphate-dependent translocase on double-stranded RNA. *Science* **323:** 1070–1074

Paul N, Springsteen G, Joyce GF (2006) Conversion of a ribozyme to a deoxyribozyme through *in vitro* evolution. *Chem Biol* **13:** 329–338

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25:** 1605–1612

Pichlmair A, Schulz O, Tan CP, Naslund TI, Liljestrom P, Weber F, Reis e Sousa C (2006) RIG-I-mediated antiviral responses to single-stranded RNA bearing 5′-phosphates. *Science* **314:** 997–1001

Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40:** 191–285

Rothenfusser S, Goutagny N, DiPerna G, Gong M, Monks BG, Schoenemeyer A, Yamamoto M, Akira S, Fitzgerald KA (2005) The RNA helicase Lgp2 inhibits TLR-independent sensing of viral replication by retinoic acid-inducible gene-I. *J Immunol* **175:** 5260–5268

Saito T, Hirai R, Loo YM, Owen D, Johnson CL, Sinha SC, Akira S, Fujita T, Gale M Jr (2007) Regulation of innate antiviral defenses through a shared repressor domain in RIG-I and LGP2. *Proc Natl Acad Sci USA* **104:** 582–587

Schlee M *et al* (2009) Recognition of 5′ triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity* **31:** 25–34

Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* **76:** 23–50

Svergun D, Barberato C, Koch MHJ (1995) CRYSOL—A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28:** 768–773

Takahasi K, Yoneyama M, Nishihori T, Hirai R, Kumeta H, Narita R, Gale M Jr, Inagaki F, Fujita T (2008) Nonself RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses. *Mol Cell* **29:** 428–440

Takeuchi O, Akira S (2010) Pattern recognition receptors and inflammation. *Cell* **140:** 805–820

Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, Hung LW (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65:** 582–601

Volkov VV, Svergun D (2003) Uniqueness of ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* **36:** 860–864

Wang Y *et al* (2010) Structural and functional insights into 5′-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nat Struct Mol Biol* **17:** 781–787

Wriggers W (2010) Using Situs for the integration of multi-resolution structures. *Biophys Rev* **2:** 21–27

Yoneyama M, Fujita T (2009) RNA recognition and signal transduction by RIG-I-like receptors. *Immunol Rev* **227:** 54–65

Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, Taira K, Akira S, Fujita T (2004) The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* **5:** 730–737

Zeng W, Sun L, Jiang X, Chen X, Hou F, Adhikari A, Xu M, Chen ZJ (2010) Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell* **141:** 315–330

2.2. Species-specific detection of the antiviral small-molecule CMA by STING

2.2.1. Summary

Since the connection between type I IFN production and the activation of the immune system was revealed lots of effort has been done to identify small-molecule compounds, which are able to trigger type I IFN production to fight infections. Indeed, PRRs and the connected pathways of the innate immune system are promising targets for the activation via small-molecule compounds. In this publication we analysed the mechanism of STING activation by the antiviral compound CMA (10-carboxymethyl-9-acridanone).

CMA is a small-molecule compound that is an approved antiviral drug in Russia and was shown to trigger type I IFN response in rodents. In experiments in mouse cell cultures we could show that the CMA indeed induces the production of type I IFNs and that this induction is dependent on the STING pathway. Unfortunately, the same experiments in various human cell lines showed no effect of CMA on the induction of type I IFNs. Further experiments in human HEK293T cells with wild type (wt) and chimeric mouse and human STING constructs were carried out. These experiments showed that CMA activates the STING pathway by direct interaction with the LBD of mouse STING but fails to do so in the context of the human STING-LBD. This observation of the species specific activation of STING by CMA was supported by *in vitro* experiments using DSF (Differential Scanning Fluorimetry) assays. In these assays CMA could stabilize the mouse STING-LBD at high temperatures but not the human STING-LBD. The crystal structure of mouse STING-LBD in complex with CMA could proof the direct interaction of CMA with mouse STING. In this structure two molecules of CMA are bound in the ligand binding cavity of the STING-LBD. Binding of CMA also triggers lid formation and compaction and thereby explains the activation of mouse STING by CMA (see Introduction).

In this publication we were able to explain how CMA stimulates type I IFN production in the mouse system by directly activating STING. The unresponsiveness of the human system to CMA could be ascribed to species specific differences in the LBD of STING. However, we could not explain the antiviral effects of CMA in the human system. A possibility is that CMA can inhibit viral replication. Nevertheless, CMA still could be the basis for the development of antiviral drugs that activate the STING pathway.

2.2.2. Contribution

The author of this thesis contributed equally to this publication. He cloned and purified the human and mouse STING constructs for the *in vitro* experiments. He performed the DSF assays. In addition, he crystallized the CMA-STING-LBD complex, collected the crystal diffraction data and solved the structure.

2.2.3. Paper

The following paper was published 2013 in The EMBO journal.

Species-specific detection of the antiviral small-molecule compound CMA by STING

Taner Cavlar*, Tobias Deimling*, Andrea Ablasser, Karl-Peter Hopfner & Veit Hornung

EMBO J. 2013 March 18; **32**(10): 1440 - 50

*: equal contribution

The EMBO Journal (2013) 32, 1440–1450
www.embojournal.org

Species-specific detection of the antiviral small-molecule compound CMA by STING

Taner Cavlar[1,4], Tobias Deimling[2,4], Andrea Ablasser[1], Karl-Peter Hopfner[2,3] and Veit Hornung[1,*]

[1]Unit for Clinical Biochemistry, Institute for Clinical Chemistry and Clinical Pharmacology, University Hospital, University of Bonn, Bonn, Germany, [2]Gene Center and Department of Biochemistry, Ludwig-Maximilians-University Munich, Munich, Germany and [3]Center for Integrated Protein Sciences, Munich, Germany

Extensive research on antiviral small molecules starting in the early 1970s has led to the identification of 10-carboxymethyl-9-acridanone (CMA) as a potent type I interferon (IFN) inducer. Up to date, the mode of action of this antiviral molecule has remained elusive. Here we demonstrate that CMA mediates a cell-intrinsic type I IFN response, depending on the ER-resident protein STING. CMA directly binds to STING and triggers a strong antiviral response through the TBK1/IRF3 route. Interestingly, while CMA displays extraordinary activity in phosphorylating IRF3 in the murine system, CMA fails to activate human cells that are otherwise responsive to STING ligands. This failure to activate human STING can be ascribed to its inability to bind to the C-terminal ligand-binding domain of human STING. Crystallographic studies show that two CMA molecules bind to the central Cyclic diguanylate (c-diGMP)-binding pocket of the STING dimer and fold the lid region in a fashion similar, but partially distinct, to c-diGMP. Altogether, these results provide novel insight into ligand-sensing properties of STING and, furthermore, unravel unexpected species-specific differences of this innate sensor.

The EMBO Journal (2013) **32,** 1440–1450. doi:10.1038/emboj.2013.86; Published online 19 April 2013
Subject Categories: immunology
Keywords: antiviral activity; innate immunity; STING; type I interferon

Introduction

The innate immune system operates to sense microbial infection. To this effect, it expresses a variety of so-called pattern-recognition receptors (PRRs) that are able to sense certain highly conserved microbial patterns, known as microbe-associated molecular patterns (MAMPs). Upon MAMP sensing, cells of the innate immune system elicit certain effector functions that are geared at eliminating the invading pathogen. As such, cytokines play an important role in orchestrating subsequent cellular immune responses or by inducing antimicrobial effector functions in non-immune cells. In this regard, cytokines of the type I interferon (IFN) family play a pivotal role in eliciting antiviral, but also antibacterial, effector functions. Soon after the discovery of type I IFNs by Isaacs and Lindenmann (1957), many research groups tried to delineate the mechanisms of its induction. Early on, it was already noted that nucleic acid preparations derived from viruses or enzymatic preparations were able to trigger potent type I IFN responses. One of the most potent triggers that emerged from these studies was the double-stranded RNA polynucleotide mimic poly(I:C) (Isaacs *et al*, 1963; Rotem *et al*, 1963), but other polynucleotide preparations, including double-stranded DNA (dsDNA), were also reported to initiate antiviral immunity by eliciting type I IFN responses. Moreover, in the early 1970s, several groups tried to develop small-molecule compounds with oral bioavailability that were able to trigger type I IFN responses, thereby blocking viral replication. In the course of these studies, the first small-molecule compound reported was the tricyclic compound tilorone (2,7-bis(2-diethylaminoethoxy)fluoren-9-one), which exhibited broad antiviral activities against many viruses (Krueger and Mayer, 1970; Mayer and Krueger, 1970). Subsequently, additional heterocyclic compounds were reported to induce type I IFNs, including various quinoline, anthraquinone and acridine derivatives. Another structurally related molecule that even surpassed most of these compounds in antiviral activity was 10-carboxymethyl-9-acridanone (CMA), discovered by Grunberg and colleagues, 1976. CMA was shown to harbour potent antiviral activity and this could be mainly ascribed to its ability to induce type I IFN production (Taylor *et al*, 1980b; Kramer *et al*, 1981; Storch and Kirchner, 1982; Brehm *et al*, 1986; Storch *et al*, 1986).

However, while most of these small-molecule compounds showed excellent type I IFN induction and antiviral activities in rodents (Kramer *et al*, 1976; Taylor *et al*, 1980a,b), these promising results failed to translate into the human system. Tilorone, for example, showed no type I IFN induction in the human system, both upon systemic or topic administration (Kaufman *et al*, 1971), and at the same time, research on CMA was abandoned by most groups. Nevertheless, CMA is currently widely distributed and applied in Russia for antiviral therapy, including hepatitis B virus, hepatitis C virus, HIV or herpes simplex virus infection (Silin *et al*, 2009).

Three independent research approaches have led to the discovery of STING as a protein that strongly induced type I IFN production upon overexpression (Ishikawa and Barber, 2008; Zhong *et al*, 2008; Sun *et al*, 2009). Through its N-terminal four-pass transmembrane region STING is tethered to the ER, whereas its C-terminal region faces the cytoplasmic lumen. STING-deficient cells show a profound defect in sensing DNA viruses, and also synthetic DNA ligands are strongly blunted in their capacity to induce

---

*Corresponding author. Unit for Clinical Biochemistry, Institute for Clinical Chemistry and Clinical Pharmacology, University Hospital, University of Bonn, Sigmund-Freud-Strasse 25, 53127 Bonn, Germany.
Tel.: +49 228 287 51203; Fax: +49 228 287 51201;
E-mail: veit.hornung@uni-bonn.de

[4]These authors contributed equally to this work.

Received: 25 October 2012; accepted: 18 March 2013; published online: 19 April 2013

pro-inflammatory gene expression in STING-deficient cells (Ishikawa *et al*, 2009). At the same time, bacteria that replicate in the cytoplasm, such as *Listeria monocytogenes*, trigger type I IFN production in a STING-dependent manner (Ishikawa *et al*, 2009). This finding was initially explained by bacterial DNA being sensed in the cytoplasm in a STING-dependent fashion. However, it turned out that type I IFN production by *L. monocytogenes* could be mainly ascribed to the cytoplasmic presence of the bacterial quorum-sensing molecule cyclic diadenylate (c-diAMPs; Woodward *et al*, 2010). Indeed, bacteria-derived c-diAMP or cyclic diguanylate (c-diGMP) had already been described as potent triggers of innate immune responses (Karaolis *et al*, 2007a,b; McWhirter *et al*, 2009). Both c-diAMP and c-diGMP function as quorum-sensing molecules in bacteria, regulating cell motility, biofilm formation and bacterial growth. Surprisingly, STING turned out to be the direct sensor for c-diGMP and c-diAMP, with its C-terminal part harbouring the binding domain (Burdette *et al*, 2011). In addition, most recently, it was discovered that cytosolic DNA sensing triggers the formation of a novel second messenger, cyclic GMP–AMP, which in turn binds to and activates STING (Sun *et al*, 2013; Wu *et al*, 2013). This finding reconciles the puzzling concept of STING serving as a sensor for microbial cyclic dinucleotides and DNA at the same time.

Several groups have recently been able to solve the crystal structure of c-diGMP binding to the C-terminal domain of STING (Huang *et al*, 2012; Ouyang *et al*, 2012; Shang *et al*, 2012; Shu *et al*, 2012; Yin *et al*, 2012). These studies have shown that the ligand-binding domain (LBD) of STING is present as a preformed dimer that forms a V-shaped structure harbouring a binding pocket for one c-diGMP molecule. While ligand-binding does not induce a major conformational change of the LBD of STING, a plausible model of activation suggests that the C-terminal tail (CTT) of STING is displaced from the pocket upon binding, so that it can interact with TBK1 (Yin *et al*, 2012). TBK1 subsequently leads to the phosphorylation of IRF3 and thereby induces transcription of antiviral genes.

In this study, we elucidate the molecular mechanism of CMA, a long-known small-molecule inducer of antiviral responses. In the murine system, CMA was found to be a potent activator of type I IFN production, yet in the human system it failed to elicit detectable antiviral responses. CMA activity depends on STING, and non-responsive human cells can be conferred responsive by overexpressing murine STING or a chimeric version of human STING that contains the LBD of murine STING. Differential scanning fluorimetry (DSF) studies using the LBDs of murine and human STING furthermore indicate that unresponsiveness of human STING to CMA is due to a lack of ligand binding. The crystal structure of CMA bound to murine STING shows that two CMA molecules bind the central c-diGMP-binding cavity in a fashion representing the inherent c-diGMP molecule symmetry, whereas differences are found in the folding of the lid region. While the structural studies cannot explain the species specificity in CMA detection, we provide additional data that suggest a differential involvement of the STING ‘lid region’ in CMA versus c-diGMP recognition. Altogether, these data reveal important insight into the species-specific recognition of a novel class of STING ligands.

Results

CMA is a potent trigger of the type I IFN response in murine macrophages

Intrigued by earlier publications on the antiviral activity of CMA (Figure 1A), we were interested in the molecular mechanisms of its activation, especially its putative receptor and associated signalling routes. We first assessed its ability to induce type I IFN production in macrophages in comparison to defined ligands for various PRR systems. To this effect, LPS (TLR4), transfected poly(I:C) (TLR3 and MDA5), transfected 5′triphosphate RNA (pppRNA/RIG-I) and transfected 45mer dsDNA (ISD/STING) were used. All ligands were tested at their previously determined optimal concentrations, and as readouts we assessed phosphorylation of IRF3, transcription of Ifnb mRNA, IFNβ protein levels and transactivation of the IFNβ promoter, using pIFNβ-Luc reporter mouse macrophages (Lienenklaus *et al*, 2009). Indeed, CMA induced robust IRF3 phosphorylation that was followed by strong Ifnb mRNA induction and translation (Figure 1B–E). CMA-mediated IFNβ production reached peak levels already 4 h after stimulation, even exceeding LPS in its readiness to trigger IFNβ synthesis (Figure 1F). Of note, for the various ligands studied, IRF3 phosphorylation, Ifnb mRNA induction and its translation did not go in parallel, which is attributable to the fact that additional ligand properties can play important roles in regulating transcription and translation of this key cytokine. Assessing other IRF3-dependent targets, such as IP-10, corroborated the notion of CMA being a potent trigger of antiviral immunity, yet NF-kB target cytokines (e.g., IL-6) were induced to a lesser degree in CMA-stimulated cells (Figure 1G,H). In line with this observation, assessing IRF3, MAPK and NF-kB activation following CMA stimulation showed that CMA led to synchronous and rapid activation of all these signalling cascades, with a predominant IRF3 signature. LPS, on the other hand, showed stronger NF-kB and MAPK activation, with a slight delay in IRF3 phosphorylation (Figure 1I). At the functional level, CMA blocked viral gene expression of VSV-based replicon particles in a dose-dependent manner (Supplementary Figure S1; Berger Rentsch and Zimmer, 2011). Altogether, these studies indicated that CMA is a rapid and potent inducer of antiviral immune responses in murine macrophages.

CMA-dependent type I IFN production requires STING

While most immortalized cell lines did not respond to CMA (e.g., immortalized murine embryonic fibroblasts (MEFs), HEK 293T cells, HeLa cells), early-passage MEFs could be stimulated with CMA. This allowed us to use MEFs from TBK1/IKKε-deficient mice, to determine the role of these canonical kinases in IRF3 phosphorylation. As expected, IRF3 phosphorylation was completed blunted in MEFs deficient for TBK1 and IKKε (Supplementary Figure S2A). We next assessed the role of TLR signalling pathways using macrophages from mice deficient in MyD88 or TRIF. Neither MyD88 nor TRIF were required for CMA sensing, which ruled out an involvement of TLRs (Supplementary Figure S2B–D). At the same time, absence of MAVS, the shared signalling adapter of RIG-I and MDA5, had no impact on CMA-mediated antiviral immunity (Supplementary Figure S2E,F). We next went on to study the role of STING

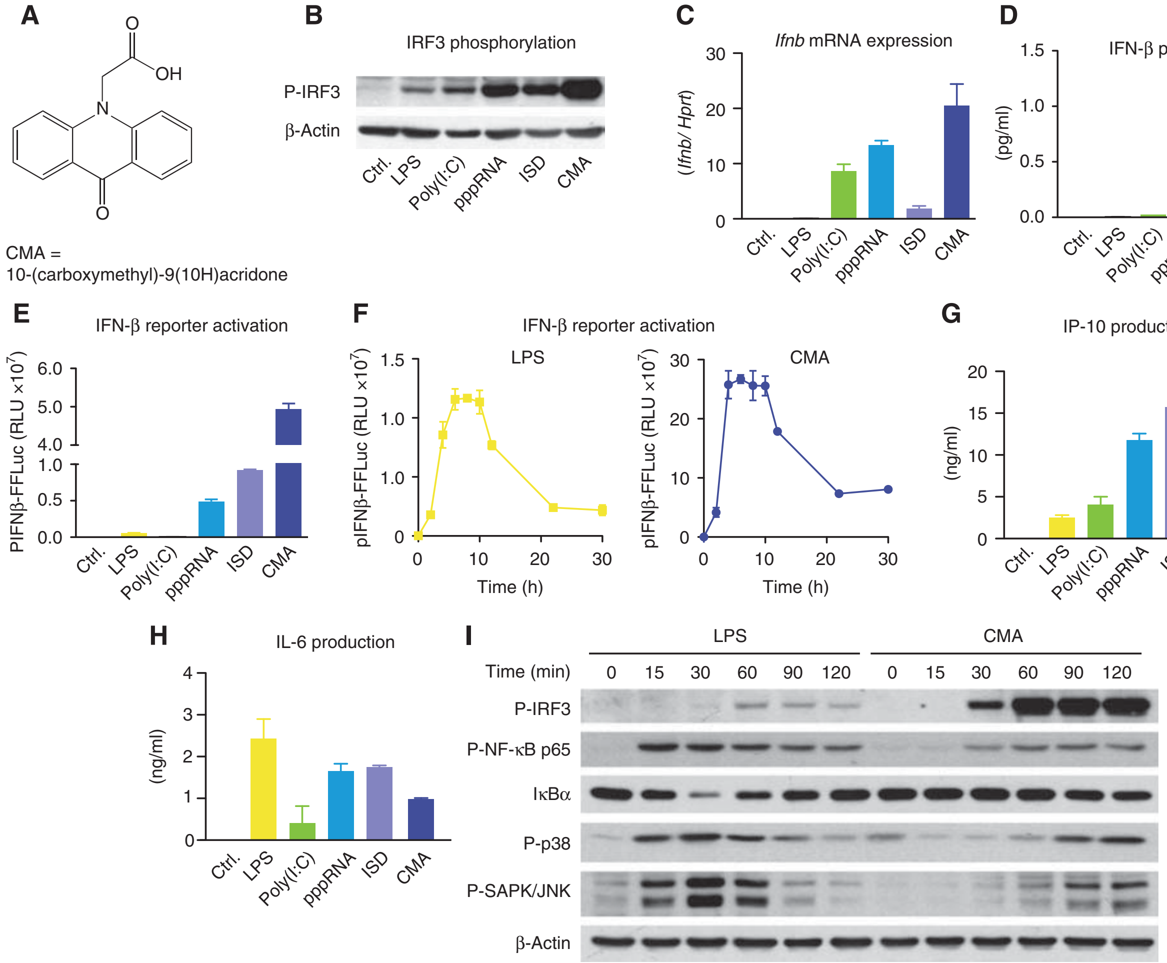


Figure 1 CMA strongly induces type I IFN in primary mouse macrophages. (**A**) The chemical structure of CMA is depicted. (B–E and G–H) Bone marrow-derived macrophages were transfected with poly(I:C), pppRNA and ISD, or stimulated with LPS or CMA (500 µg/ml). (**B**) After 2 h, cells were collected and subjected to SDS–PAGE, and western blotting for phospho-IRF3 (P-IRF3) was performed. (**C**) Four hours after stimulation, transcription of the IFNβ gene (*Ifnb*) was assessed by quantitative RT–PCR, with normalization to *HPRT1*. (**D**) Eighteen hours after stimulation, IFNβ was measured in the supernatants by enzyme-linked immunosorbent assay (ELISA). (**E**) pIFNβ-firefly-luciferase (pIFNβ-FFLuc) macrophages were stimulated as indicated. After 18 h, cells were lysed with passive lysis buffer, and FFLuc activity was measured in the lysates. (**F**) pIFNβ-FFLuc macrophages were stimulated with LPS or CMA (500 µg/ml). Luciferase activity was assessed at the indicated time points (in hours). (**G**,**H**) Eighteen hours after stimulation, IP-10 (**G**) and IL-6 (**H**) were measured in the supernatants by ELISA. (**I**) Bone marrow-derived macrophages were stimulated with LPS or CMA (500 µg/ml). Total protein was collected at indicated time points (in minutes) after stimulation and was assessed for P-IRF3, phospho-NF-κB-p65 (P-NF-κB p65), IkBα, phospho-p38 (P-p38) or phospho-SAPK/JNK (P-SAPK/JNK). Representative results out of three independent experiments are depicted. Source data for this figure is available on the online supplementary information page.

in CMA-triggered type I IFN production. To this effect, we used macrophages from a mutant mouse strain, Goldenticket, that harbours a missense mutation (I199N) in STING (Sauer *et al*, 2011). Indeed, macrophages from these mice showed no detectable activation of IRF3 or NF-kB (Figure 2A) and also a complete absence of cytokine production (Figure 2B). In addition, the antiviral activity elicited by CMA treatment was completely abolished when STING-deficient macrophages were stimulated (Supplementary Figure S3). In line with the critical requirement of STING in CMA sensing, a human HEK 293T cell line engineered to stably express murine STING responded to CMA, with a strong increase in pIFNβ or pELAM reporter activity (Supplementary Figure S4). As expected, this cell line also showed a strong gain-of-function signal with regards to c-diGMP sensing, which is usually non-functional in unmodified HEK 293T cells. Altogether, these results indicated that STING was required and also sufficient for CMA recognition.

CMA fails to activate human STING

Despite the fact that CMA had been reported to induce type I IFN responses in human cells (Silin *et al*, 2009), we were unable to elicit type I IFN production in various cell types of the human system. Human PBMCs readily responded to 5′pppRNA, poly(I:C), DNA, c-diGMP and LPS stimulation, yet CMA failed to induce detectable cytokine responses, even at high doses (Figure 3A–C). A similar picture was seen when primary human fibroblasts were used (Figure 3 D,E). In line with this finding, when we transiently transfected HEK 293T cells with a human STING construct, no CMA response could be detected, despite the fact that overexpression of human STING rendered 293T cells sensitive to c-diGMP (Figure 4A). At the same time, transient overexpression of murine STING made HEK 293T cells responsive towards both c-diGMP and CMA (Figure 4B). The fact that human STING displayed c-diGMP-dependent signalling capacity in HEK 293T cells suggested that the C-terminal

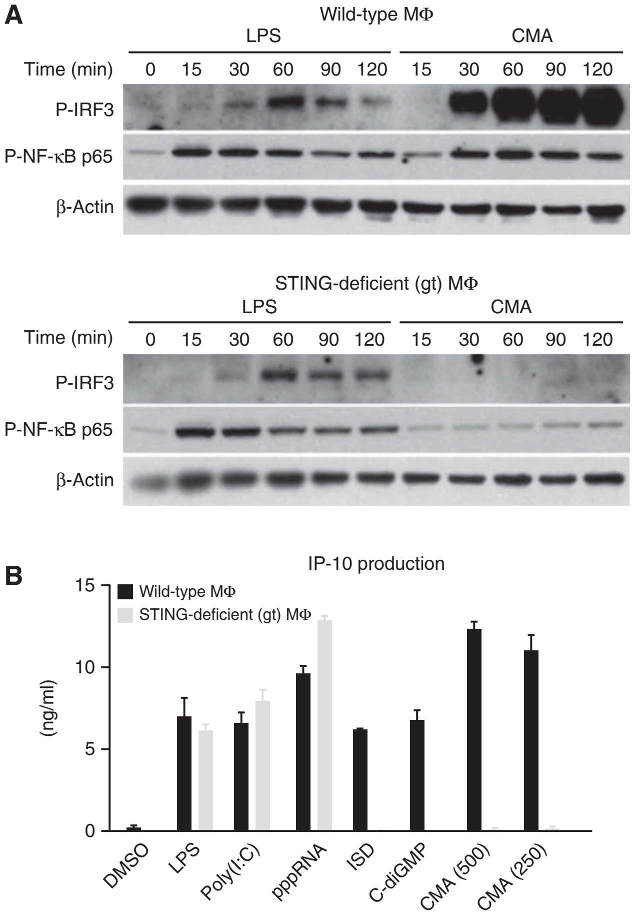


Figure 2 Loss of STING leads to complete abrogation of CMA-induced cytokine production. (**A**) Bone marrow-derived wild-type and STING-deficient macrophages were stimulated with LPS and CMA (500 μg/ml). Phospho-IRF3 (P-IRF3) and phospho-NF-κB-p65 (P-NF-κB p65) were assessed at indicated time points (in minutes) after stimulation. (**B**) Bone marrow-derived wild-type and STING-deficient macrophages were transfected with poly(I:C), pppRNA, ISD and c-diGMP, or stimulated with LPS and CMA (500 or 250 μg/ml) in duplicates. Supernatants were collected after 18 h, and IP-10 levels were measured by enzyme-linked immunosorbent assay. Representative results out of two independent experiments are depicted, whereas data are presented as mean values + s.e.m. Source data for this figure is available on the online supplementary information page.

LBD of STING could be held responsible for the insensitivity towards CMA. To address this question, we constructed chimeric STING constructs, in which the N- and C-terminal domains of the murine or human STING were exchanged: mmSTING(1-137)-hSTING(139-379) or hSTING(1-138)-mmSTING(138-378) (Figure 4C,D). Testing these constructs in HEK 293T cells revealed that as long as the murine LBD was present (hSTING(1-138)-mmSTING (138-378); Figure 4D), CMA was able to trigger type I IFN production. Altogether, these results indicated that unresponsiveness of human cells towards CMA could be explained by species-dependent differences in the LBD of STING.

Using the murine STING construct, we next wanted to address the role of functionally relevant point mutations for the recognition of CMA. The previously described null-mutant I199N most likely perturbs the structure of STING and thereby abolishes c-diGMP binding and signalling (Burdette *et al*, 2011). In line with this notion, overexpression of mmSTING-I199N showed no pIFNβ-Luc transactivation upon CMA- or c-diGMP-mediated stimulation (Supplementary Figure S5A,B). Interestingly, mmSTING-R231A, a mutant that has previously been described to completely blunt c-diGMP-mediated STING activation despite binding, showed normal activity upon CMA stimulation (Supplementary Figure S5C). These data demonstrated that c-diGMP sensing could be dissociated from CMA recognition at the receptor level.

CMA does not bind the human STING LBD

The failure of CMA to activate human STING could be explained by several scenarios. Foremost, we wanted to rule out the possibility that human STING does not bind to CMA. To address this question, we expressed and purified the LBD of murine and human STING in *E. coli* and tested its ability to associate with various ligands, using DSF. DSF indirectly assesses the association of a low molecular weight compound to a purified protein by measuring the stability of a protein–compound complex as a function of temperature (Niesen *et al*, 2007). As indicated by a robust thermal shift, the LBD of murine STING associated with its known ligands c-diGMP and c-diAMP, and also binding to CMA, could be observed (Figure 5A). At the same time, also the LBD of human STING displayed association with c-diGMP and c-diAMP, as indicated by a dose-dependent shift in thermal stability. However, addition of CMA had no detectable impact on the thermal stability of human STING, indicating that CMA does not bind human STING (Figure 5B). Despite high homology, murine and human STING show least conservation in their CTT region. As such, we additionally wanted to rule out that this part of STING was required for CMA binding. To this effect, we generated truncated versions of the murine STING LBD lacking the CTT. However, testing this variant showed similar binding properties as the full-length LBD, indicating that the CTT is not required for ligand binding by STING (Supplementary Figure S6).

The crystal structure shows a c-diGMP-like STING interaction for CMA

Recently, five independent groups determined the crystal structure of the LBD of human STING bound to c-diGMP (Huang *et al*, 2012; Ouyang *et al*, 2012; Shang *et al*, 2012; Shu *et al*, 2012; Yin *et al*, 2012). In these studies it was shown that the LBD of STING constitutes a preformed dimer that is stabilized through homotypic interaction at a hydrophobic interface. This STING dimer forms a V-shaped pocket, in which one c-diGMP molecule is buried at its bottom. To see how CMA binds murine STING, we crystallized murine STING LBD in the presence of CMA and determined the crystal structure to 2.75 Å resolution (Supplementary Table S1). The final model comprises two STING LBDs along with two well-resolved CMA molecules (Figure 6A,B). Murine STING LBD forms a dimer with high overall structural similarity to the previously determined human STING LBDs (Supplementary Figure S7). Two CMA molecules are located in the deep central c-diGMP-binding pocket at the dimer interface. The acridone ring moieties of both CMAs partially stack to each other (∼4 Å distance) in a parallel, laterally shifted orientation, resembling the twofold symmetry of c-diGMP. They are situated near the phosphate–ribose

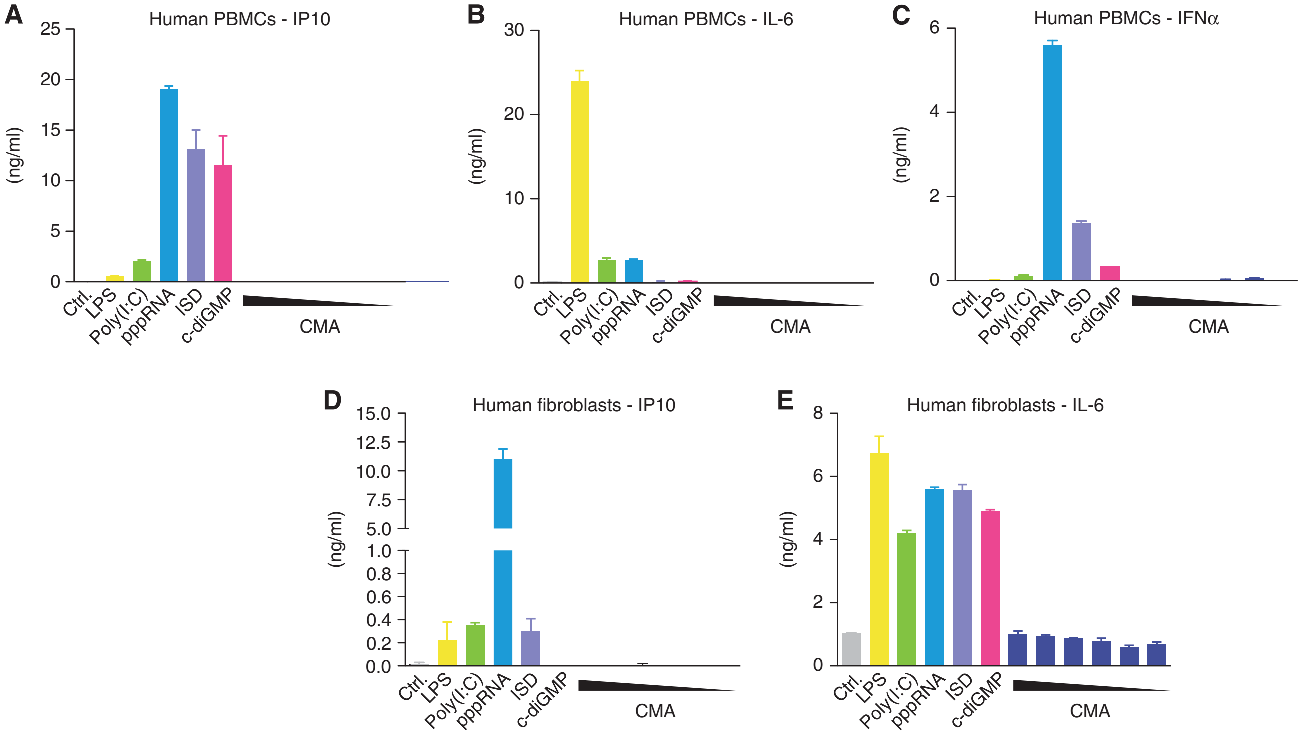


Figure 3 Human cells do not respond to CMA. (**A–E**) Primary human PBMCs and fibroblasts were transfected with poly(I:C), pppRNA, ISD and c-diGMP, or stimulated with LPS and decreasing concentrations of CMA (4000—125 μg/ml in twofold dilutions). IP-10, IL-6 and IFNα levels in the supernatants of stimulated PBMCs (**A–C**) and fibroblasts (**D**,**E**) were determined by enzyme-linked immunosorbent assay. Representative results out of three independent experiments are depicted.

binding site for c-diGMP, but due to the flat shape can wedge deep into the helical bundle core of the LBD dimer. As a result, CMA binds directly to Thr226 at the bottom of the ligand-binding pocket, whereas c-diGMP binds via a water molecule (Figure 6C,D).

Significantly, we find two well-ordered lid domains in a four-stranded antiparallel β-sheet that closes the CMA/c-diGMP-binding pocket (Figure 6A,B). The lids are generally in a similar conformation than the folded lids in the human STING–c-diGMP complex reported by Huang *et al* (2012) (PDB-code: 4F5D), but there are also differences. The carboxymethyl groups of the CMA moieties face and stabilize the conformation of the lid by forming salt bridges to both Arg237 residues. The equivalent human Arg238 residues stack between the two c-diGMP guanine moieties, thereby stabilizing the lid. The space of the guanine moieties is unoccupied by CMA. As a result, murine Arg237 and Tyr167, plus Tyr240, directly stack instead of sandwiching guanine as observed in the human c-diGMP complex. The direct stacking induces or enables further closure of the V-shaped binding cleft compared to the c-diGMP-bound conformation, which may contribute to active signalling. However, the tips of the lids fold differently in the presence of CMA and c-diGMP. In particular, Arg231 (human Arg232), which binds the phosphates of c-diGMP via a magnesium ion or water molecule (Huang *et al*, 2012), is pushed to the surface due to steric hindrance from the acridone ring. In summary, CMA binds to the c-diGMP-binding pocket, with two stacked CMA molecules mimicking the symmetric c-diGMP. CMA induces a conformation in the LBD dimer that is similar, if not more pronounced than the proposed signalling conformation of human LBD, with folded lids in the presence of c-diGMP, but also results in a somewhat altered lid conformation. Altogether, the structure can explain why CMA activates murine STING.

Discussion

In line with early reports on the antiviral activity of CMA (Storch and Kirchner, 1982; Storch *et al*, 1986), our data confirm the extraordinary type I IFN-inducing capacity of CMA in the murine system. Further analysis of signalling cascades involved in IFNβ transactivation show a quick and strong phosphorylation of IRF3, whereas the activation of the NF-κB and MAPK pathways is slightly delayed and less prominent compared to TLR signalling. This observation is also reflected by the cytokine profile, as CMA induces high levels of IFNβ as opposed to IL-6. Studies using knockout cells revealed that CMA triggers antiviral immunity cells through the STING–TBK1–IRF3 route, which could be further corroborated by gain-of-function experiments in cells devoid of functional mmSTING.

Unexpectedly, human cells showed complete unresponsiveness towards CMA, although c-diGMP-dependent STING activity was observed. We could ascribe this to species-dependent differences within the C-terminal LBD of STING, as HEK 293T cells expressing a chimeric STING protein equipped with the murine LBD readily responded towards CMA. Further studies could pinpoint the failure of human STING to respond to CMA, to the inability of the human LBD to form a stable complex. Of note, these experiments do not formally proof that human STING cannot bind to CMA, yet in light of the functional data this appears to be the most likely scenario. The species-specific activation of murine STING by

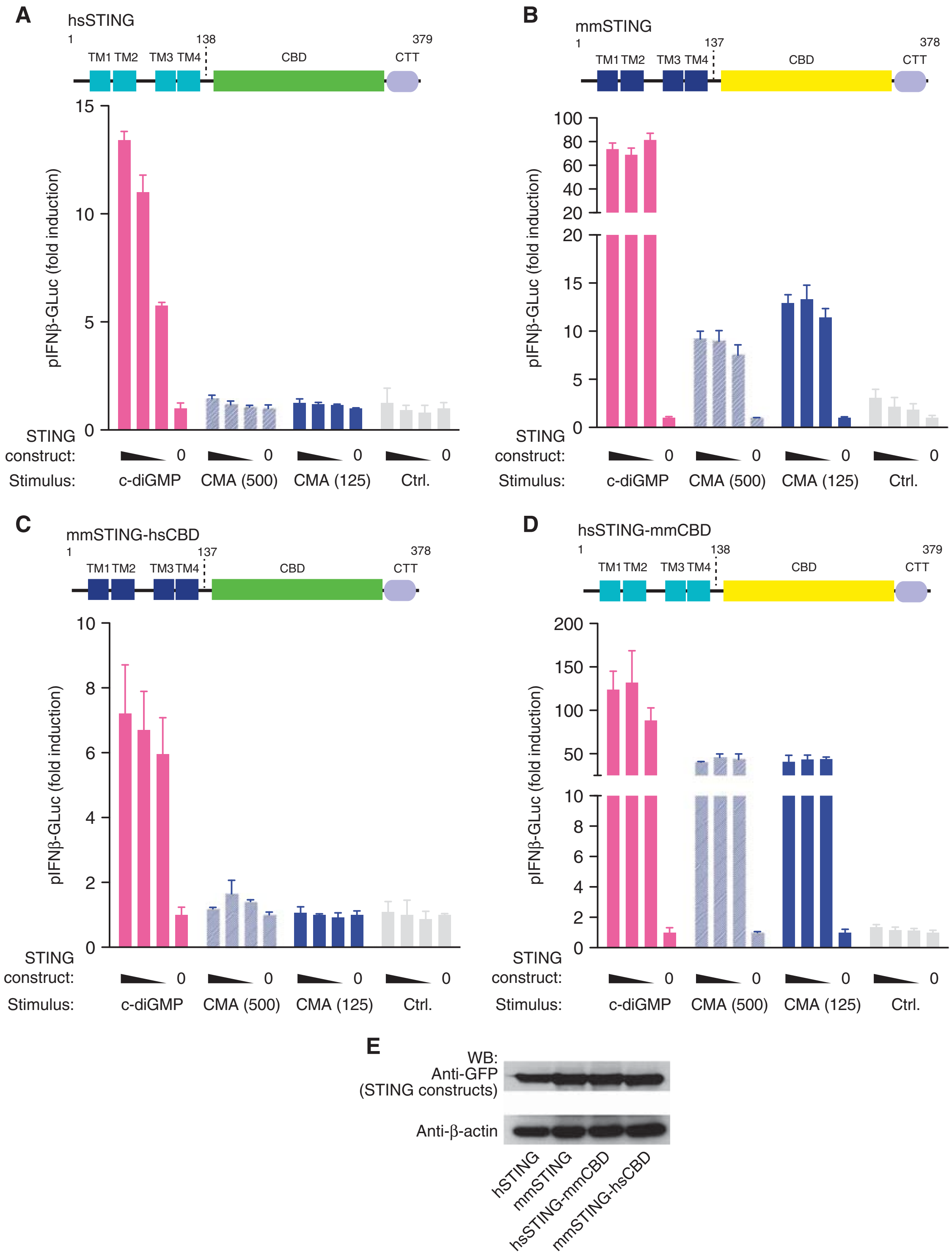


Figure 4 Species-specificity CMA activity is determined by the C-terminal LBD of STING. (**A–D**) 293T cells were transiently transfected with the indicated STING constructs (25, 12.5, 6.25 and 0 ng), whereas 12.5 ng of pIFNα-GLuc reporter plasmid were included. For titrations, an empty pCI vector served as a stuffer to obtain 200 ng total plasmid DNA. After 24 h, cells were transfected with c-diGMP or stimulated with CMA (500 and 125 μg/ml). Luciferase activity was measured after an additional period of 24 h in the supernatant, and data were normalized to the condition without STING overexpression. Plasmids coding for full-length hSTING (**A**), mmSTING (**B**), mmSTING(1-137)-hSTING(139-379) (**C**) and hSTING(1-138)-mmSTING(138-378) (**D**) were tested. (**E**) Expression of the above described constructs was studied in 293T cells 24 h after transfection (200 ng per 96-well plate) using western blot, whereas β-actin served as a loading control. Representative results out of three independent experiments are depicted. Source data for this figure is available on the online supplementary information page.

CMA is surprising, especially given the fact that several reports on antiviral activity in the human system both *in vitro* and *in vivo* exist (Vershinina *et al*, 2002; Zarubaev *et al*, 2003). One possibility is that these antiviral activities are due to direct inhibition of viral replication, independent of the innate immune response (e.g., inhibition of viral

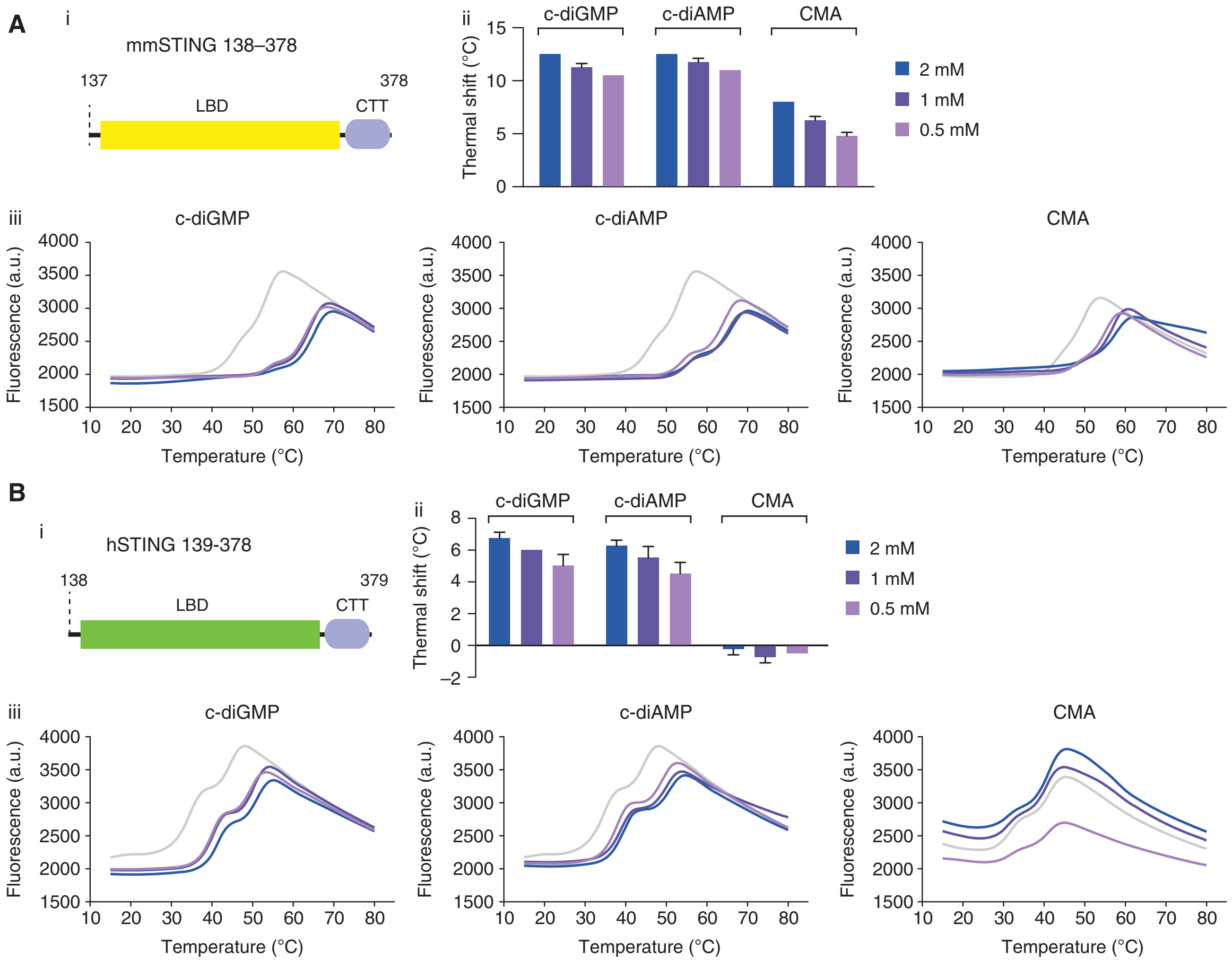


Figure 5 DSF implicates direct binding of CMA to murine, but not to human STING. (**A**,**B**) The interaction of STING with c-diGMP, c-diAMP and CMA were analysed by thermal shift assay. Purified murine STING (**A**) and human STING (**B**) were tested with different concentrations of c-diGMP/c-diAMP/CMA; (i) Schematic views of the protein domains used for binding studies are shown; (ii) thermal shifts of (iii) fluorescence intensity versus temperature are shown. Representative results out of two independent experiments are depicted for the temperature curves (ii), whereas mean values + s.e.m. out of two independent experiments are depicted for the thermal shift graphs (iii).

polymerases). At the same time, it is conceivable that higher local concentrations of CMA can be achieved during systemic application or that other cell types are activated, both leading to a STING-dependent type I IFN response. Nevertheless, our results clearly question a predominant role for type I IFN induction by CMA in the human system.

FAA (flavone-8-acetic acid) and DMXAA (5,6-dimethyl-9-oxo-9H-xanthen-4-yl)-acetic acid) are two additional tricyclic small-molecule compounds that have been reported to trigger potent type I IFN responses in the murine system (Hornung *et al*, 1988; Futami *et al*, 1991; Perera *et al*, 1994; Roberts *et al*, 2007). Both compounds have been pursued for their potent antitumour activity in murine tumour models, and for both compounds it was shown that activation of the immune system plays a pivotal role in therapeutic activity (Ching and Baguley, 1987; Hornung *et al*, 1988; Pang *et al*, 1998). FAA, however, failed to induce type I IFN responses in human cells (Futami *et al*, 1991) and, moreover, it failed to display antitumour activity in clinical trials (Bibby and Double, 1993). In parallel, despite strong activity in the murine system, DMXAA by itself also showed limited immunostimulatory capacity in human cells (Patel *et al*, 1997; Philpott *et al*, 2001; Gobbi *et al*, 2006). At the same time, DMXAA in combination with a platinum-based chemotherapy did not show any efficacy in a large phase III clinical trial for the treatment of non-small cell lung cancer (Lara *et al*, 2011). Most recently, it has been reported that DMXAA triggers type I IFN responses in a STING-dependent fashion (Brunette *et al*, 2012; Prantner *et al*, 2012). In these reports, only murine cells were tested and direct binding of DMXAA to STING was not assessed. Nevertheless, given the similarity of CMA and DMXAA at the molecular level, we would expect that DMXAA also directly engages STING in a species-specific manner. As such, it appears likely that the failure of DMXAA in the human system can be attributed to its inability to trigger STING activation.

At the structural level, the species-specific recognition of CMA is unlikely to be attributable to the binding pocket itself. This region, which is mainly formed by α-helices, α1 and α3, is more or less invariant between the human and the murine system, with all direct interactors of CMA being identical. It is possible that more subtle structural differences prevent

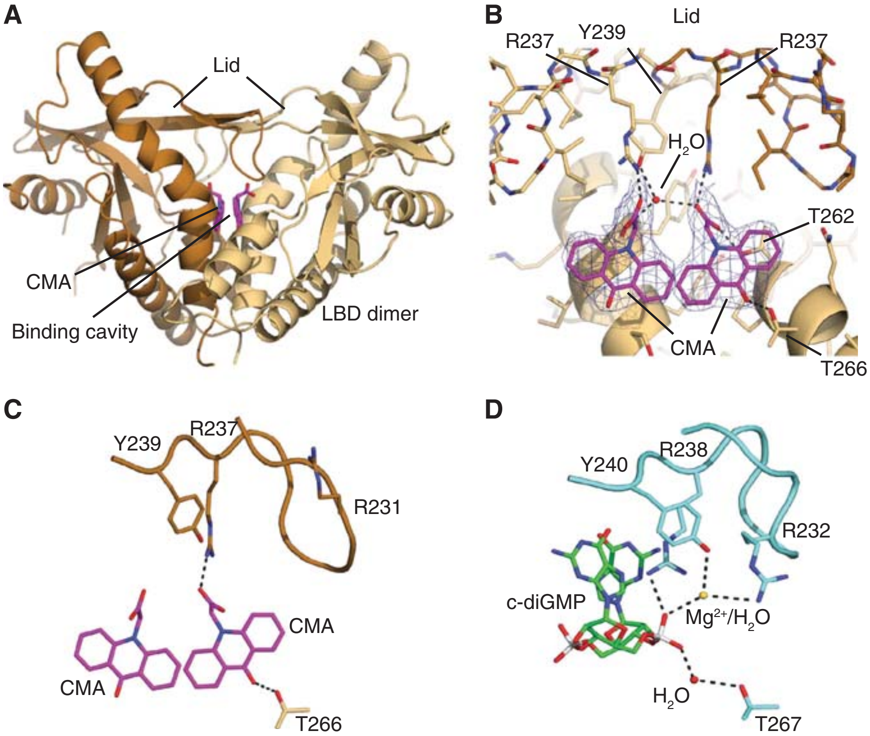


Figure 6 Structural basis for CMA recognition. (**A**) Ribbon model of the mouse STING dimer (light and dark brown) with highlighted secondary structure. The two bound CMA molecules are shown as magenta stick models. (**B**) Close-up view of the CMA-binding site with superimposed 2mFo-DFc electron density (blue; contoured at 1.4). One STING protomer is shown in light brown. For the other protomer (brown), only the lid is displayed. Folding of the lid via hydrogen bonds to Arg237 and Tyr239 suggests how CMA activates mouse STING. (**C**,**D**) Side-by-side comparison of CMA bound to mouse STING (**C**) and c-diGMP bound to human STING (**D**) showing selected interactions. CMA folds the lid differently from c-diGMP, due to steric clash with Arg231 (human Arg232), which binds c-diGMP via a magnesium ion/water molecule.

binding of CMA to human STING. An alternative, probably more likely, explanation concerns loops β2 and β3 of the LBD, which are critically involved in the coordination of c-diGMP. While flexible in the apo form of human STING, this loop region undergoes a conformational rearrangement upon c-diGMP binding, thereby functioning as a ‘lid’ that keeps c-diGMP within the binding pocket (Huang *et al*, 2012). Two conserved residues within the tip of this loop (human: Arg238, Tyr240) have been shown to interact with c-diGMP, and point mutagenesis data support the critical role of the loop region in c-diGMP recognition. In support of a similar activation mechanism, CMA binds Arg237 of murine STING. Interestingly, the previously reported R231A mutation within this loop region impairs c-diGMP-mediated IFN induction even though DNA- and CMA-dependent IFN induction are not affected (Burdette *et al*, 2011). These data suggest that the loop region is differently involved in CMA sensing, as opposed to c-diGMP recognition. Consistent with the mutational data, Arg231 is not involved in recognition of CMA, and the tip of the lid, where this conserved residue is located, is folded differently compared to the human STING bound to c-diGMP (Figure 6C). As such, species-dependent differences that may stabilize or allow this altered conformation could account for the observed unresponsiveness of hSTING towards CMA. Yet, additional studies will be required to pinpoint this phenomenon to an exact structural determinant.

This is the first report of a direct STING ligand that is not a cyclic dinucleotide. Most intriguingly, CMA is not sensed as a monomeric ligand within the ligand-binding pocket of the STING dimer; yet, two molecules positioned in a rotational symmetry are required to activate STING. We assume that other reported tricyclic small-molecule antivirals (e.g., tilorone, DMXAA, *etc*) are also sensed as a dimeric ligand in a fashion similar to CMA. Moreover, it is tempting to speculate that beyond these synthetic compounds, natural ligands exist that also follow a CMA-like recognition mode. In this regard, several bioflavonoids reported to exert antiviral activity might be interesting candidates. At the same time, the novel class of STING ligands described here might open novel avenues to develop compounds that block STING activation, with favorable drug-like properties. Given the likely involvement of STING in sensing endogenous DNA in the course of sterile inflammatory conditions (Ahn *et al*, 2012; Gall *et al*, 2012), pharmacological targeting of STING might constitute a reasonable therapeutic venture.

Materials and methods

Reagents

Poly(I:C) and CMA were purchased from Sigma Aldrich. Ultra-pure LPS from *E. coli* was purchased from Invivogen. c-diGMP and c-diAMP were from BioLog. GeneJuice Transfection Reagent was from Novagen. Lipofectamine 2000 was from Life Technologies. Goat anti-rabbit-IgG-HRP, goat-anti-mouse-IgG-HRP and anti-β-actin-IgG-HRP were from Santa Cruz Biotechnology. Passive lysis buffer was purchased from Promega.

Plasmids

Expression plasmids coding for human or murine STING were cloned into pEFBOS coding for an N-terminal GFP or an N-terminal mCherry tag. Chimeric STING constructs, human STING (AA1–

138)–murine STING (AA138–378) and murine STING (AA1–137)–human STING (AA139–379) were generated by ligation-independent cloning (Aslanidis and de Jong, 1990; Schmid-Burgk *et al*, 2012). The STING mutants (I199N, R231A in pEFBOS-mCherry) were generated by point mutagenesis PCR. pCI-empty was used as a stuffer. All primer sequences used for cloning are available upon request.

Cell culture

PBMCs were isolated from whole blood of healthy donors. After Ficoll density-gradient centrifugation (Biochrom), red blood cells were lysed using lysing buffer (BD Biosciences). PBMCs were seeded with a density of 4×10^6 per ml at 100 μl in a 96-well plate containing RPMI supplemented with 10% (v/v) FCS, sodium pyruvate, penicillin and streptomycin (all from Life Technologies). 293T and primary human fibroblasts were cultured in DMEM supplemented with 10% (v/v) FCS, sodium pyruvate (all from Life technologies) and Ciprofloxacin (Bayer Schering Pharma). Primary macrophages were generated from mouse bone marrow cells that were cultured for 7 days in DMEM with the same additives as described above and 30% (v/v) L929 supernatant.

Immunoblotting

Primary macrophages were lysed in 1× Laemmli buffer and denatured at 95°C for 5 min. Probes were separated by 10% SDS–PAGE and transferred onto nitrocellulose membranes. Blots were incubated with anti-Phospho-IRF3 (number 4947), anti-Phospho-NF-kappaB-p65 (number 3033), anti-Phospho-p38 (number 4511), anti-IkappaB-alpha (number 4814) or anti-Phospho-SAPK/JNK (number 9255) from Cell Signaling Technology.

Transfection

For transfection experiments, primary macrophages were seeded with a density of 1×10^5 per ml. Cells were transfected with poly (I:C) (2 μg/ml), pppRNA (1.33 μg/ml), ISD (2 μg/ml) and c-diGMP (8.66 μg/ml) using Lipofectamine 2000 (Life Technologies), according to the manufacturer's instructions. LPS (200 ng/ml) and CMA were directly added to the medium. Human PBMCs were transfected as described above, at a density of 4×10^6 per ml, human fibroblasts at a density of 1.5×10^5 per ml. For western blot experiments, cells were lysed after 2 h, if not indicated otherwise. For cytokine assays, supernatants were collected after 18–20 h. For RNA isolation, cells were lysed after 4 h.

Enzyme-linked immunosorbent assay

Cell culture supernatants were assayed for human IP-10 (BD Biosciences), human IL-6 (BD Biosciences), human IFNα (eBioscience), mouse IFNβ (BioLegend), mouse IP-10 (R&D Systems) or mouse IL-6 (BD Biosciences), according to the manufacturer's instructions.

Quantitative real-time PCR analysis

RNA from macrophages was reverse transcribed using the RevertAid First Strand cDNA Synthesis kit (Fermentas) and quantitative PCR analysis was performed on an ABI 7900HT. All murine gene expression data are presented as relative expression to HPRT1. Primer sequences are available upon request.

Luminescence assays

pIFNβ-FFLuc macrophages (Lienenklaus *et al*, 2009) were stimulated as indicated and were lysed with passive lysis buffer 18–20 h after stimulation, to determine luminescence. For VSV*ΔG(Luc) replicon assays, macrophages were treated with ligands or supernatants as indicated and subsequently infected at an MOI of 10 with the VSV*ΔG(Luc) replicon virus particles. Firefly-luciferase activity was measured in the lysates using a 2104 EnVision Multilabel reader from Perkin-Elmer.

Plasmid overexpression experiments

293T cells were seeded with a density of 2×10^5 cells per ml at 100 μl 96-well plate. IFNβ (12.5 ng) promoter–reporter plasmid pIFNβ–GLuc or pELAM-GLuc, and STING constructs were transfected using GeneJuice according to the manufacturer's instructions. For titration experiments, empty pCI vector was used as a stuffer. After 24 h, the cells were transfected with c-diGMP or were stimulated with CMA. After further 20 h, GLuc activity was measured in the supernatants using coelenterazine as a substrate.

Cell viability assays

Cell viability was assessed using CellTiter-Blue (Promega) according to the manufacturer's instructions. The assay was performed immediately after (Gaussia) or before (Firefly) luciferase measurement.

Differential scanning fluorimetry

Thermal shift assays were carried out using a CFX 96TM Real-Time System (Biorad) for recording the fluorescence signal (HEX: Ex/Em: 450–490/560–580 nm) as a function of temperature. The temperature gradient was set from 15 to 80°C, with an increment of 0.5°C and incubation steps of 15 s. Each 20 μl reaction (buffer: 20 mM Tris, 150 mM NaCl, (10% DMSO for reactions with CMA), pH 7.5), with or without 2 mM/1 mM/0.5 mM of c-diGMP/c-diAMP/CMA contained 1 mg/ml of STING and a dilution of 1:500 of SYPRO Orange dye (Invitrogen).

Cloning of human and mouse STING constructs for protein expression

Human STING AA139–379 (R220H + H232R) was cloned from a human macrophage cDNA library into pET28-SUMO1-eGFP vector via *Bam*HI and *Not*I restriction sites. The mouse STING constructs AA138–378 and AA138–341 were cloned from a mouse lung tissue cDNA library into pET28-SUMO1-eGFP vector via *Age*I and *Not*I restriction sites. These plasmids were used to transform *E. coli* Rosetta (DE3) protein expression strain cells (Novagen).

Expression and purification of human and mouse STING constructs

For all used STING constructs, the expression and purification procedure was the following: *E. coli* Rosetta (DE3) cells were grown in 3 l of LB media supplemented with Kanamycin (50 mg/l) and Chloramphenicol (34 mg/l) at 37°C for 3 h, to $OD_{600} = 0.8$. Expression of N-terminal His_6-SUMO1-tagged STING was induced by adding IPTG (Roth) to a final concentration of 0.2 mM. Expression was done overnight at 18°C. Cells were collected by centrifugation and were resuspended in lysis buffer (50 mM Tris, 500 mM NaCl, 10 mM imidazole, 5% glycerol, 2 mM β-mercaptoethanol, pH 7.5) and lysed by sonication. The soluble His_6-SUMO1-STING was purified by Ni-affinity chromatography. The His_6-SUMO1-tag was removed by proteolytic cleavage with SenP2 protease during dialysis overnight (20 mM Tris, 150 mM NaCl, 3% glycerol, 2 mM β-mercaptoethanol, pH 7.5) and a second Ni-affinity chromatography step. To remove additional contaminants, a HiTrap Q FF (GE Healthcare) purification was applied. Finally, the STING containing flow-through was used in a HiLoad^{TM} 26/60 Superdex 75 prep grade (GE Healthcare) size-exclusion chromatography step (20 mM Tris, 150 mM NaCl, pH 7.5). Purified STING was concentrated (10–15 mg/ml) with a 10 kDa cut-off centrifugal concentrator device (Millipore) and was flash frozen in liquid nitrogen for storage (−80°C).

Crystallization of mmSTING AA149–348 with CMA

A protein solution of mmSTING (8 mg/ml) was saturated with CMA (by adding solid powder due to its low solubility in aqueous solution) and incubated on ice for 1 h. Prior to crystallization, the protein solution was centrifugated and filtered to remove solid CMA. Crystals were grown using the hanging drop vapour diffusion method with drops of 1:1 ratio protein:reservoir. The reservoir solution contained 0.1 M HEPES, pH 7, 1.65 M ammonium sulphate, 2% PEG 400 (v/v) and saturating amounts of CMA. mmSTING crystals appeared after ∼4 days at 20°C. For cryoprotection, the crystals were soaked in reservoir solution containing 12% (±)-1,3-butanediol (v/v) before flash freezing in liquid nitrogen.

Data collection and structure determination

X-ray diffraction data were collected at beamline X06SA at the Swiss Light Source (Villigen, Switzerland). Data processing was carried out using XDS (Kabsch, 2010). The structure was solved by molecular replacement with PHASER (McCoy *et al*, 2007) from the ccp4 package (Winn *et al*, 2011), using a mouse search model created from the human STING structure (4EMT (Shu *et al*, 2012)) by ClustalW2 alignment (Goujon *et al*, 2010; Larkin *et al*, 2007) and modelling with the SWISS-MODEL structure homology-modelling server (Peitsch, 1995; Arnold *et al*, 2006; Kiefer *et al*, 2009). The structure was refined by rounds of manual model building carried out with COOT (Emsley and Cowtan, 2004) and refinement using Phenix (Adams *et al*, 2010). The final structure with R/Rfree = 21/23.7 shows good stereochemistry and no outliers in the Ramachandran plot. Figures were created with PyMOL (Schrödinger, 2010). Coordinates and structure factors have been deposited with the Protein Data Bank (accession number 4JC5).

Supplementary Data

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

Acknowledgements

We thank Dr U Kalinke for providing us MAVS-deficient macrophages; Dr S Akira for providing us TBK1/IKKε-deficient MEFs, and MyD88- and TRIF-deficient macrophages; Dr S Weiss for providing us pIFNβ-FFLuc mice; Dr R Vance for providing us goldenticket (Gt) mice; and Dr G Zimmer for providing us VSV*ΔG(Luc) replicon particles. We thank EMBL for supplying the pET28-SUMO1-eGFP vector and the SenP2 vector. We thank Stefan Emming for help with cloning and protein purification, and Dr Gregor Witte for help with model refinement. We thank the Swiss Light Source and the European Synchrotron Radiation Facility for generous beam time and excellent on-site support. This work was supported by grants from the German Research Foundation (SFB670), the European Research Council (ERC–2009–StG 243046) to VH, NIH grant U19AI083025 to K-PH and a BIF fellowship to TC. VH is a member of the excellence cluster ImmunoSensation.

Author contributions: TC performed all experiments, except the DSF studies and the crystallographic studies, which were performed by TD. AA generated constructs for the study. K-PH and VH conceived or designed the experiments. TC, TD, AA, K-PH and VH analysed the data, TC and VH wrote the manuscript, and VH conceived and supervised the study.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213–221

Ahn J, Gutman D, Saijo S, Barber GN (2012) STING manifests self DNA-dependent inflammatory disease. *Proc Natl Acad Sci USA* **109:** 19386–19391

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22:** 195–201

Aslanidis C, de Jong PJ (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18:** 6069–6074

Berger Rentsch M, Zimmer G (2011) A vesicular stomatitis virus replicon-based bioassay for the rapid and sensitive determination of multi-species type I interferon. *PLoS One* **6:** e25858

Bibby MC, Double JA (1993) Flavone acetic acid–from laboratory to clinic and back. *Anticancer Drugs* **4:** 3–17

Brehm G, Storch E, Kirchner H (1986) Characterization of interferon induced in murine macrophage cultures by 10-carboxymethyl-9-acridanone. *Nat Immun Cell Growth Regul* **5:** 50–59

Brunette RL, Young JM, Whitley DG, Brodsky IE, Malik HS. Stetson DB (2012) Extensive evolutionary and functional diversity among mammalian AIM2-like receptors. *J Exp Med* **209:** 1969–1983

Burdette DL, Monroe KM, Sotelo-Troha K, Iwig JS, Eckert B, Hyodo M, Hayakawa Y, Vance RE (2011) STING is a direct innate immune sensor of cyclic di-GMP. *Nature* **478:** 515–518

Ching LM, Baguley BC (1987) Induction of natural killer cell activity by the antitumour compound flavone acetic acid (NSC 347 512). *Eur J Cancer Clin Oncol* **23:** 1047–1050

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126–2132

Futami H, Eader LA, Komschlies KL, Bull R, Gruys ME, Ortaldo JR, Young HA, Wiltrout RH (1991) Flavone acetic acid directly induces expression of cytokine genes in mouse splenic leukocytes but not in human peripheral blood leukocytes. *Cancer Res* **51:** 6596–6602

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

Gobbi S, Belluti F, Bisi A, Piazzi L, Rampa A, Zampiron A, Barbera M, Caputo A, Carrara M (2006) New derivatives of xanthenone-4-acetic acid: synthesis, pharmacological profile and effect on TNF-alpha and NO production by human immune cells. *Bioorg Med Chem* **14:** 4101–4109

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* **38:** W695–W699

Hornung RL, Young HA, Urba WJ, Wiltrout RH (1988) Immunomodulation of natural killer cell activity by flavone acetic acid: occurrence via induction of interferon alpha/beta. *J Natl Cancer Inst* **80:** 1226–1231

Huang YH, Liu XY, Du XX, Jiang ZF, Su XD (2012) The structural basis for the sensing and binding of cyclic di-GMP by STING. *Nat Struct Mol Biol* **19:** 728–730

Isaacs A, Cox RA, Rotem Z (1963) Foreign nucleic acids as the stimulus to make interferon. *Lancet* **2:** 113–116

Isaacs A, Lindenmann J (1957) Virus interference. I. The interferon. *Proc R Soc Lond B Biol Sci* **147:** 258–267

Ishikawa H, Barber GN (2008) STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* **455:** 674–678

Ishikawa H, Ma Z, Barber GN (2009) STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* **461:** 788–792

Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132

Karaolis DK, Means TK, Yang D, Takahashi M, Yoshimura T, Muraille E, Philpott D, Schroeder JT, Hyodo M, Hayakawa Y, Talbot BG, Brouillette E, Malouin F (2007a) Bacterial c-di-GMP is an immunostimulatory molecule. *J Immunol* **178:** 2171–2181

Karaolis DK, Newstead MW, Zeng X, Hyodo M, Hayakawa Y, Bhan U, Liang H, Standiford TJ (2007b) Cyclic di-GMP stimulates protective innate immunity in bacterial pneumonia. *Infec Immun* **75:** 4942–4950

Kaufman HE, Centifanto YM, Ellison ED, Brown DC (1971) Tilorone hydrochloride: human toxicity and interferon stimulation. *Proc Soc Exp Biol Med* **137:** 357–360

Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* **37:** D387–D392

Kramer MJ, Cleeland R, Grunberg E (1976) Antiviral activity of 10-carboxymethyl-9-acridanone. *Antimicrob Agents Chemother* **9:** 233–238

Kramer MJ, Taylor JL, Grossberg SE (1981) Induction of interferon in mice by 10-carboxymethyl-9-acridanone. *Methods Enzymol* **78:** 284–287

Krueger RE, Mayer GD (1970) Tilorone hydrochloride: an orally active antiviral agent. *Science* **169:** 1213–1214

Lara Jr PN, Douillard JY, Nakagawa K, von Pawel J, McKeage MJ, Albert I, Losonczy G, Reck M, Heo DS, Fan X, Fandi A, Scagliotti G (2011) Randomized phase III placebo-controlled trial of carboplatin and paclitaxel with or without the vascular disrupting agent vadimezan (ASA404) in advanced non-small-cell lung cancer. *J Clin Oncol* **29:** 2965–2971

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23:** 2947–2948

Lienenklaus S, Cornitescu M, Zietara N, Lyszkiewicz M, Gekara N, Jablonska J, Edenhofer F, Rajewsky K, Bruder D, Hafner M, Staeheli P, Weiss S (2009) Novel reporter mouse reveals constitutive and inflammatory expression of IFN-beta *in vivo*. *J Immunol* **183:** 3229–3236

Mayer GD, Krueger RF (1970) Tilorone hydrochloride: mode of action. *Science* **169:** 1214–1215

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40:** 658–674

McWhirter SM, Barbalat R, Monroe KM, Fontana MF, Hyodo M, Joncker NT, Ishii KJ, Akira S, Colonna M, Chen ZJ, Fitzgerald KA, Hayakawa Y, Vance RE (2009) A host type I interferon response is induced by cytosolic sensing of the bacterial second messenger cyclic-di-GMP. *J Exp Med* **206:** 1899–1911

Niesen FH, Berglund H, Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2:** 2212–2221

Ouyang S, Song X, Wang Y, Ru H, Shaw N, Jiang Y, Niu F, Zhu Y, Qiu W, Parvatiyar K, Li Y, Zhang R, Cheng G, Liu ZJ (2012) Structural analysis of the STING adaptor protein reveals a hydrophobic dimer interface and mode of cyclic di-GMP binding. *Immunity* **36:** 1073–1086

Pang JH, Cao Z, Joseph WR, Baguley BC, Ching LM (1998) Antitumour activity of the novel immune modulator 5,6-dimethylxanthenone-4-acetic acid (DMXAA) in mice lacking the interferon-gamma receptor. *Eur J Cancer* **34:** 1282–1289

Patel S, Parkin SM, Bibby MC (1997) The effect of 5,6-dimethylxanthenone-4-acetic acid on tumour necrosis factor production by human immune cells. *Anticancer Res* **17:** 141–150

Peitsch MC (1995) Protein modeling by E-mail. *Nat Biotechnol* **13:** 658–660

Perera PY, Barber SA, Ching LM, Vogel SN (1994) Activation of LPS-inducible genes by the antitumor agent 5,6-dimethylxanthenone-4-acetic acid in primary murine macrophages. Dissection of signaling pathways leading to gene induction and tyrosine phosphorylation. *J Immunol* **153:** 4684–4693

Philpott M, Ching LM, Baguley BC (2001) The antitumour agent 5,6-dimethylxanthenone-4-acetic acid acts *in vitro* on human mononuclear cells as a co-stimulator with other inducers of tumour necrosis factor. *Eur J Cancer* **37:** 1930–1937

Prantner D, Perkins DJ, Lai W, Williams MS, Sharma S, Fitzgerald KA, Vogel SN (2012) 5,6-Dimethylxanthenone-4-acetic acid (DMXAA) activates Stimulator of Interferon Gene (STING)-dependent innate immune pathways and is regulated by mitochondrial membrane potential. *J Biol Chem* **287:** 39776–39788

Roberts ZJ, Goutagny N, Perera PY, Kato H, Kumar H, Kawai T, Akira S, Savan R, van Echo D, Fitzgerald KA, Young HA, Ching LM, Vogel SN (2007) The chemotherapeutic agent DMXAA potently and specifically activates the TBK1-IRF-3 signaling axis. *J Exp Med* **204:** 1559–1569

Rotem Z, Cox RA, Isaacs A (1963) Inhibition of virus multiplication by foreign nucleic acid. *Nature* **197:** 564–566

Sauer JD, Sotelo-Troha K, von Moltke J, Monroe KM, Rae CS, Brubaker SW, Hyodo M, Hayakawa Y, Woodward JJ, Portnoy DA, Vance RE (2011) 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. *Infec Immun* **79:** 688–694

Schmid-Burgk JL, Xie Z, Frank S, Virreira Winter S, Mitschka S, Kolanus W, Murray A, Benenson Y (2012) Rapid hierarchical assembly of medium-size DNA cassettes. *Nucleic Acids Res* **40:** e92

Schrödinger LLC (2010) *The PyMOL Molecular Graphics System Version 1.3*

Shang G, Zhu D, Li N, Zhang J, Zhu C, Lu D, Liu C, Yu Q, Zhao Y, Xu S, Gu L (2012) Crystal structures of STING protein reveal basis for recognition of cyclic di-GMP. *Nat Struct Mol Biol* **19:** 725–727

Shu C, Yi G, Watts T, Kao CC, Li P (2012) Structure of STING bound to cyclic di-GMP reveals the mechanism of cyclic dinucleotide recognition by the immune system. *Nat Struct Mol Biol* **19:** 722–724

Silin DS, Lyubomska OV, Ershov FI, Frolov VM, Kutsyna GA (2009) Synthetic and natural immunomodulators acting as interferon inducers. *Curr Pharm Des* **15:** 1238–1247

Storch E, Kirchner H (1982) Induction of interferon in murine bone marrow-derived macrophage cultures by 10-carboxymethyl-9-acridanone. *Eur J Immunol* **12:** 793–796

Storch E, Kirchner H, Brehm G, Huller K, Marcucci F (1986) Production of interferon-beta by murine T-cell lines induced by 10-carboxymethyl-9-acridanone. *Scand J Immunol* **23:** 195–199

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

Sun W, Li Y, Chen L, Chen H, You F, Zhou X, Zhou Y, Zhai Z, Chen D, Jiang Z (2009) ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. *Proc Natl Acad Sci USA* **106:** 8653–8658

Taylor JL, Schoenherr C, Grossberg SE (1980a) Protection against Japanese encephalitis virus in mice and hamsters by treatment with carboxymethylacridanone, a potent interferon inducer. *J Infect Dis* **142:** 394–399

Taylor JL, Schoenherr CK, Grossberg SE (1980b) High-yield interferon induction by 10-carboxymethyl-9-acridanone in mice and hamsters. *Antimicrob Agents Chemother* **18:** 20–26

Vershinina MY, Narovlyansky AN, Deryabin PG, Amchenkova AM, Ivanova AM, Scherbenko VE, Nagurskaya EV, Bechalo VA, Timofeeva TY, Sanin AV, Ershov FI (2002) Regulation of cytokine mRNAs by interferon and interferon inducers. *Russ J Immunol* **7:** 161–166

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67:** 235–242

Woodward JJ, Iavarone AT, Portnoy DA (2010) c-di-AMP secreted by intracellular Listeria monocytogenes activates a host type I interferon response. *Science* **328:** 1703–1705

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

Yin Q, Tian Y, Kabaleeswaran V, Jiang X, Tu D, Eck MJ, Chen ZJ, Wu H (2012) Cyclic di-GMP sensing via the innate immune signaling protein STING. *Mol Cell* **46:** 735–745

Zarubaev VV, Slita AV, Krivitskaya VZ, Sirotkin AK, Kovalenko AL, Chatterjee NK (2003) Direct antiviral effect of cycloferon (10-carboxymethyl-9-acridanone) against adenovirus type 6 *in vitro*. *Antiviral Res* **58:** 131–137

Zhong B, Yang Y, Li S, Wang YY, Li Y, Diao F, Lei C, He X, Zhang L, Tien P, Shu HB (2008) The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. *Immunity* **29:** 538–550

2.3. cGAS produces a 2’-5’-linked cyclic dinucleotide second messenger that activates STING

2.3.1. Summary

It has been known for a long time that cytosolic DNA is an important MAMP that triggers type I IFN production. Just recently cGAS was revealed as the responsible PRR for sensing this class of MAMPs. Binding of cytosolic dsDNA activates cGAS leading to the production of a cyclic dinucleotide second messenger that induces STING-dependent signalling and the production of type I IFNs. However, the exact nature of this cyclic dinucleotide was unclear. In this publication we were able to characterize this second messenger using a variety of biochemical and biophysical methods.

In a previous publication cGAMP was suggested to be the second messenger that is produced by cGAS. Since the data in this publication was not able to unambiguously proof this suggestion, we wanted to test if cGAMP is really produced by cGAS. Therefore, we performed experiments in HEK293T cells expressing mouse wild-type (wt) or R231A STING. The R231A mutant is known to be responsive to the product of cGAS but unresponsive to c-di-GMP. We stimulated these cells endogenously by additional overexpression of either cGAS, tDGC (thermophilic diguanylate cyclase domain -> c-di-GMP) or DcnV (-> cGAMP), respectively. In wt STING cells the stimulation led to the production of type I IFNs in all cases. In R231A STING cells type I IFN induction was only observed when cGAS was overexpressed. These results revealed that cGAMP cannot be the second messenger that is produced by cGAS. Indeed, the protein-depleted lysates of cGAS overexpressing HEK293T cells showed an additional, unique peak in a RP-HPLC (Reversed-Phase High-Performance Liquid Chromatography). It eluted with a different retention time as the cGAMP peak in the control experiment. Mass spectrometry analysis revealed that the molecular mass of the cGAS product matches the mass of cGAMP but the fragmentation patterns were different. A series of additional nuclease digestion assays coupled with TLC (Thin Layer Chromatography) and ESI-LC-MS (Electrospray Ionization Liquid Chromatography Mass Spectroscopy) identified 2,3-

cGAMP as the product of the cGAS reaction. This was further supported by H^1-NMR data.

A low responsiveness of human STING towards bacterial cyclic dinucleotides compared to mouse STING was reported earlier. Therefore, we decided to compare cGAMP and 2,3-cGAMP activation of STING for the human and the mouse protein. We performed assays with HEK293T cells expressing either human or mouse STING. Stimulation with 2,3-cGAMP led to induction of IFN expression in both cases, while cGAMP only induced an IFN response in cells expressing mouse STING. DSF (Differential Scanning Fluorimetry) assays with mouse or human STING-LBD and 2,3-cGAMP or cGAMP supported these results. In summary, these observations indicated that the endogenous second messenger 2,3-cGAMP is indeed a high potent activator of mouse and human STING.

By analyzing the side products of the cGAS reaction and by recycling them for a second round of cGAS mediated cyclization reaction we could postulate a two step mechanism of 2,3-cGAMP formation. In the first step, cGAS utilizes GTP and ATP to form the linear intermediate pppGMP-2'-p-5'-AMP with the unique guanine-2'-p-5'-adenosine phosphodiester linkage. In the second step the cyclization reaction is completed and the final product is formed. In both steps pyrophosphate is released.

In this publication we could identify the eukaryotic second messenger 2,3-cGAMP and shed light on its formation and involvement in the signalling of cytosolic DNA. This expands our understanding of the processes involved in various bacterial and viral infections and can help to develop therapies for several diseases.

2.3.2. Contribution

The author of this thesis contributed to this publication by cloning and purifying human and mouse STING and cGAS constructs for *in vitro* experiments. In addition, he performed the DSF assays.

2.3.3. Paper

The following paper was published 2013 in nature.

cGAS produces a 2’-5’-linked cyclic dinucleotide second messenger that activates STING

Andrea Ablasser, Marion Goldeck, Taner Cavlar, Tobias Deimling, Gregor Witte, Ingo Röhl, Karl-Perter Hopfner, Janos Ludwig & Veit Hornung

Nature 2013 June 20; **498**(7454): 380 - 4

LETTER

doi:10.1038/nature12306

cGAS produces a 2′-5′-linked cyclic dinucleotide second messenger that activates STING

Andrea Ablasser[1], Marion Goldeck[1], Taner Cavlar[1], Tobias Deimling[2], Gregor Witte[2], Ingo Röhl[3], Karl-Peter Hopfner[2,4], Janos Ludwig[1] & Veit Hornung[1]

Detection of cytoplasmic DNA represents one of the most fundamental mechanisms of the innate immune system to sense the presence of microbial pathogens[1]. Moreover, erroneous detection of endogenous DNA by the same sensing mechanisms has an important pathophysiological role in certain sterile inflammatory conditions[2,3]. The endoplasmic-reticulum-resident protein STING is critically required for the initiation of type I interferon signalling upon detection of cytosolic DNA of both exogenous and endogenous origin[4–8]. Next to its pivotal role in DNA sensing, STING also serves as a direct receptor for the detection of cyclic dinucleotides, which function as second messenger molecules in bacteria[9–13]. DNA recognition, however, is triggered in an indirect fashion that depends on a recently characterized cytoplasmic nucleotidyl transferase, termed cGAMP synthase (cGAS), which upon interaction with DNA synthesizes a dinucleotide molecule that in turn binds to and activates STING[14,15]. We here show *in vivo* and *in vitro* that the cGAS-catalysed reaction product is distinct from previously characterized cyclic dinucleotides. Using a combinatorial approach based on mass spectrometry, enzymatic digestion, NMR analysis and chemical synthesis we demonstrate that cGAS produces a cyclic GMP-AMP dinucleotide, which comprises a 2′-5′ and a 3′-5′ phosphodiester linkage >Gp(2′-5′)Ap(3′-5′)>. We found that the presence of this 2′-5′ linkage was required to exert potent activation of human STING. Moreover, we show that cGAS first catalyses the synthesis of a linear 2′-5′-linked dinucleotide, which is then subject to cGAS-dependent cyclization in a second step through a 3′-5′ phosphodiester linkage. This 13-membered ring structure defines a novel class of second messenger molecules, extending the family of 2′-5′-linked antiviral biomolecules.

Recently, it has been demonstrated that upon intracellular DNA delivery, a cytoplasmic enzyme dubbed cyclic GMP-AMP synthase (cGAS) produces a ribo-dinucleotide, which in turn binds to and activates STING[14,15]. Given the striking analogy to bacterial cyclic dinucleotide recognition and its determined molecular mass, it was suggested that this molecule constitutes a cyclic adenosine monophosphate-guanosine monophosphate (cGAMP) with a symmetric 12-membered ring formed by 3′-5′ linked nucleotide residues (>Gp(3′-5′)Ap(3′-5′)>, cGAMP(3′-5′)). On the other hand, it was shown that STING-dependent DNA sensing can be differentiated from bacterial cyclic di-GMP recognition through a point mutation at a conserved arginine residue (R231A) within the lid region of murine STING[9]. R231 functions to indirectly bind the phosphate of the phosphodiester bond of cyclic di-GMP/AMP through a Mg^{2+} or H_2O molecule, yet this coordination seems to be dispensable for STING activation in response to DNA transfection. We have recently identified a novel STING ligand (10-carboxymethyl-9-acridanone, CMA) that also triggers STING activation independently of the R231 residue[16]. In fact, the crystal structure of CMA bound to murine STING revealed that the lid region binds CMA differently than cyclic di-GMP and that R231 is not involved in CMA binding. We were intrigued by the differential role of R231 for DNA and cyclic di-GMP sensing, given the fact that modelling studies using cGAMP(3′-5′) instead of cyclic di-GMP could not readily explain the reported differential role of this residue at the structural level. To explore this further, we expressed cGAS in HEK293T cells together with either wild-type murine STING or its R231A mutant. As a control, we induced endogenous cyclic di-GMP production using a codon-optimized version of the thermophilic diguanylate cyclase domain (tDGC) (amino acids 83–248) of *Thermotoga maritima*[17] and a codon-optimized version of the recently discovered bacterial cGAMP(3′-5′) synthetase (DncV) from *Vibrio cholerae*[18]. As expected, overexpression of the cyclic di-GMP synthetase, the cGAMP synthetase and cGAS induced

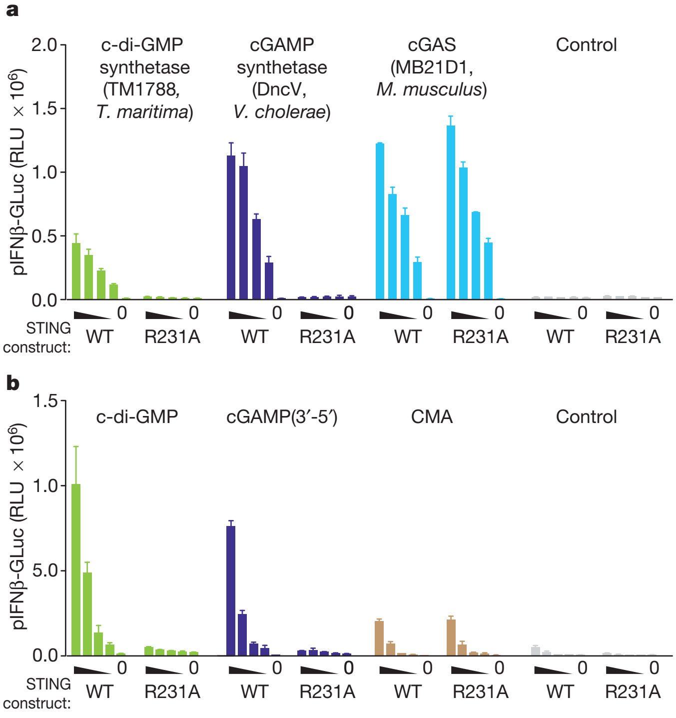


Figure 1 | The R231A STING mutant uncouples cyclic di-GMP sensing from cGAS-induced activation. **a**, Overexpression of dinucleotide synthetases. HEK293T cells were transfected with different dinucleotide synthetases (100 ng) together with decreasing amounts of wild-type (WT) mmSTING or the R231A mutant (10, 5, 2.5, 1.25 and 0 ng) and a pIFNβ-luciferase reporter (pIFNβ-GLuc). Reporter activity was measured 16 h after transfection. RLU, relative light units. **b**, Direct stimulation with synthetic compounds. HEK293T cells were transfected with WT mmSTING or the R231A mutant in conjunction with pIFNβ-GLuc. The next day CMA was added or synthetic cyclic di-GMP or synthetic cGAMP(3′-5′) was transfected as indicated and pIFNβ-GLuc activity was assayed 16 h later. Representative data of two (**a**) or three (**b**) independent experiments are shown (mean values + s.e.m.).

---

[1] Institute for Clinical Chemistry and Clinical Pharmacology, University Hospital, University of Bonn, 53127 Bonn, Germany. [2] Department of Biochemistry and Gene Center, Ludwig-Maximilians-University, 81377 Munich, Germany. [3] Axolabs GmbH, 95326 Kulmbach, Germany. [4] Center for Integrated Protein Sciences, 81377 Munich, Germany.

a robust type I interferon (IFN) response in HEK293T cells expressing wild-type murine STING. Moreover, in line with previous reports, expression of the R231A point mutant completely abolished type I IFN production in response to endogenous cyclic di-GMP production, but not upon overexpression of cGAS (Fig. 1a and Supplementary Fig. 1). Surprisingly, however, induction of endogenous cGAMP production using DncV was also completely blunted for the R231A mutant. Next we stimulated HEK293T cells overexpressing wild-type murine STING or the R231A mutant directly with synthetic compounds. As previously reported, CMA-mediated activation of STING did not require coordination through R231 and in accordance with the synthetase data from above, synthetic cyclic di-GMP only activated cells expressing wild-type murine STING, but not the R231A mutant (Fig. 1b). Unexpectedly, synthetic cGAMP(3′-5′) was also completely blunted in its stimulatory activity when transfected into cells expressing STING(R231A). Altogether, these results confirmed previous reports on DNA/cGAS-mediated STING activation being distinct from cyclic dinucleotide sensing with regards to the involvement of the lid region of STING. At the same time, however, these results questioned the concept of cGAMP(3′-5′) being the cGAS-dependent second messenger molecule activating STING.

To follow up on this observation, we generated cytoplasmic lysates from cGAS overexpressing HEK293T cells and untreated HEK293T cells and subjected the protein-depleted, low-molecular-weight fraction to reversed-phase high-performance liquid chromatography (RP-HPLC). In comparison to untreated HEK293T cells, cGAS-overexpressing HEK293T cells showed an additional, unique peak with a retention time of 46 min (Fig. 2a, *), whereas synthetic cGAMP(3′-5′) spiked into cell lysate eluted at a far higher retention time (Fig. 2a, **). Comparing endogenously produced cyclic di-GMP to synthetic cyclic di-GMP under the same conditions revealed no difference in retention time, excluding the possibility of the purification process affecting the physiochemical properties of the compounds (Supplementary Fig. 2). Fractionation of the cell-derived, cGAS-specific low-molecular-weight product (*) and transfer into STING competent LL171 cells revealed potent stimulatory activity, within the same range as synthetic cGAMP(3′-5′) (Fig. 2b and Supplementary Fig. 3). Analogous results were obtained when purified cGAS was incubated *in vitro* with GTP and ATP (Fig. 2c). A cGAS-dependent peak could be detected at the same retention time as in cell lysates from cGAS overexpressing HEK293T cells and only this peak exerted stimulatory activity in LL171 cells (Fig. 2d). Thin-layer chromatography (TLC) as an alternative separation technique revealed that the cell-derived and the *in-vitro*-synthesized cGAS product showed a similar chromatographic mobility to that of the synthetic cGAMP(3′-5′) (Fig. 2e). Despite the big difference in chromatographic properties under RP-HPLC conditions, electrospray ionization-liquid chromatography-mass spectrometry (ESI-LC-MS) analysis revealed the same molecular mass (m/z (M-H) = 673.1) for both the cell-derived cGAS product and synthetic cGAMP(3′-5′) (Fig. 2f). In addition, while the MS/MS fragmentation pattern of the cGAS-derived molecule was consistent with a ribo-dinucleotide made up of guanosine and adenosine, these studies reproducibly displayed a clear difference compared to synthetic cGAMP(3′-5′) (Supplementary Fig. 4). Most intriguingly, the MS/MS fragmentation studies pointed to the presence of a 2′-5′ phosphodiester bond between guanosine and adenosine (Supplementary Figs 4–6 and Supplementary Notes 1 and 2).

On the basis of these observations, we considered several candidate molecules as products of cGAS (Supplementary Fig. 7). Among these, a cyclic dinucleotide with one or two 2′-5′ phosphodiester bonds seemed to be most likely. To address this hypothesis, we performed a series of enzyme digests coupled to TLC and ESI-LC-MS. First, we treated synthetic cGAMP(3′-5′) and the *in-vivo*- and *in-vitro*-synthesized

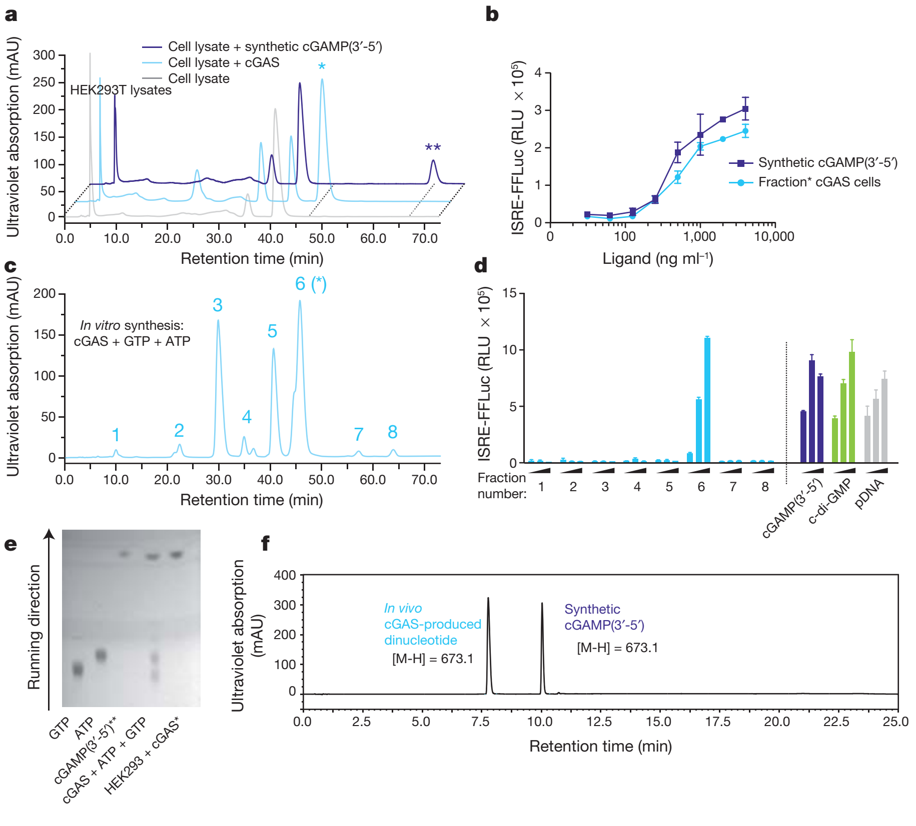


Figure 2 | The cGAS reaction product is distinct from cGAMP(3′-5′). **a**, RP-HPLC chromatograms of lysates of untreated HEK293T cells (grey), of cGAS overexpressing HEK293T cells (light blue) or of synthetic cGAMP(3′-5′) spiked into untreated HEK293T lysate (dark blue). Asterisks highlight differential elution peaks. **b**, IFN-stimulated response element (ISRE) activity in LL171 cells. Endogenous cGAS product was purified from **a** and transfected into LL171 cells, whereas synthetic cGAMP(3′-5′) served as a control. ISRE-reporter activity was measured 14 h later. **c**, Chromatogram of an *in vitro* cGAS assay. The asterisk indicates the fraction that elutes at the same retention time as the endogenous product from **a**. **d**, ISRE activity in LL171 cells. Peaks 1–8 from **c** were fractionated and transfected into LL171 cells that were then studied for ISRE-reporter activity using respective control stimuli. **e**, TLC analysis of *in-vitro*- and *in-vivo*-synthesized cGAS product with ATP, GTP and synthetic cGAMP(3′-5′) as controls. **f**, ESI-LC-MS analysis of *in-vivo*-produced cGAS product and synthetic cGAMP(3′-5′). Representative data of two (**e**) or three (**a**–**d**, **f**) independent experiments are shown (mean values + s.e.m.).

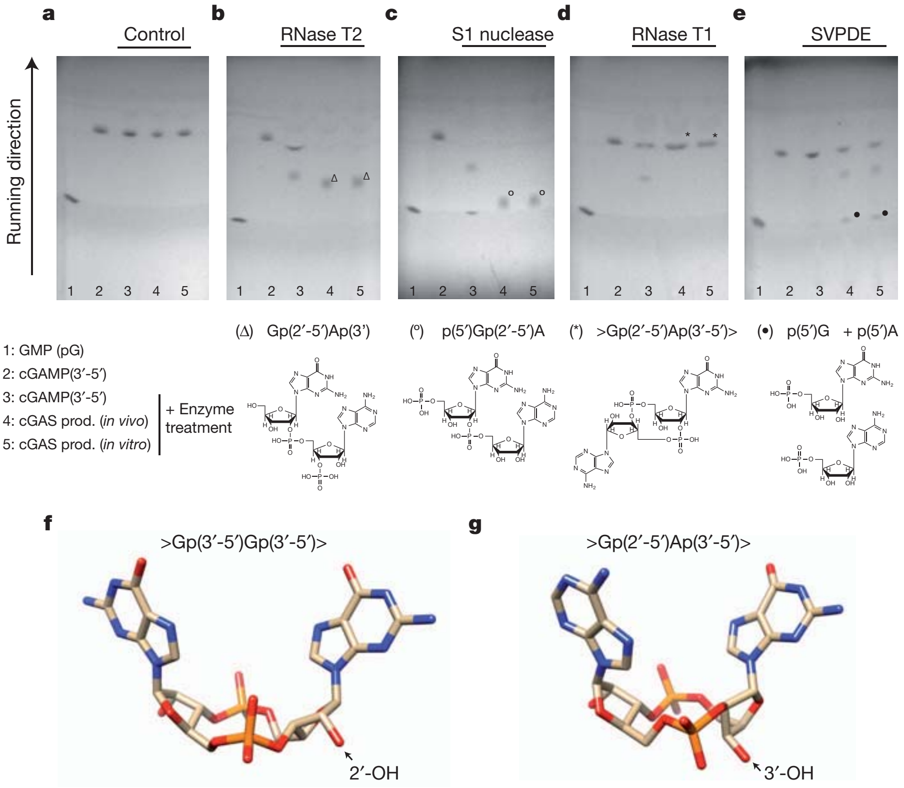


Figure 3 | The second messenger produced by cGAS is >Gp(2′-5′)Ap(3′-5′)>. **a**–**e**, TLC analysis of GMP (lane 1), synthetic cGAMP (3′-5′) (lane 2) and enzyme-treated synthetic cGAMP(3′-5′) (lane 3), *in-vivo*-synthesized cGAS product (lane 4) and *in-vitro*-synthesized cGAS product (lane 5). Enzyme treatments of molecules analysed in lanes 3–5 were control (**a**), RNase T2 (**b**), S1 nuclease (**c**), RNase T1 (**d**) and SVPDE (**e**). The resulting reaction products from lanes 4 and 5 as confirmed by ESI-LC-MS analysis are depicted below. Representative data out of two independent experiments are shown. **f**, **g**, Comparison of the structure of cyclic di-GMP (4F9G.pdb) and a model for cGAMP(2′-5′) based on NMR-derived ribose conformations.

cGAS product using S1 nuclease and ribonuclease T2. Both enzymes can cleave internal 3′-5′ phosphodiester linkages. Synthetic cGAMP(3′-5′) could be processed into mononucleotides by both enzymes, whereas the cGAS-derived cyclic dinucleotide was only cleaved into a linear dinucleotide (Fig. 3a–c and Supplementary Fig. 9a, b). These results suggested that one of the internucleotide bonds was not a 3′-5′ phosphodiester. To address which one of the two phosphodiester bonds was not hydrolysable by the enzymes above, we took advantage of the nucleotide specificity of ribonuclease T1, which catalyses the endonucleolytic cleavage of 3′-5′ phosphodiester bonds only after guanosine. As expected, ribonuclease T1 processed synthetic cGAMP(3′-5′) into a linear dinucleotide, consistent with the presence of a Gp(3′-5′)A phosphodiester bond (Fig. 3d and Supplementary Fig. 9c). The cGAS-derived dinucleotide, however, was not processed, indicating that the

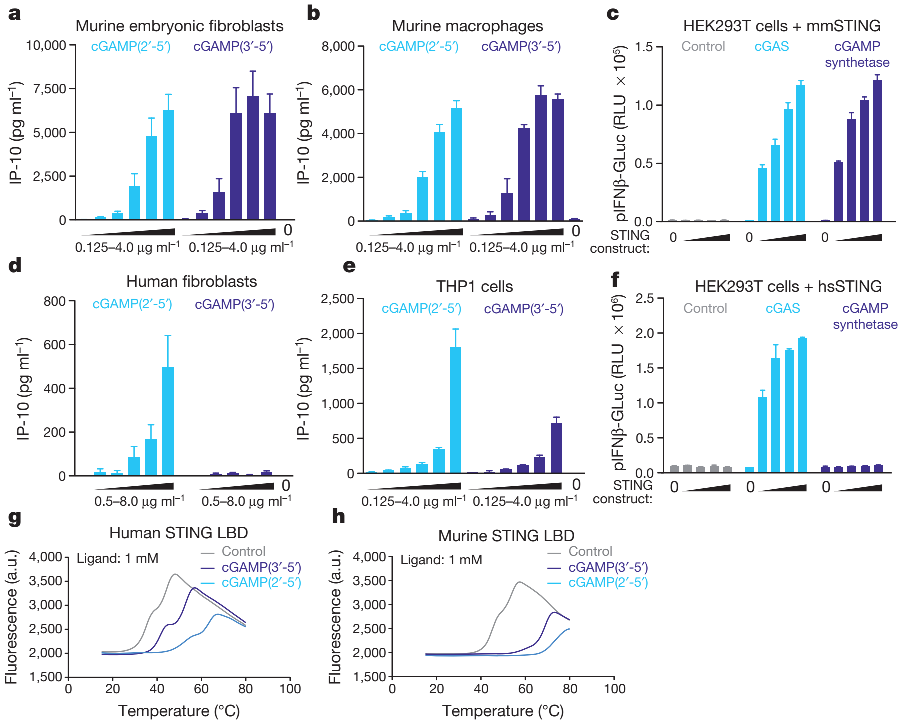


Figure 4 | cGAMP(2′-5′) is a potent activator of human and murine STING. **a**, **b**, **d**, **e**, IP-10 production of murine embryonic fibroblasts, murine macrophages, human fibroblasts and THP1 cells transfected with increasing amounts of cGAMP(2′-5′) or cGAMP(3′-5′). **c**, **f**, HEK293T cells transfected with human or murine STING (0, 3.13, 6.25, 12.5 and 25 ng) as indicated, together with cGAS or DncV cGAMP synthetase (100 ng) subsequently analysed for pIFNβ-GLuc activity. **g**, **h**, Interaction of the human and murine STING LBD with 1 mM cGAMP(2′-5′) or cGAMP(3′-5′) analysed by DSF. Mean + s.e.m. of two (**e**) or three (**a**, **b**, **d**) independent experiments or one representative experiment out of three independent experiments (**c**, **f**, **g**, **h**) are depicted (**c**, **f**, mean values + s.e.m.).

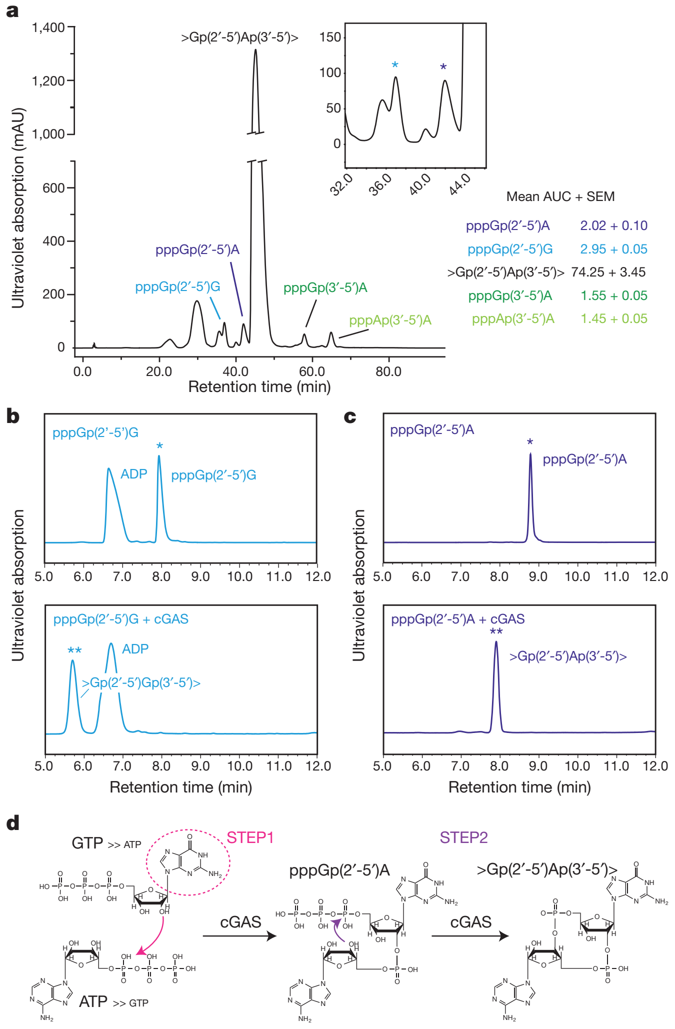


Figure 5 | >Gp(2′-5′)Ap(3′-5′)> is synthesized in a two-step process. **a**, RP-HPLC chromatogram of a cGAS+ATP+GTP *in vitro* reaction upon termination of the reaction. The insertion represents an enlargement of the chromatogram indicating the position of pppGp(2′-5′)G and pppGp(2′-5′)A. Right panel demonstrates mean area under the curve (AUC) + s.e.m. for depicted dinucleotides out of two independent experiments. **b**, **c**, ESI-LC-MS chromatograms before (top) and after (bottom) *in vitro* incubation of pppGp(2′-5′)G (**b**) and pppGp(2′-5′)A (**c**) with cGAS. Asterisks indicate the position of the substrates (*) and the resulting products (**). Note that the fraction of pppGp(2′-5′)G was contaminated with ADP. Data are representative of two independent experiments. **d**, Schematic model of the two-step process of cGAS-catalysed cyclic dinucleotide synthesis.

GpA phosphodiester bond was not 3′-5′. In line with this notion, substitution of GTP by 2′dGTP during the *in vitro* enzymatic reaction completely blunted synthesis of dinucleotides by cGAS, whereas addition of 3′dGTP gave rise to small, but consistent amounts of a cGAMP product (Supplementary Fig. 10). In a reverse approach, we made use of snake venom phosphodiesterase I (SVPDE), which can hydrolyse 5′-mononucleotides from 3′-hydroxy-terminated ribo-oligonucleotides. Consistent with its two internal 3′-5′ phosphodiester bonds, synthetic cGAMP was not processed by SVPDE, whereas the cGAS-derived product was hydrolysed in a two-step process into its mononucleotide components (Fig. 3e and Supplementary Fig. 9d). Altogether, these results clearly identified the cGAS-derived dinucleotide product as a cyclic GA dinucleotide with a 2′-5′ phosphodiester linkage between guanosine and adenosine and a 3′-5′ phosphodiester linkage between adenosine and guanosine >Gp(2′-5′)Ap(3′-5′)> (cGAMP(2′-5′)). Comparison of synthetic cGAMP(3′-5′) and the cGAS-derived product by ^{1}H-NMR spectroscopy supported the notion of differing 3′-5′/2′-5′ linkages between guanosine and adenosine within the cyclic GA dinucleotide and revealed crucial information on the ribose conformation of the distinct nucleotide elements in both molecules (Supplementary Fig. 11 and Supplementary Note 3). Relying on these data we were able to formulate a model of cGAMP(2′-5′) based on the previously determined structure of cyclic di-GMP bound to STING (Fig. 3f, g and Supplementary Note 3). To provide an additional proof of cGAMP(2′-5′) being the actual cGAS-derived second messenger molecule, we chemically synthesized >Gp(2′-5′)Ap(3′-5′)> and its isomer >Gp(3′-5′)Ap(3′-5′)> (Supplementary Fig. 12). As expected, synthetic cGAMP (2′-5′) had the same physiochemical properties as the *in vitro* cGAS-generated dinucleotide (Supplementary Fig. 13).

Transfection of both DNA and cyclic dinucleotides (3′-5′linked) has been shown to induce indistinguishable transcriptional responses in murine cells in a STING-dependent manner. However, despite being equally responsive towards DNA challenge, recent reports indicated that human cells are less responsive to intracellular delivery of cyclic dinucleotides or other STING ligands[19]. To determine whether these species-specific properties of STING would also apply for the recognition of cGAMP, we first sought to compare both cGAMP isomers with regard to their biological activity in human and murine cells. Transfection of both cGAMP(2′-5′) and cGAMP(3′-5′) into murine embryonic fibroblasts and macrophages strongly induced production of the antiviral cytokine IP-10, confirming previous reports on the stimulatory potency of cyclic dinucleotides for murine STING (Fig. 4a, b). However, to our surprise, we reproducibly observed a marked difference in responsiveness towards the two cGAMP isomers, when we tested them for functional activity in human fibroblasts and the monocytic cell line THP-1 (Fig. 4d, e). Here, cGAMP(2′-5′) was more active than cGAMP(3′-5′) with regards to production of IP-10. In fact, human fibroblasts were almost unresponsive towards transfection with cGAMP(3′-5′), even at high concentrations of the cyclic dinucleotide being delivered. Similar results were obtained when we studied HEK293T cells overexpressing cGAS or the cGAMP synthetase DncV derived from *V. cholerae* (Fig. 4c and f). Whereas expression of murine STING rendered HEK293T responsive towards both the cGAS and the cGAMP synthetase products, only cGAS expression was able to activate human STING. Finally, we performed binding studies with the carboxy-terminal ligand-binding domain (LBD) of mouse and human STING using differential scanning fluorometry (DSF). These data revealed that cGAMP(2′-5′) showed stronger complex formation with both human and mouse STING compared to cGAMP(3′-5′) (Fig. 4g, h). Of note, this preference for cGAMP(2′-5′) over cGAMP(3′-5′) was more prominent when the LBD of human STING was tested. Together, these data indicate that, in contrast to cGAMP(3′-5′), cGAMP(2′-5′) is highly potent in the human system and, like DNA, is not affected by species-specific properties of STING.

Next we wanted to delineate the mechanism of cGAS-dependent cyclic dinucleotide synthesis. Enzymatic cGAS reactions in the presence of excess substrate (ATP/GTP) displayed, among other dinucleotide species, one predominant peak in ESI-LC-MS analysis, which was represented by pppGp(2′-5′)A (Supplementary Fig. 10). However, upon termination of the reaction we identified four distinct linear dinucleotide species at almost similar quantities, next to >Gp(2′-5′)Ap(3′-5)> as the major species (Fig. 5a). These molecules included: pppGp(2′-5′)A, pppGp(3′-5′)A, pppGp(2′-5′)G and pppAp(3′-5′)A (Fig. 5a). When we incubated ATP or GTP alone with cGAS, either pppAp(3′-5′)A or pppGp(2′-5′)G predominated by approximately 10:1 over their respective phosphodiester linkage isomers (Supplementary Fig. 14). Of note, whereas GTP by itself gave rise to substantial amounts of a cyclic dinucleotide (>Gp(2′-5′)Gp(3′-5′)>), ATP by itself was unable to trigger synthesis of detectable amounts of a cyclic dinucleotide. This observation indicated that a 2′-5′ dinucleotide constitutes the substrate for the subsequent cyclization reaction and that pppGp(2′-5′)A represented the precursor molecule for cGAMP(2′-5′). To prove this hypothesis, we fractionated the four major dinucleotide species obtained during enzymatic cGAS reactions and incubated them again with cGAS. Interestingly, we found that both 2′-5′-linked dinucleotide species were quantitatively converted into cyclic dinucleotides, whereas the 3′-5′-linked dinucleotides were only scarcely, if at all, converted (Fig. 5b and Supplementary Fig. 15). Interestingly, the second phosphodiester bond was linked exclusively via 3′-5′ for all cyclization reactions. Together these results unequivocally identified pppGp(2′-5′)A as the precursor of cGAS-dependent cGAMP(2′-5′) synthesis.

On the basis of these results we postulate the following two-step synthesis model (Fig. 5d). (1) In the presence of ATP and GTP cGAS first catalyses the generation of a linear dinucleotide, with the attacking nucleotide determining the type of phosphodiester bond being generated. 5′-GTP preferentially results in a 2′-5′ linkage, whereas 5′-ATP results in a 3′-5′ linkage leading to either pppGp(2′-5′)R or pppAp(3′-5′)R, respectively. In this first synthesis step, cGAS shows a preference for GTP over ATP being the attacking nucleotide, and ATP over GTP for the nucleotide being attacked. (2) Whereas pppGp(2′-5′)R species are quantitatively cyclized by cGAS in a second step, pppRp(3′-5′)A dinucleotides are poor, if at all, substrates for cyclization. Of note, this second step exclusively generates a 3′-5′ linkage, at least for the dinucleotide species studied. The fact that only scarce amounts of cyclic di-GMP are found during *in vitro* reactions might be attributed to lower supply of its precursor molecule pppGp(2′-5′)G and presumably the preference of pppGp(2′-5′)A over pppGp(2′-5′)G during the cyclization step. All in all, this model explains the nearly exclusive generation of >Gp(2′-5′)Ap(3′-5)> by cGAS in the presence of ATP and GTP.

Previously it has been shown that the response of STING towards DNA and cyclic dinucleotides can be uncoupled[9]. This observation can now be rationalized by our finding that cGAS produces a novel class of second messenger being a 2′-5′/3′-5′-linked cyclic dinucleotide, which is structurally and physiochemically distinct from bacteria-derived cyclic dinucleotides. In fact, this report describes the enzymatic production of a cyclic 2′-5′/3′-5′-linked dinucleotide and thereby adds, at the functional level, cGAS to the oligoadenylate synthetase (OAS) family of enzymes that are unique in their ability to synthesize 2′-5′ phosphodiester bonds. Indeed, this functional similarity seems quite plausible given the sequence homology of cGAS to OAS1, which produces 2′-5′-linked oligoadenylates upon binding double stranded RNA[15,20]. Another striking analogy is that cGAS as well as the OAS enzymes both require nucleic acid binding to be activated to synthesize their products in a template-independent fashion. Thus, our results now unequivocally unify these two innate sensing systems and suggest both processes to be evolutionary linked.

The unorthodox chemical linkage within cGAMP(2′-5′) provides a unique feature that may be targeted by specific cellular regulation mechanisms. At the same time, the cGAS-dependent, two-step synthesis of cGAMP(2′-5′) could be amenable for the development of specific inhibitors for the treatment of autoimmune diseases that engage the cGAS–STING axis.

Note added in proof: After submission of the revised version of this manuscript, Gao *et al.*[21] and Diner *et al.*[22] reported the same finding, that cGAMP(2′-5′) is the cGAS-derived second messenger molecule that activates STING.

METHODS SUMMARY

Cell stimulation. If not otherwise indicated, cells were transfected using Lipofectamine 2000 (Invitrogen) with cyclic dinucleotides at a final concentration of $2\,\mu\mathrm{g\,ml^{-1}}$ or DNA (pCI vector) at $1.33\,\mu\mathrm{g\,ml^{-1}}$.

***In vitro* assay for cGAS activity.** For *in vitro* synthesis of the cGAS reaction product $2\,\mu\mathrm{M}$ recombinant cGAS was mixed with $3\,\mu\mathrm{M}$ dsDNA in Buffer A (100 mM NaCl, 40 mM Tris pH 7.5, 10 mM $MgCl_2$) with 1 mM ATP and 1 mM GTP.

Reverse phase-HPLC. Cell lysates and enzymatic reaction mixtures were applied to a 4.1×250 mm PRP-1 column (Hamilton) and separated in a linear gradient of 0% buffer B for 8 min, followed by an increase of buffer B from 0 to 75% in 62 min at a flow rate of $1\,\mathrm{ml\,min^{-1}}$. Buffer A was 20 mM triethylammonium hydrogen carbonate (TEAB) and buffer B 20 mM TEAB in 20% methanol.

Full Methods and any associated references are available in the online version of the paper.

Received 1 April; accepted 20 May 2013.

1. Hornung, V. & Latz, E. Intracellular DNA recognition. *Nature Rev. Immunol.* **10,** 123–130 (2010).
2. Gall, A. *et al.* Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease. *Immunity* **36,** 120–131 (2012).
3. Ahn, J., Gutman, D., Saijo, S. & Barber, G. N. STING manifests self DNA-dependent inflammatory disease. *Proc. Natl Acad. Sci. USA* **109,** 19386–19391 (2012).
4. Ishikawa, H. & Barber, G. N. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* **455,** 674–678 (2008).
5. Zhong, B. *et al.* The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. *Immunity* **29,** 538–550 (2008).
6. Jin, L. *et al.* MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex class II and mediates transduction of apoptotic signals. *Mol. Cell. Biol.* **28,** 5014–5026 (2008).
7. Sun, W. *et al.* ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. *Proc. Natl Acad. Sci. USA* **106,** 8653–8658 (2009).
8. Ishikawa, H., Ma, Z. & Barber, G. N. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* **461,** 788–792 (2009).
9. Burdette, D. L. *et al.* STING is a direct innate immune sensor of cyclic di-GMP. *Nature* **478,** 515–518 (2011).
10. Huang, Y. H., Liu, X. Y., Du, X. X., Jiang, Z. F. & Su, X. D. The structural basis for the sensing and binding of cyclic di-GMP by STING. *Nature Struct. Mol. Biol.* **19,** 728–730 (2012).
11. Ouyang, S. *et al.* Structural analysis of the STING adaptor protein reveals a hydrophobic dimer interface and mode of cyclic di-GMP binding. *Immunity* **36,** 1073–1086 (2012).
12. Shang, G. *et al.* Crystal structures of STING protein reveal basis for recognition of cyclic di-GMP. *Nature Struct. Mol. Biol.* **19,** 725–727 (2012).
13. Shu, C., Yi, G., Watts, T., Kao, C. C. & Li, P. Structure of STING bound to cyclic di-GMP reveals the mechanism of cyclic dinucleotide recognition by the immune system. *Nature Struct. Mol. Biol.* **19,** 722–724 (2012).
14. Wu, J. *et al.* Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science* **339,** 826–830 (2013).
15. 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).
16. Cavlar, T., Deimling, T., Ablasser, A., Hopfner, K. P. & Hornung, V. Species-specific detection of the antiviral small-molecule compound CMA by STING. *EMBO J.* **32,** 1440–1450 (2013).
17. Rao, F. *et al.* Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase. *Anal. Biochem.* **389,** 138–142 (2009).
18. Davies, B. W., Bogard, R. W., Young, T. S. & Mekalanos, J. J. Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. *Cell* **149,** 358–370 (2012).
19. 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. *J. Immunol.* **190,** 5216–5225 (2013).
20. Kristiansen, H., Gad, H. H., Eskildsen-Larsen, S., Despres, P. & Hartmann, R. The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. *J. Interferon Cytokine Res.* **31,** 41–47 (2011).
21. 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).
22. Diner, E. J. The innate immune DNA sensory cGAS produces a noncanonical cyclic dinucleotide that activates human STING. *Cell Rep.* **3,** 1355–1361 (2013).

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

Acknowledgements We thank M. Pelegrin for providing us with LL171 cells. K.-P.H. is supported by the National Institutes of Health (U19AI083025), the European Research Council Advanced Grant 322869, and the Center for Integrated Protein Science Munich (CIPSM). A.A. and V.H. are members of the excellence cluster ImmunoSensation. V.H. is supported by grants from the German Research Foundation (SFB670) and the European Research Council (ERC 243046).

Author Contributions A.A., M.G., T.C., G.W., T.D., I.R., J.L., K.-P.H. and V.H. designed experiments and analysed the data. A.A., M.G., T.C., G.W., T.D. and I.R. performed experiments. A.A. and V.H. wrote the manuscript. V.H. supervised the project.

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. Correspondence and requests for materials should be addressed to A.A. (andrea.ablasser@uni-bonn.de) and V.H. (veit.hornung@uni-bonn.de).

METHODS

Reagents. Cyclic di-GMP and cyclic GAMP(3′-5′) were obtained from Biolog. DNA oligonucleotides corresponding to ISD were obtained from Metabion and annealed in PBS. 10-carboxymethyl-9-acridanone was purchased from Sigma Aldrich. ATP and GTP were obtained from Fermentas.

Cell culture. HEK293T cells, THP1 cells, human fibroblasts (hTERT-BJ1 cells), mouse embryonic fibroblasts, bone marrow-derived macrophages and LL171 cells (L929 cells containing a stable IFN-stimulated response element-luciferase reporter plasmid (ISRE-Luc)) were cultured in DMEM supplemented with 10% (v/v) FCS, sodium pyruvate (all Life Technologies) and Ciprofloxacin (Bayer Schering Pharma). All mouse cells used in this study and human hTERT-BJ1 cells and THP1 cells show responsiveness towards DNA stimulation and thus could be used for the exploration of DNA sensing pathways.

Plasmids. Expression plasmids coding for murine STING (amino-terminal green fluorescent protein (GFP)-tag)[16], murine STING R231A[16] and murine cGAS are based on pEFBOS[23]. Murine cGAS was amplified from cDNA by PCR (forward 5′-ATTACTCGAGATGGAAGATCCGCGTAGA-3′ and reverse 5′-ATTAAGATCTCTATCAAAGCTTGTCAAAAATTGGAAACCC-3′) and cloned into pEFBOS using XhoI and BglII/BamHI. A codon-optimized version of the diguanylate cyclase domain (amino acids 83–248) of TM1788 (*Thermotoga maritima* MSB8) harbouring a point mutation (R158A) to enhance c-diGMP production was cloned into pEFBOS-C-term-Flag/His using XhoI and BamHI[17]. In addition, a codon-optimized version of the *Vibrio cholerae* cGAMP synthetase (DncV; amino acids 1–438) was cloned into pEFBOS-C-term-Flag/His using XhoI and BamHI[18].

Immunoblotting. Cells were lysed in 1× Laemmli buffer and denatured at 95 °C for 5 min. Cell lysates were separated by 10% SDS–PAGE and transferred onto nitrocellulose membranes. Blots were incubated with anti-β-actin-IgG–horseradish peroxidase (HRP) and anti-GFP-IgG/anti-rabbit-IgG–HRP (all Santa Cruz Biotechnology).

Cell stimulation. LL171 cells (0.15×10^6 per ml), murine BMDM (1×10^6 per ml), MEFs (0.15×10^6 per ml), hTERT-BJ1 cells (0.2×10^6 per ml) and THP1 cells (0.6×10^6 per ml) were transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Plasmid DNA (empty pCI vector) was transfected at a final concentration of 1.33 μg ml^{-1}. Unless otherwise indicated, cyclic dinucleotides were transfected at a final concentration of 2 μg ml^{-1}. Cells were stimulated 14 h before final read-out was performed.

Luciferase assay. LL171 cells were lysed in 5× passive lysis buffer (Promega) for 10 min at room temperature. The total cell lysate was incubated with firefly luciferase substrate at a 1:1 ratio and luminescence was measured on an EnVision 2104 Multilabel Reader (Perkin Elmer). pIFNβ-GLuc activity was measured in HEK293T cell supernatants using Coelenterazine as a substrate.

ELISA. Cell culture supernatants were assayed for mouse IP-10 (R&D Systems) and human IP-10 (BD Biosciences) according to the manufacturer's instructions.

***In vitro* assay for cGAS activity.** For *in vitro* synthesis of the cGAS reaction product 2 μM recombinant cGAS was mixed with 3 μM dsDNA (ISD) in Buffer A (100 mM NaCl, 40 mM Tris pH 7.5, 10 mM $MgCl_2$). Reaction was started by addition of 1 mM ATP and 1 mM GTP. After 2–4 h incubation at 37 °C the reaction was stopped and filtered using Amicon Ultra-15 filter devices (10,000 or 30,000 relative molecular mass cut-off).

Preparation of HEK293T cell lysates. HEK293T cells (0.33×10^6 per ml) were transfected with 3.2 μg plasmid using GeneJuice (Novagen). After 20 h cells were collected, washed twice with PBS and pelleted by centrifugation at 500*g* at 4 °C. The cell pellet was lysed (lysis buffer: 1 mM $CaCl_2$, 3 mM $MgCl_2$, 1 mM EDTA, 1% Triton X 100, 10 mM Tris pH 7.5) for 20 min at 4 °C. The cell lysate was briefly centrifuged (1,000*g*, 10 min, 4 °C) and the resultant supernatant was further purified via two sequential rounds of phenol-chloroform extraction. The extract was then filtered by centrifugation using Amicon Ultra-15 filter devices (10,000 or 30,000 relative molecular mass cut-off). In some experiments the final extract was concentrated via centrifugation under vacuum (Eppendorf Vacufuge).

Reversed phase-HPLC. Cell lysates and enzymatic reaction mixtures were applied to a 4.1 × 250 mm PRP-1 column (Hamilton) and separated in a linear gradient of 0% buffer B for 8 min, followed by an increase of buffer B from 0 to 75% in 62 min at a flow rate of 1 ml min^{-1}. Buffer A was 20 mM TEAB and buffer B 20 mM TEAB in 20% methanol. The product fractions were collected, evaporated and desalted by repeated co-evaporation with methanol. The residue was dissolved in PBS and the product concentration was determined by measuring ultraviolet absorbance (A_{260}). This HPLC method was mainly employed for preparative runs of cell lysates or *in vitro* synthesis products. Please note the differing retention times of this method compared to the analytical ESI-LC-MS runs.

Enzymatic reactions. 0.07 A_{260} of cGAMP(3′-5′) and cGAMP generated either *in vivo* or *in vitro* were dissolved in 6.5 μl incubation buffer and treated with 1 μl of the following enzymes: RNaseT1 (Fermentas, 100 mM Tris-HCl pH 7.4, 10 mM EDTA, 1 h, 37 °C), S1 nuclease (Fermentas, 50 mM NaOAc pH 4.5, 300 mM NaCl, 2 mM $ZnSO_4$, 1 h, 37 °C), RNase T2 (MoBiTec, 125 mM NH_4Ac pH 4.5, 1 h, 37 °C), SVPDE (Sigma, isolated from *Crotalus adamanteus*, 50 mM Tris-HCl pH 8.8, 10 mM $MgCl_2$, 30 min, 37 °C). The digestion products were analysed by TLC and ESI-LC-MS.

Thin layer chromatography (TLC). TLC was performed on 5 × 10 cm LuxPlate Si60 silica-covered glass plates (Merck). The samples (1–2 μl) were spotted onto the plate and separation was performed in *n*-propanol/ammonium hydroxide/water (11:7:2 v/v/v). The plate was air-dried and bands were visualized with a short-wavelength (254 nm) ultraviolet light source.

ESI-LC-MS and ESI-LC-MS/MS. All reagents used were purchased from Sigma Aldrich. The ESI-LC/MS analysis was performed using a Dionex Ultimate 3000 RS system (Thermo Fisher Scientific) coupled to an IonTrap mass spectrometer (LCQ Deca XP$^+$, Thermo Finnigan) equipped with an electrospray source operating in negative ionization mode. The ionization source parameters were set to: ion transfer capillary temperature 310 °C, spray voltage 4 kV and internal source fragmentation 15 kV. All samples were chromatographed on a Waters XBridge C18 OST column (2.1 × 50 mm; 2.5 μm particle size) at 30 °C column temperature. Separation of the analytes was achieved using a gradient of 10 mM TEAB in water as eluent A and 10 mM TEAB in 20% MeOH as eluent B with a flow rate of 0.25 ml min^{-1}. The HPLC gradient starts at 0% B, hold for 3 min and then increases over 16.5 min to 90% B.

Full-scan mass spectrometry spectra were acquired in a mass range from m/z 150 to 1,000 with isotopic resolution for the singly charged molecular ions. Tandem MS-MS and MS-MS-MS spectra were recorded from isolated ions in the ion trap applying collision induced dissociation (CID) applying helium as collision gas. For tandem MS-MS spectra the singly charged molecular ion was isolated and subsequently fragmented with 28% normalized collision energy. For tandem MS-MS-MS spectra the G-depurinated daughter ion of the cyclic dinucleotides with $m/z = 522.0$ (−guanine base) generated in the first CID fragmentation stage, was isolated and subsequently fragmented with 30% normalized collision energy.

Periodate oxidation assay. 0.1 mg ml^{-1} of the dinucleotide fractionated from cell culture lysates was incubated with 20 mM sodium periodate for 60 min at room temperature in the dark. After the incubation 10-vol% of 2 M triethylammonium acetate was added to the mixture that then was analysed by ESI-LC-MS.

2′→3′ isomerization assay. Approximately 0.1 mg ml^{-1} of the dinucleotide fractionated from cell culture lysates was incubated for 2 h at 90 °C in the presence of 10 mM EDTA and 20 mM Tris-HCl at pH 8. After the incubation 10-vol% of 2 M triethylammonium acetate was added to the mixture that then was analysed by LC-MS.

Chemical synthesis of >Gp(2′-5′)Ap(3′-5′)> and >Gp(3′-5′)Ap(3′-5′)>. The chemical synthesis of >Gp(2′-5′)Ap(3′-5′)> and >Gp(3′-5′)Ap(3′-5′)> was performed according to the strategy described in ref. 24 using the commercially available 3′-TBDMS protected 2′-guanosine phosphoramidite (Chemgenes) or 2′-TBDMS protected 3′-guanosine phosphoramidite (Sigma-Aldrich) for introduction of the 2′-5′ and 3′-5′ phosphodiester bond linkage, respectively. The 3′-adenosine phosphoramidite and all other reagents were purchased from Sigma-Aldrich. After base deprotection and removal of the TBDMS protecting groups >Gp(2′-5′)Ap(3′-5′)> or >Gp(3′-5′)Ap(3′-5′)> was purified by RP-HPLC as described above and the product was verified by ESI-LC-MS/MS.

Differential scanning fluorometry. Purification of human and murine STING ligand binding domains and differential scanning fluorometry to evaluate their thermal stabilization by nucleotide ligands was performed as previously described[16].

23. Mizushima, S. & Nagata, S. pEF-BOS, a powerful mammalian expression vector. *Nucleic Acids Res.* **18,** 5322 (1990).
24. Gaffney, B. L., Veliath, E., Zhao, J. & Jones, R. A. One-flask syntheses of c-di-GMP and the [R_p,R_p] and [R_p,S_p] thiophosphate analogues. *Org. Lett.* **12,** 3269–3271 (2010).

2.4. Structural mechanism of cytosolic DNA sensing by cGAS

2.4.1. Summary

cGAS is a PRR responsible for sensing cytosolic DNA. Upon DNA binding cGAS utilizes ATP and GTP to produce the second messenger 2,3-cGAMP that activates STING and downstream signalling. In this publication we were able to solve the crystal structures of apo- and substrate bound cGAS-Mab21 domain. This enabled us to explain the first step in the cyclization reaction that leads to the formation of 2,3-cGAMP and the role of cGAS as broad range DNA sensor.

The Mab21 domain of cGAS harbours the dinucleotide cyclization activity. We were able to solve the structure of the apo domain using a SAD (Single Wavelength Anomalous Dispersion) phasing approach with a dataset obtained from crystals of selenomethionine-labelled protein. The structure comprises two lobes. Lobe 1 possesses an NTase fold and lobe 2 consists of a bundle four α-helices. Both lobes are separated by the active site cleft. On the opposing site of the active site cleft the two lobes are connected by a long α-helix (spine helix) that flanks a flat highly positively charged surface called platform. A peculiar protrusion consisting of highly conserved histidines and cysteines coordinating a zinc^{2+} ion (Zn-thumb) flanks the platform on the other side.

Using a mutant construct that is incapable of catalyzing the cyclization reaction we were additionally able to solve the structure of porcine cGAS-Mab21 domain trapped in the substrate bound state. In this state, the dsDNA is positioned along the platform. It is stabilized by various side- and main-chain contacts of platform and spine helix residues with the sugar-phosphate backbone. Further stabilization is achieved by inserting two arginine fingers into the minor groove. The Zn-thumb is located near the major groove. This binding mode is sequence independent and explains the role of cGAS as broad range DNA sensor. In the active site, we could identify the bound ATP and GTP nucleotides. Their orientation in respect to each other explains the first step of the cyclization reaction which leads to the formation of a linear intermediate.

Further, comparison of apo- and holo cGAS revealed that dsDNA binding causes the nicking of the spine helix. The resulting movement is translated into a slight closure and the rearrangement of the active site cleft. We termed this the “switching mechanism”. It describes the transition from the inactive state to the active state of cGAS. The crystal structure derived models were further supported by EMSAs (Electrophoretic Mobility Shift Assays) and TLC (Thin Layer Chromatography) coupled activity assays *in vitro* and by luciferase activity based reporter assays *in vivo*.

In addition, we compared the cGAS-Mab21 domain with the RNA sensor molecule OAS1. Both structures are remarkably similar. But in contrast to OAS1, cGAS binds B-form DNA instead of A-form RNA. We hypothesize that the Zn-thumb which is only present in cGAS acts as a “ruler” that contributes to the DNA specificity of cGAS. Nevertheless, because of the similarities of our results and the existing model of OAS1 we suggested that DNA and RNA sensing could be evolutionarily connected.

In this publication we could solve the structure of porcine cGAS-Mab21 domain in apo- and the substrate bound form. These structures are in accordance with the role of cGAS as broad range DNA sensor and explain the first step in the formation of 2,3-cGAMP. We describe a switching mechanism that is the basis for the transition of the inactive form of cGAS to the active form. Furthermore, we structurally unify RNA and DNA sensing in the innate immune system and propose an evolutionary connection.

2.4.2. Contribution

The author of this thesis contributed equally to this publication. He cloned cGAS mutants for cell based assays and for the *in vitro* experiments. He purified cGAS mutants for EMSAs, TLC coupled activity assays and for crystallization. He helped with performing the TLC coupled activity assays. Additionally, he crystallized the substrate bound Mab21 domain, collected the diffraction data and solved the structure.

2.4.3. Paper

The following paper was published 2013 in nature.

Structural mechanism of cytosolic DNA sensing by cGAS

Filiz Civril*, Tobias Deimling*, Carina C. De Oliveira Mann, Andrea Ablasser, Manuela Moldt, Gregor Witte, Veit Hornung & Karl-Peter Hopfner

Nature 2013 June 20; **498**(7454): 332 - 7

*: equal contribution

ARTICLE

doi:10.1038/nature12305

Structural mechanism of cytosolic DNA sensing by cGAS

Filiz Civril[1]*, Tobias Deimling[1]*, Carina C. de Oliveira Mann[1], Andrea Ablasser[2], Manuela Moldt[1], Gregor Witte[1], Veit Hornung[2] & Karl-Peter Hopfner[1,3]

Cytosolic DNA arising from intracellular bacterial or viral infections is a powerful pathogen-associated molecular pattern (PAMP) that leads to innate immune host defence by the production of type I interferon and inflammatory cytokines. Recognition of cytosolic DNA by the recently discovered cyclic-GMP-AMP (cGAMP) synthase (cGAS) induces the production of cGAMP to activate the stimulator of interferon genes (STING). Here we report the crystal structure of cGAS alone and in complex with DNA, ATP and GTP along with functional studies. Our results explain the broad DNA sensing specificity of cGAS, show how cGAS catalyses dinucleotide formation and indicate activation by a DNA-induced structural switch. cGAS possesses a remarkable structural similarity to the antiviral cytosolic double-stranded RNA sensor 2′-5′oligoadenylate synthase (OAS1), but contains a unique zinc thumb that recognizes B-form double-stranded DNA. Our results mechanistically unify dsRNA and dsDNA innate immune sensing by OAS1 and cGAS nucleotidyl transferases.

Recognition of pathogen- or danger-associated molecular patterns (PAMPs or DAMPs) is crucial for host defence. Innate immunity ensures this recognition through germline-encoded pattern recognition receptors (PRRs) and triggers signalling cascades that result in production of proinflammatory cytokines and type I interferons (IFN-α and IFN-β)[1,2]. Cytosolic DNA arising from intracellular bacteria or viral infections is a powerful PAMP and is also implicated as a DAMP in autoimmune diseases[1,3,4]. Over the past years, a variety of PRRs for cytosolic DNA have been reported: DNA-dependent activator of IFN-regulatory factors (DAI, also known as ZBP1)[5], absent in melanoma 2 (AIM2)[6–8], RNA polymerase III[9,10], leucine-rich repeat (in Flightless I) interacting protein-1 (LRRFIP1)[11], DExD/H box helicases (DDX41, DHX9 and DHX36)[12,13] and IFN-inducible protein IFI16[14]. However, these PRRs are either cell-type- or DNA-sequence-specific, are possible accessory factors (DExD/H proteins), or trigger different pathways such as caspase-1 activation (AIM2) or a β-catenin-dependent signalling pathway (LRRFIP1)[15].

Although the DNA sensor for type I IFN production with broad specificity and cell distribution was not identified until recently, it was known that IRF3 and NFκB activation in response to DNA requires STING (stimulator of interferon genes, encoded by gene *TMEM173* the protein is also known as MITA, MPYS or ERIS), a transmembrane protein that is resident on the endoplasmic reticulum[16–18]. STING colocalizes with DNA *in vivo* but binds DNA only with low affinity *in vitro*[19], suggesting the presence of an additional sensor. Furthermore, STING is a direct PRR for cyclic dinucleotides such as c-di-AMP and c-di-GMP[20], which are signalling molecules in prokaryotes and trigger IFN in response to, for example, intracellular bacteria[21,22].

Recent results identified human c-GMP-AMP (cGAMP) synthase (cGAS, also known as C6ORF150 and male abnormal 21 domain containing 1 (MB21D1)) as a broad-specificity cytosolic DNA sensor[23]. In the presence of DNA cGAS produces cGAMP, which is an endogenous second messenger that activates STING[18], explaining how STING can stimulate IFN in response to both cyclic dinucleotides and DNA. To reveal the mechanism of DNA-stimulated cGAMP synthesis, we determined the crystal structure of porcine cGAS$^{\mathrm{Mab21}}$ (residues 135–497, comprising the highly conserved, DNA-stimulated nucleotidyl transferase (NTase) domain) with and without a 14-mer dsDNA ligand and nucleotide substrates, along with functional studies *in vitro* and in living cells.

Crystal structure of cGAS$^{\mathrm{Mab21}}$

cGAS is a 60 kDa protein composed of an unstructured, not well conserved amino-terminal stretch of approximately 130–150 residues followed by a highly conserved Mab21 domain that belongs to the nucleotidyl transferase (NTase) superfamily[24]. To overproduce and crystallize cGAS, it was necessary to genetically remove the unstructured N-terminal tail. The resulting cGAS$^{\mathrm{Mab21}}$ used in this study (residues 155/161–522 for human cGAS and residues 135–497 for porcine cGAS) possesses DNA-dependent dinucleotide synthesis activity in the presence of a 50-mer dsDNA that induces IFN in THP1 cells (Fig. 1a and Supplementary Fig. 1a, b). Whereas cGAS also produces cGAMP in the presence of a 40-mer dsDNA, no activity was observed when we omitted either GTP or ATP from the reaction mixture or substituted dsDNA with single-stranded DNA (Supplementary Fig. 1a).

We determined the crystal structure of porcine cGAS$^{\mathrm{Mab21}}$ by single-wavelength anomalous dispersion to 2.5 Å resolution using a selenomethionine derivative. After density modification, we could build an initial model, which was completed and refined against the 2.0 Å resolution native data, resulting in good *R*-factors and stereochemistry (Supplementary Fig. 1c and Supplementary Table 1).

The Mab21 domain of cGAS comprises two lobes, separated by a deep cleft (Fig. 1b). Lobe 1 possesses the NTase fold with a two-leaved highly twisted β-sheet (β1–β8) that is flanked on the outside by two long α-helices (αA and αB). At the inner side, lining the cleft, β1 and β6 harbour the signature catalytic site residues (E200, D202, D296) of the NTase superfamily that coordinate the catalytic Mg^{2+} ions and nucleotides. Lobe 2 is a bundle of four α-helices (αE–αH), connected

---

[1]Department of Biochemistry and Gene Center, Ludwig-Maximilians-University, 81377 Munich, Germany. [2]Institute for Clinical Chemistry & Clinical Pharmacology, Unit for Clinical Biochemistry, University Hospital, University of Bonn, 53127 Bonn, Germany. [3]Center for Integrated Protein Sciences, 81377 Munich, Germany.
*These authors contributed equally to this work.

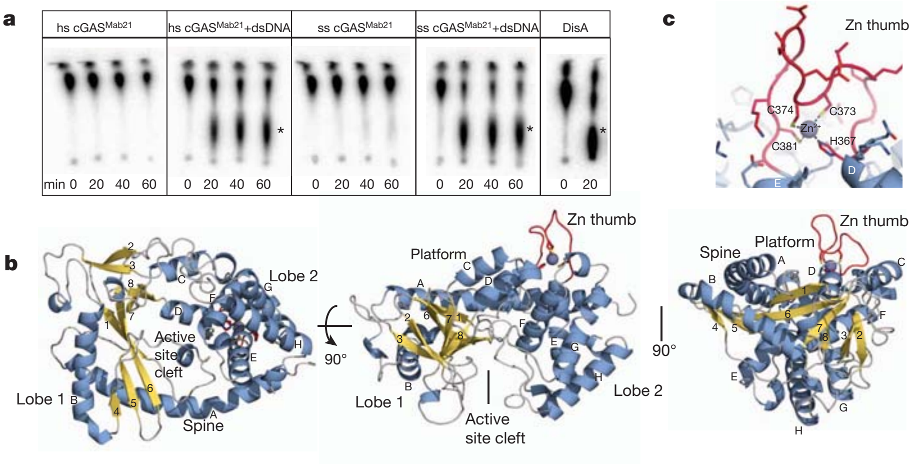


Figure 1 | Crystal Structure of cGAS$^{\text{Mab21}}$. **a**, Activity assays of human and porcine cGAS$^{\text{Mab21}}$ alone or in presence of dsDNA. *Bacillus subtilis* DisA, a c-di-AMP synthase is used as positive control. The dinucleotide products are indicated with asterisks. **b**, Side and top views of cGAS$^{\text{Mab21}}$. The model is shown as ribbon representation with annotated domains and secondary structure (blue α-helices, yellow β-strands). **c**, Close-up view of the 'zinc thumb'.

to lobe 1 by a long 'spine' (αA), two linker helices (αC, αD) and by a long active site loop connecting αA and β1.

The molecular surface opposite the active site is a fairly flat, slightly concave 'platform', formed predominantly by αA, αC, αD and the nucleotide-binding loop. An intriguing protrusion (residues 367–382) is situated at one end of the platform. This protrusion contains highly conserved histidine and cysteines (H367, C373, C374 and C381), which together coordinate a Zn^{2+} ion (Fig. 1c). We denote this loop 'Zn thumb'. Its sequence is inserted between lobes 1 and 2 and is a highly conserved and characteristic feature of cGAS orthologues (Supplementary Fig. 1d), indicating an important functional role.

The cGAS–DNA–GTP–ATP complex

To reveal the structure of the activated conformation of cGAS, we co-crystallized cGAS$^{\text{Mab21(td)}}$ with a self-complementary 14-mer oligonucleotide, ATP, GTP and MgCl$_2$. To trap an activated conformation of cGAS$^{\text{Mab21}}$ with DNA and bound nucleotides we mutated the NTase catalytic residues E200 and D202 to Q and N, respectively, thereby preventing catalysis during crystallization. The resulting transferase-deficient (td) variant is denoted cGAS$^{\text{Mab21(td)}}$. The structure of the cGAS$^{\text{Mab21(td)}}$–DNA–GTP–ATP complex was determined by molecular replacement using the coordinates for apo cGAS$^{\text{Mab21}}$ as search model. $2F_o - F_c$ and $F_o - F_c$ maps revealed interpretable density for 13 out of 14 base pairs of the dsDNA duplex and for both nucleotides bound at the active site (Supplementary Fig. 2). The structure was refined at 3.1 Å resolution, resulting in a model with good *R*-factors and stereochemistry (Supplementary Table 1).

DNA is bound along the platform between the spine on one side and the Zn thumb on the other side (Fig. 2a). cGAS binds DNA predominantly by sequence-independent interactions to both phosphate-sugar backbone strands along the minor groove (Fig. 2b, c). Hereby, cGAS binds seven nucleotides at the core of the platform, which are recognized by at least eleven residues via specific side- and/or main-chain contacts. In addition to the phosphate and sugar contacts, two arginine fingers (R150 and R192) are inserted into the minor groove, additionally stabilizing the interaction in a fairly sequence-independent manner. Besides binding to the array of conserved positively charged residues at the bottom of the platform, DNA is also bound by the spine and the Zn thumb. The continuous helix of the spine in apo-cGAS$^{\text{Mab21}}$ is interrupted in the DNA complex and a DNA backbone phosphate is bound at the central kink of the spine helix. On the other side of the platform, the Zn thumb contacts the DNA backbone near the major groove. We do not see close, direct polar contacts between Zn thumb and DNA, but do not want to rule out water-mediated interactions here (Supplementary Fig. 2a).

The Zn thumb does not substantially change conformation or location between apo and DNA-bound cGAS. It seems to be a rather rigid element, in which the zinc ion serves as a structural stabilizer of the protruding loop, similar to Zn^{2+} in regulatory domains of RIG-I-like receptors[25]. The location of the Zn thumb at the backbone near the major groove suggests that it may assist in binding to B-form DNA. In support of this, we do not see a substantial perturbation of the bound DNA from canonical B-form DNA.

Altogether, our structure suggests a specific recognition of B-form dsDNA by cGAS through an extended B-DNA binding platform and

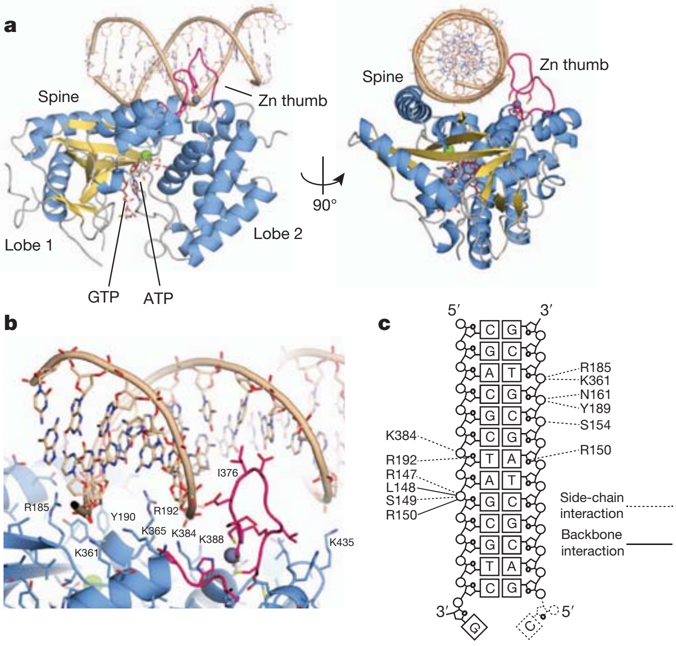


Figure 2 | The cGAS$^{\text{Mab21}}$–DNA–GTP–ATP complex. **a**, Side and top views of cGAS$^{\text{Mab21}}$ (colour code of Fig. 1b) in complex with dsDNA (brown), GTP and ATP (ruby stick models). DNA binds along the platform between spine and Zn thumb. **b**, Close-up view of the DNA binding site with selected annotated residues. DNA is bound mainly via the minor groove. A notable exception is the Zn thumb near the major groove. **c**, Schematic representation of DNA–cGAS contacts.

flanking 'Zn thumb' across both lobes of the enzyme. The observed mode of binding is consistent with the key role of cGAS in sensing very different types of DNA in a sequence-independent manner[18,23].

Structure–function analysis

To validate the structural results, we mutated several conserved positively charged residues at the DNA-binding platform of human cGAS, two active site residues, two zinc ligands in the Zn thumb, or the entire Zn thumb and tested for nucleotidyl-transferase activity *in vitro* by thin-layer chromatography (TLC) (Fig. 3a). cGAS produces a product that migrates approximately in the range of c-di-AMP synthesized by DisA[26], consistent with formation of a dinucleotide. The conserved active site residues of NTases (human E225+D227; porcine E200+D202 and human G212+S213) are essential for *in vitro* activity of cGAS$^{\mathrm{Mab21}}$. Moreover, mutation of conserved positively charged residues at the centre and flanking regions of the platform (K173+R176 and K407+K411) either diminish or abolish activity, in accordance with this site being important for DNA sensing. Finally, disruption of the zinc-binding site of the thumb (human C396+C397, Zn thumbless) abolishes DNA-induced NTase activity *in vitro*, highlighting the functional importance of the conserved Zn thumb in DNA binding.

To test the effect of active site, platform and thumb mutations in living cells, we measured the transactivation of an IFN-β promoter reporter by transiently expressing human cGAS variants in HEK293T cells that stably expressed murine STING (Fig. 3b). Induction of IFN-β by cGAS$^{\mathrm{Mab21}}$ (human cGAS$^{155-552}$) in these cells is only moderately reduced compared to wild-type cGAS, showing that the Mab21 domain structurally addressed in this study is the catalytic active functional core of the sensor. The activity of full-length cGAS was abolished when residues of the NTase active site were mutated (E225Q/A+D227N/A or G212A+S213A). Mutating charged platform residues (K173A+R176A; K407A+K411A) substantially reduced the activity of cGAS in living cells. Likewise, disrupting the zinc-binding site of the thumb (C396A+C397A, Zn thumbless) severely compromised cGAS activity. These data validate the *in vitro* biochemical data and emphasize the importance of the structure-derived motifs and elements in living cells.

To see whether Zn thumb and conserved platform surface residues are important for dsDNA binding and activity, we performed electrophoretic mobility shift assays (Fig. 3c). Both porcine and human wild-type cGAS$^{\mathrm{Mab21}}$ bind efficiently to dsDNA and, surprisingly, also to dsRNA (Supplementary Fig. 3a, c). The mutations in platform and thumb either did not affect DNA/RNA binding under these conditions, or reduced but did not abolish it (Supplementary Fig. 3b). However, both mutants fail to show DNA-stimulated activity under conditions where they still bind DNA, and dsRNA fails to stimulate activity under conditions where it binds robustly to the protein (Supplementary Fig. 3c, d). Thus, although these analyses validate the functional relevance of the DNA binding platform and Zn thumb on activating cGAS, they suggest that DNA or RNA interactions per se are not sufficient to activate the enzyme, indicating for instance the necessity for a precise DNA-induced structural switch.

NTase and DNA induced structural switch

To reveal the mechanism of activation of cGAS by DNA, we first analysed the NTase mechanism. We see clear electron density for two nucleotide triphosphate moieties (Supplementary Fig. 2b). The two bases partially stack in an approximately 90° rotated orientation and inserted into a hydrophobic/aromatic pocket, sandwiched between

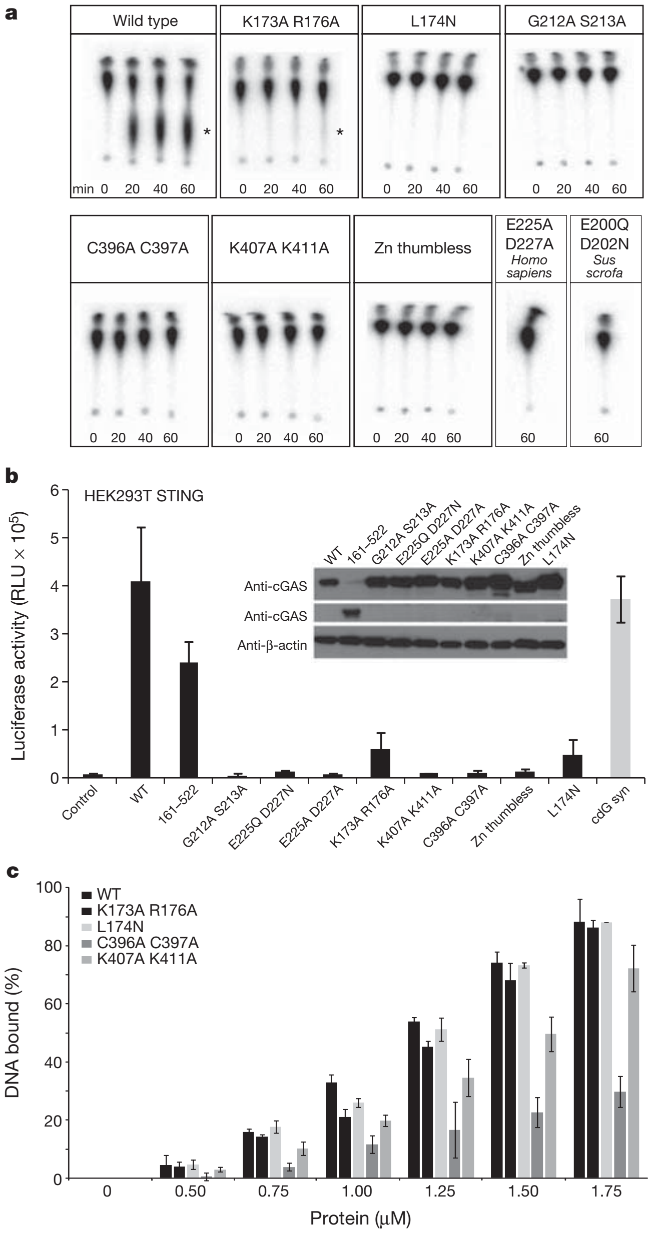


Figure 3 | Platform and Zn thumb are involved in dsDNA-dependent activity. a, NTase assays performed with different cGAS$^{\mathrm{Mab21}}$ mutants (2 μM) in presence of 3 μM dsDNA (50-mer). Human wild-type cGAS$^{\mathrm{Mab21}}$ (positive control) synthesizes dinucleotide, DNA binding site mutant K173A+R176A show reduced activity. K407A+K411A DNA binding site mutant, C396A+C397A Zn thumb mutant, Zn thumbless, L174N structural switch mutant, active site mutants E200Q+D202N of porcine cGAS$^{\mathrm{Mab21}}$ and E225A+D227A and G212A+S213A of human cGAS$^{\mathrm{Mab21}}$ are inactive. The asterisk indicates the dinucleotide product. **b**, IFN-β stimulation of cGAS mutants in HEK293T cells stably expressing murine STING. HEK293T cells were transfected with plasmids encoding indicated constructs along with the IFN-β promoter reporter plasmid pIFN-β-GLUC. Luciferase activity is plotted: mean ± s.d. ($n = 3$). Both full-length and the crystallized region (cGAS$^{\mathrm{Mab21}}$ human 155–522) induce IFN-β promoter transactivation. Active site mutations (G212A+S213A and E225Q/A+D227N/A) abolish IFN-β stimulation. DNA-binding site mutants (K173A+R176A, K407A+K411A), Zn thumb mutants (C396A+C397A, Zn thumbless) and structural switch mutant (L174N) either reduce or abolish IFN-β stimulation. Empty vector was used as negative control whereas cyclic-di-GMP synthase (cdG syn) expressing vector was used as positive control. Inset: western blot showing wild-type and mutant protein levels with β-actin as loading control. **c**, Electrophoretic mobility shift analysis of 50-mer dsDNA (0.2 μM) bound to cGAS$^{\mathrm{Mab21}}$ mutants at indicated concentrations. Plotted bars, mean ± s.d. ($n = 3$). Whereas K407A+K411A DNA binding site mutant and C396A+C397A Zn thumb mutant show slightly reduced but not impaired affinity to dsDNA, no detectable binding change was observed with the other mutants.

I298 (lobe 1) and Y413 (lobe 2). The current resolution of the diffraction data does not allow us to unambiguously determine which base is adenine and which guanine. Binding of R353 at nucleobase 1 (the 'receiving substrate' of NTases) near O6 and N7 would argue for this being guanine. In general, nucleobase 1 (interpreted as guanine here) is in hydrogen bonding distance to S355, S357 and T186, suggesting that this nucleotide is specifically recognized. In contrast, we do not observe direct hydrogen-bonding contacts of the protein to nucleobase 2 (the 'transferred' nucleotide in NTases; interpreted as adenine here). Nevertheless, this recognition might be mediated via water molecules such as in 3′ terminal uridylyl transferases[27].

The structure provides a mechanism for attack of nucleotide 1 on nucleotide 2, consistent with the mechanism of other NTases, for example, CCA adding enzyme[28]. The triphosphate chain of nucleotide 2 is well coordinated via S188 (lobe 1), S412 (lobe 2) and Mg^{2+} bound to E200 (Q in $cGAS^{Mab21(td)}$) and D202 (N in $cGAS^{Mab21(td)}$). As a consequence, the relative orientation of lobes 1 and 2 is important for the phosphate coordination of nucleotide 2. In our conformation, the α-phosphate of nucleotide 2 is well placed and oriented to promote nucleophilic attack of the sugar 2′ OH from nucleotide 1 to form the 2′-5′ linkage (Fig. 4a, see ref. 29). The attacking OH of nucleotide 1 is polarized and activated by D296, consistent with the conserved features of NTases[24]. A second Mg^{2+} could be important for this catalytic step. However, distinct localization will require higher resolution.

cGAS is proposed to form a cyclic-dinucleotide, which would require a second catalysis step and an additional attack of the OH of nucleotide 2 on the phosphate of nucleotide 1. Such an attack will require an almost 180° flip of the sugar moiety of nucleotide 2 to place its α-phosphate appropriately. In principle this is possible: in the course of our studies we determined the crystal structure of $cGAS^{Mab21}$ bound to UTP in the absence of DNA and do observe an appropriate flip of the sugar moiety (Supplementary Fig. 4). In any case, our structure satisfactorily explains the catalysis of formation of a specific (at present linear) dinucleotide by cGAS, but formation of a cyclic dinucleotide needs to be addressed in future studies.

To reveal a potential activation mechanism of cGAS, we superimposed apo-cGAS, $cGAS^{Mab21}$–UTP and $cGAS^{Mab21(td)}$–DNA–GTP–ATP complex (Fig. 4b, c and Supplementary Fig. 5a, b). We used $cGAS^{Mab21}$–UTP because UTP binding orders the β-sheets on lobe 1 and we can also visualize conformational changes specifically induced by dsDNA rather than the nucleotides.

Although UTP binding to cGAS ordered to some extend the nucleotide-binding loop in the active site, it did not substantially change the overall structure and active site geometry of cGAS (Supplementary Fig. 5b). In contrast, DNA phosphate binding to the spine (Fig. 4b) triggers a substantial structural switch in the spine helix (Fig. 4c) that closes lobes 1 and 2 and rearranges the active site loop, allowing magnesium coordinating of E200 to position and activate nucleotide 2.

To test the role of this DNA-induced structural switch we mutated human L174 to N. L174 (porcine L148) is repositioned in response to DNA binding to stabilize the nucleotide-binding loop, but does not directly bind DNA or NTPs (Supplementary Fig. 5c). Although L174N shows fairly normal DNA binding (Fig. 3c and Supplementary Fig. 3b), it lacks DNA-stimulated cGAMP synthetase activity *in vitro* (Fig. 3a) and shows decreased interferon stimulation in cells (Fig. 3b). Thus, the structural and biochemical data suggest that cGAS is activated by a DNA-induced structural switch that rearranges the NTase active site.

Conclusion

Here we provide the structure and mechanism of activation of the cytosolic DNA sensor cyclic-GMP-AMP synthase that readily explain the synthesis of a linear dinucleotide intermediate by cGAS in response to DNA binding. The backbone binding of a canonical B-DNA by cGAS is consistent with a broad specificity innate immune PRR for cytosolic DNA and the structural elements of cGAS such as the positioning of residues involved in minor-groove binding, arginine fingers and the Zn thumb suggest that cGAS specifically responds to B-form DNA. This might explain the function of other innate immune DNA sensors to detect non-canonical DNA structures, such as DAI[5]. A structural switch transmitted by proper B-form DNA binding to the active site could also explain the lack of activation by dsRNA or in mutants that still bind DNA: slightly different conformations of RNA-bound or DNA-bound mutant cGAS would not trigger robust cGAMP synthesis as even small differences in the active site geometry can strongly affect catalytic rates of enzymes.

In future, it will be important to address the specificity for other DNA structures in the activation of cGAS in more detail to see which types of DNA structures can activate cGAS. It will also be important to investigate additional requirements for efficient DNA sensing *in vivo*, because although shorter dsDNA molecules can stimulate $cGAS^{Mab21}$ *in vitro*, DNA larger than 50-mer is required for efficient IFN stimulation *in vivo*[14,19]. One possibility is that fraying of shorter DNA molecules prevents efficient stimulation or that the positively charged N terminus contributes to sensing of longer DNA molecules. In addition, STING might have a direct role in DNA binding in a larger context *in vivo*[19], although we do not see strong DNA binding *in vitro* and IFN stimulation in response to DNA in HEK293T cells in the absence of cGAS (Supplementary Fig. 6).

Interestingly, cGAS has remarkable fold similarity to the antiviral protein oligoadenylate synthase 1 (OAS1)[30,31] (Fig. 5). OAS1 synthesizes 2′-5′ linked oligoadenylate chains in response to binding to

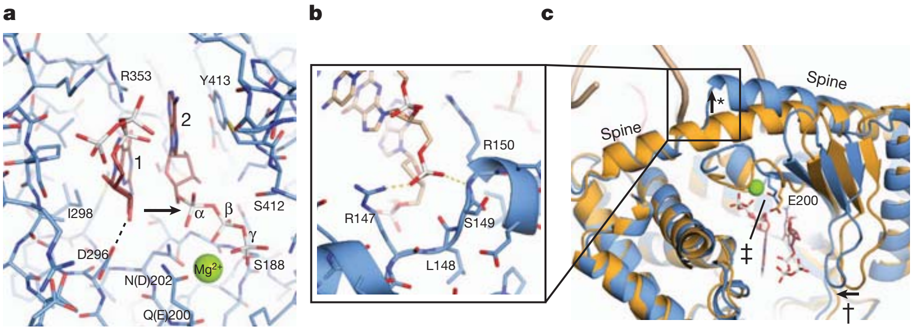


Figure 4 | NTase and DNA-induced structural switch. **a**, Close-up view of the NTase active site. Selected residues that are implicated in binding and catalysis are annotated. Both base moieties partially stack to each other and are further bound by stacking to Y413 and recognition by R353. E200 (mutated to Q in our structure) and D202 (mutated to N in our structure) bind an active site magnesium that coordinates phosphates of nucleotide 2. The attacking OH of nucleotide 1 is activated by D296 for nucleophilic attack on the α-phosphate of nucleotide 2 (arrow). **b**, Close-up view of DNA backbone phosphate binding at the spine. **c**, This DNA phosphate binding triggers a change in the spine helix (*), which allows a closure of the active site cleft (†) and repositioning of the substrate binding loop for Mg^{2+} coordination of E200 (‡).

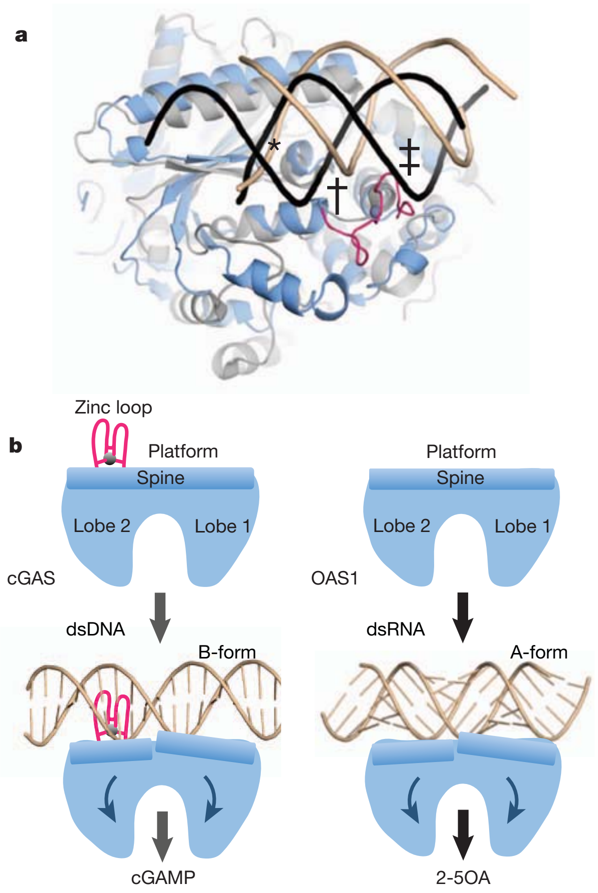


Figure 5 | Model for DNA sensing by cGAS. a, Superposition of cGAS–DNA (blue) with OAS1–RNA (grey) shows key elements for nucleic duplex selectivity. Both enzymes bind one DNA (brown)/RNA (black) backbone at the same protein site (*). The Zn thumb specifically recognizes the position of the second DNA strand in B-form (†).However, it would clash with A-form RNA/DNA (‡). **b**, Unified activation model for cytosolic double-stranded nucleic acid sensing by cGAS and OAS1 NTases by a ligand induced structural switch. 2-5OA, 2′-5′ linked oligoadenylate chains.

cytosolic dsRNA. The structural similarity not only embraces the overall fold, several active site features and arrangement of lobes 1 and 2, but also certain structural elements of the platform, including the long 'spine' helix. Like cGAS, OAS1 binds dsRNA along the 'platform' and triggers a structural change that is transmitted to the active site[31]. However, whereas OAS1 is activated by A-form RNA, cGAS is activated by B-form DNA. The Zn thumb in cGAS, missing in OAS1, probably acts as a molecular 'ruler' to specifically trigger activation in response to B-form but not A-form nucleic acids (Fig. 5). Despite these differences, cGAS shows a structural switch induced by dsDNA that is very similar to that of OAS1 induced by dsRNA[31] (Fig. 5). Thus, our results structurally unify dsDNA and dsRNA sensing by cGAS and OAS1 NTases, respectively, in the innate immune system and suggest that both processes are evolutionarily connected. *Note added in proof*: After submission of the revised version of this manuscript, Gao *et al.*[32] reported related structures of cGAS and its complexes with DNA and nucleotides.

METHODS SUMMARY

Proteins were produced in *Escherichia coli* and purified by affinity, ion exchange and size exclusion chromatography. Apo, UTP- and DNA–ATP–GTP-bound cGAS$^{\mathrm{Mab21}}$ and its catalytic inactive form were crystallized by hanging or sitting drop vapour diffusion. The structure of apo cGAS$^{\mathrm{Mab21}}$ was determined by single-anomalous dispersion phasing on selenomethionine derivatized protein. The other structures were determined by molecular replacement using apo cGAS$^{\mathrm{Mab21}}$ as search model. NTase assays were performed by thin layer chromatography and phosphor imaging. DNA and RNA binding were assessed by electrophoretic mobility shift assays. Analysis of cGAS mutants in living cells were performed in HEK 293T cells stably expressing full-length murine STING and transfected with an IFN-β promoter reporter plasmid.

Full Methods and any associated references are available in the online version of the paper.

Received 21 March; accepted 20 May 2013.

1. Rathinam, V. A. K. & Fitzgerald, K. A. Cytosolic surveillance and antiviral immunity. *Curr. Opin. Virol.* **1,** 455–462 (2011).
2. Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. *Cell* **140,** 805–820 (2010).
3. Keating, S. E., Baran, M. & Bowie, A. G. Cytosolic DNA sensors regulating type I interferon induction. *Trends Immunol.* **32,** 574–581 (2011).
4. Krug, A. Nucleic acid recognition receptors in autoimmunity. *Handb. Exp. Pharmacol.* **183,** 129–151 (2008).
5. Takaoka, A. *et al.* DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* **448,** 501–505 (2007).
6. Bürckstümmer, T. *et al.* An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. *Nature Immunol.* **10,** 266–272 (2009).
7. Fernandes-Alnemri, T., Yu, J. W., Datta, P., Wu, J. & Alnemri, E. S. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* **458,** 509–513 (2009).
8. Hornung, V. *et al.* AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* **458,** 514–518 (2009).
9. Ablasser, A. *et al.* RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. *Nature Immunol.* **10,** 1065–1072 (2009).
10. Chiu, Y. H., Macmillan, J. B. & Chen, Z. J. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell* **138,** 576–591 (2009).
11. Yang, P. *et al.* The cytosolic nucleic acid sensor LRRFIP1 mediates the production of type I interferon via a β-catenin-dependent pathway. *Nature Immunol.* **11,** 487–494 (2010).
12. 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* **107,** 15181–15186 (2010).
13. Zhang, Z. *et al.* The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. *Nature Immunol.* **12,** 959–965 (2011).
14. Unterholzner, L. *et al.* IFI16 is an innate immune sensor for intracellular DNA. *Nature Immunol.* **11,** 997–1004 (2010).
15. Rathinam, V. A. & Fitzgerald, K. A. Innate immune sensing of DNA viruses. *Virology* **411,** 153–162 (2011).
16. Ishikawa, H., Ma, Z. & Barber, G. N. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* **461,** 788–792 (2009).
17. Ishikawa, H. & Barber, G. N. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* **455,** 674–678 (2008).
18. Wu, J. *et al.* Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science* **339,** 826–830 (2013).
19. Abe, T. *et al.* STING recognition of cytoplasmic DNA instigates cellular defense. *Mol. Cell* **50,** 5–15 (2013).
20. Burdette, D. L. *et al.* STING is a direct innate immune sensor of cyclic di-GMP. *Nature* **478,** 515–518 (2011).
21. McWhirter, S. M. *et al.* A host type I interferon response is induced by cytosolic sensing of the bacterial second messenger cyclic-di-GMP. *J. Exp. Med.* **206,** 1899–1911 (2009).
22. Woodward, J. J., Iavarone, A. T. & Portnoy, D. A. c-di-AMP secreted by intracellular *Listeria monocytogenes* activates a host type I interferon response. *Science* **328,** 1703–1705 (2010).
23. 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).
24. Kuchta, K., Knizewski, L., Wyrwicz, L. S., Rychlewski, L. & Ginalski, K. Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. *Nucleic Acids Res.* **37,** 7701–7714 (2009).
25. Cui, S. *et al.* The C-terminal regulatory domain is the RNA 5′-triphosphate sensor of RIG-I. *Mol. Cell* **29,** 169–179 (2008).
26. Witte, G., Hartung, S., Buttner, K. & Hopfner, K. P. Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. *Mol. Cell* **30,** 167–178 (2008).
27. Stagno, J., Aphasizheva, I., Rosengarth, A., Luecke, H. & Aphasizhev, R. UTP-bound and apo structures of a minimal RNA uridylyltransferase. *J. Mol. Biol.* **366,** 882–899 (2007).
28. Xiong, Y. & Steitz, T. A. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* **430,** 640–645 (2004).
29. Ablasser, A. *et al.* cGAS produces a 2′-5′-linked cyclic dinucleotide second messenger that activates STING *Nature* http://dx.doi.org/10.1038/nature12306 (30 May 2013).

30. Hartmann, R., Justesen, J., Sarkar, S. N., Sen, G. C. & Yee, V. C. Crystal structure of the 2′-specific and double-stranded RNA-activated interferon-induced antiviral protein 2′-5′-oligoadenylate synthetase. *Mol. Cell* **12,** 1173–1185 (2003).
31. Donovan, J., Dufner, M. & Korennykh, A. Structural basis for cytosolic double-stranded RNA surveillance by human oligoadenylate synthetase 1. *Proc. Natl Acad. Sci. USA* **110,** 1652–1657 (2013).
32. 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).

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

Acknowledgements We thank A. Butryn for comments on the manuscript. We thank the Max-Planck-Crystallization facility for initial crystal screening and the Swiss Light Source, European Synchrotron Radiation Facility and the German electron synchrotron Petra III for beam time and on-site assistance. This work was funded by the National Institutes of Health (U19AI083025), the European Research Council Advanced Grant 322869, and the Center for Integrated Protein Science Munich (CIPSM) to K.-P.H., by DFG grant 3717/2-1 to G.W., by GRK1721 to K.-P.H. and G.W., by DFG grant SFB670 and ERC grant 243046 to V.H.; C.C.O.M. is supported by GRK1721.

Author Contributions F.C. crystallized and determined the structure of cGAS, performed biochemical assays, interpreted data and wrote the manuscript. T.D. crystallized and refined the DNA complex. C.C.O.M., A.A., T.D. and M.M. performed biochemical assays. G.W. performed biochemical assays, interpreted data and helped with structure determination. V.H. supervised the cell-based experiments and interpreted data. K.-P.H designed the research, helped with structure determination, interpreted data and wrote the manuscript.

Author Information Coordinates and structure factors have been deposited at the Protein Data Bank (4JLX, 4JLZ and 4KB6). 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. Correspondence and requests for materials should be addressed to K.-P.H. (hopfner@genzentrum.lmu.de).

METHODS

Constructs and cloning. The sequence encoding full-length or truncated *Homo sapiens* and *Sus scrofa* cGAS were amplified from total cDNA (courtesy of S. Bauersachs) and cloned into pIRESneo3 (Clontech) or a modified pET21 (Novagen), respectively. The mutants were generated by site-directed mutagenesis using PfuUltra (Stratagene). Zn thumbless mutant was created by replacing residues 390–405 (*Homo sapiens*) by three Gly-Ser replicates.

Protein production and purification. All proteins were produced in *E. coli* Rosetta (DE3) or B834 (DE3) strains for native or selenomethionine derivative proteins, respectively. Bacteria were grown until a D_{600} of 0.6 to 0.8 was reached and expression was induced at 18 °C for 16 to 18 h with 0.1 mM IPTG. Proteins were purified by Ni-NTA agarose resin and incubated with tobacco etch virus (TEV) protease (ratio 1:50) at 4 °C overnight to remove the 6xHis-MBP-tag. The proteins were further purified by cation exchange chromatography followed by size exclusion chromatography using a Superdex 200 column (GE Healthcare), equilibrated in 20 mM Tris pH 7.5, 150 mM NaCl and 1 mM DTT. Purified proteins were concentrated to 10 mg ml^{-1} for crystallization. Human STING 139–379 was purified as described[33]. All purified proteins were frozen in liquid N_2 and stored at -80 °C.

Crystallization of cGAS^{Mab21}. Purified porcine cGAS (10 mg ml^{-1}) was crystallized by hanging drop vapour diffusion in 20% PEG3350 and 200 mM sodium malonate. The crystals appeared after one day at 20 °C and were flash frozen after addition of glycerol to a final concentration of 15% (v/v). The selenomethionine derivatized protein was crystallized in 100 mM Bis-Tris propane pH 6.3, 18% PEG3350 and 200 mM sodium malonate and cryo protected with 20% ethane-1,2-diol before flash freezing. UTP bound crystals were obtained by adding 20 mM $MgCl_2$ and 1:10 (v/v) of 50 mM of nucleotide in 100 mM Tris pH 7.5 to the protein before crystallization.

For crystallizing the DNA–GTP–ATP–cGAS complex 20 mM $MgCl_2$, 2 mM of both nucleotides and 14 bp dsDNA (5′-CGACGCTAGCGTCG-3′) in a molar ratio of 1:1.2 protein:DNA were added to the inactive porcine cGAS$^{Mab21(td)}$ (E200Q+D202N) (10 mg ml^{-1}). Crystals were obtained by hanging drop vapour diffusion in 50 mM sodium cacodylate pH 7.0, 2.5 mM spermine, 60 mM $MgCl_2$ and 3% (v/v) PEG 400 after one day at 20 °C. The crystals were soaked in reservoir solution containing 25% (v/v) glycerol before flash freezing.

Data collection and refinement. X-ray diffraction data of cGAS and cGAS-UTP were collected at X06SA beamline (Swiss Light Source, Switzerland) and diffraction data of the cGAS$^{Mab21(td)}$–GTP–ATP–DNA complex were collected at PetraIII beamline P14 (EMBL/DESY, Hamburg, Germany) at 100 K. The selenomethionine derivative data were collected at the selenium peak wavelength ($\lambda = 0.97961$ Å). Data processing was carried out with XDS[34]. AutoSHARP was used to locate Se sites (SAD data set) and to produce an initial solvent flattened map[35]. An initial model was built using iterative cycles of Buccaneer[36] and ARP/wARP classic[37]. The model was optimized by alternating manual building with Coot[38] and refinement using Phenix[39] against a 2.0 Å native data set. The structure of UTP-bound cGAS and the DNA–GTP–ATP–cGAS complex structure were determined using molecular replacement with Phaser[40] and optimized by manual building with Coot and refinement with Phenix or Autobuster[41]. Data collection and refinement statistics are listed in Supplementary Table 1.

NTase assays. NTase assays were performed as described in ref. 26. Reaction mixtures with the indicated concentrations of protein and DNA (40-mer: 5′-GGATACGTAACAACGCTTATGCATCGCCGCCGCTACATCC-3′, 50-mer: 5′-GGATACGTAACAACGCTTATGCATCGCCGCCGCTACATCCCTGAGCTGAC-3′) (unless indicated 50-mer dsDNA is used) or RNA (sequence as 50-mer DNA) in 0.1 M NaCl, 40 mM Tris pH 7.5 and 10 mM $MgCl_2$ were started by addition of 100 µM ATP and 100 µM GTP containing 1:600 [α^{32}P]ATP and/or [α^{32}P]GTP (3,000 Ci mmol^{-1}, Hartmann Analytic). Analysis of the reaction products was done using thin layer chromatography (PEI-Cellulose F plates, Merck) with 1 M $(NH_4)_2SO_4$/1.5 M KH_2PO_4 pH 3.8 as running buffer for the TLC plates. Assays were performed at 35 °C. The dried TLC plates were analysed by phosphor imaging (GE Healthcare).

Electrophoretic mobility shift assays. 0.2 µM of dsDNA or dsRNA (same sequences used for NTase assays) was incubated with indicated amount of purified protein for 30 min on ice. As reaction buffer 20 mM Tris pH 8.0 and 200 mM NaCl was used. Samples were separated by 1% agarose gel prepared with Gel-Red (Biotium) as suggested by the manufacturer. The gel images were analysed using ImageJ.

Reporter assays. HEK 293T cells stably expressing full-length murine STING (2×10^4 cells in each well of a 96-well plate) were transiently transfected with 25 ng IFN-β promoter reporter plasmid (pIFN-β-GLUC) in conjunction with 200 ng cGAS expression vectors using GeneJuice (Novagen) as indicated by the manufacturer. A codon-optimized version the diguanylate cyclase domain (83–248) of TM1788 (*Thermotoga maritima* MSB8) harbouring a point mutation (R158A) to enhance c-di-GMP production was cloned into pEFBOS to contain a carboxy-terminal haemagglutinin (HA) tag[42]. This construct (c-di-GMP-synthase) was used to induce c-di-GMP production within 293T cells upon transient overexpression, which served as positive control. 14 h post transfection luciferase activity was assessed.

THP-1 cells were stimulated with 200 ng of either 50-mer dsDNA (as in NTase assays) or tri-phosphate-RNA complexed with Lipofectamine 2000 (Life Technologies) according to the manufacturer's instructions. Supernatants were collected 18 h after stimulation and assayed for IP-10 production via ELISA. 90-mer DNA used is as described in ref. 19. CMA was purchased from Sigma Aldrich.

Immunoblotting. Cells were lysed in 1× Laemmli buffer and denatured at 95 °C for 5 min. Probes were separated by 10% SDS–PAGE and transferred onto nitrocellulose membranes. Blots were incubated with anti-cGAS (Sigma Aldrich), anti-phospho-IRF3 (Cell Signaling Technology) or anti-β-actin-IgG–horseradish peroxidase (HRP). Goat anti-rabbit-IgG–HRP was purchased from Santa Cruz Biotechnology.

33. Cavlar, T., Deimling, T., Ablasser, A., Hopfner, K. P. & Hornung, V. Species-specific detection of the antiviral small-molecule compound CMA by STING. *EMBO J.* **32,** 1440–1450 (2013).
34. Kabsch, W. XDS. *Acta Crystallogr.* **66,** 125–132 (2010).
35. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364,** 215–230 (2007).
36. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr.* **62,** 1002–1011 (2006).
37. Morris, R. J., Perrakis, A. & Lamzin, V. S. ARP/wARP's model-building algorithms. I. The main chain. *Acta Crystallogr.* **58,** 968–975 (2002).
38. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr.* **60,** 2126–2132 (2004).
39. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.* **66,** 213–221 (2010).
40. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
41. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr.* **60,** 2210–2221 (2004).
42. Rao, F. *et al.* Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase. *Anal. Biochem.* **389,** 138–142 (2009).

2.5. Crystal and solution structure of RIG-I SF2 domain

2.5.1. Summary

RIG-I activation by RNA binding involves dramatic conformational changes. In this publication we could support the current model of the conformational flexibility of RIG-I by solving the human RIG-I SF2 domain in an unusual elongated conformation. Additional SAXS studies revealed that this conformation was induced to a large degree by crystal packing. However, this RIG-I SF2 structure expands the conformational spectrum needed for sensing RNA and forming the active state.

2.5.2. Contribution

The author of this thesis contributed to this publication by purifying and crystallizing the protein. He collected SEC-RALS-, SAXS- and crystal diffraction data. Additionally, he analyzed the SAXS data and participated in solving and refinement of the structure, and contributed to writing of the manuscript.

2.5.3. Paper

The following paper was published 2014 in Acta Crystallographica F.

Crystal and solution structure of RIG-I SF2 domain

Tobias Deimling, Sheng Cui, Katja Lammens, Karl-Peter Hopfner & Gregor Witte

Acta Crystallogr F Struct Biol Commun. 2014 Aug 1; **70**(Pt 8): 1027 - 31

Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

Crystal and solution structure of the human RIG-I SF2 domain

Tobias Deimling,[a] Sheng Cui,[b] Katja Lammens,[a] Karl-Peter Hopfner[a]* and Gregor Witte[a]*

[a]Gene Center and Department of Biochemistry, Ludwig-Maximilians University Munich, Feodor-Lynen-Strasse 25, 81377 Munich, Germany, and [b]MOH Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences and Peking Union Medical College, 9 Dong Dan San Tiao, Beijing 100730, People's Republic of China

Correspondence e-mail:
hopfner@genzentrum.lmu.de,
witte@genzentrum.lmu.de

Received 16 April 2014
Accepted 27 May 2014

PDB reference: RIG-I helicase domain, 4on9

RIG-I is a pathogen-recognition receptor that recognizes viral 5′-triphosphates carrying double-stranded RNA. Upon binding to these microbe-associated molecular patterns (MAMPs), RIG-I forms oligomers and promotes downstream processes that result in type I interferon production and induction of an antiviral state. Here, the crystal structure of the human RIG-I superfamily 2 ATPase domain crystallized in an unusually elongated and open conformation is reported. The elongated structure is probably induced in part by crystal packing, but nevertheless indicates that the domain is intrinsically very flexible. This flexibility might allow substantial structural changes upon substrate binding and oligomerization.

1. Introduction

The innate immune system detects invading pathogens *via* microbe-associated molecular patterns (MAMPs). Proteins belonging to the group of pattern-recognition receptors (PRR) have evolved to specifically sense MAMPs such as viral RNA or bacterial cell-wall components. The formation of PRR–MAMP complexes triggers signalling cascades that lead to host defence mechanisms by activation of nuclear factor-κB and interferon regulatory factors. The resulting production of inflammatory cytokines and type I interferons induces an antiviral state and activates the adaptive immune response.

PRRs can be membrane-bound, such as the Toll-like receptors (TLRs), or cytoplasmic, with the RIG-I-like receptor family being one prominent example. The latter comprises the name-giving RIG-I (retinoic acid inducible gene-I), MDA-5 (melanoma differentiation-associated gene 5) and LGP-2 (Laboratory of Genetics and Physiology 2). These three proteins are involved in the sensing of a variety of viruses. RIG-I, for example, senses *Rabies virus*, *Hepatitis C virus*, influenza viruses and others (Goubau *et al.*, 2013; Hornung *et al.*, 2006; Kato *et al.*, 2006; Loo *et al.*, 2008; Saito *et al.*, 2008), while MDA-5 detects, for example, picornaviruses (Goubau *et al.*, 2013; Gitlin *et al.*, 2006). Structurally, RIG-I, MDA-5 and LGP-2 are closely related to DEAD-box helicases and form a family within the superfamiliy 2 RNA helicases/ATPases (Fairman-Williams *et al.*, 2010). Apart from the helicase (SF2) domain, RIG-I comprises two N-terminal CARDs (caspase activation and recruitment domains), which are essential for downstream signalling (Gack *et al.*, 2007; Meylan *et al.*, 2005; Saito *et al.*, 2007; Zeng *et al.*, 2010), and a C-terminal regulatory domain (RD) involved in substrate binding and specificity (Hornung *et al.*, 2006; Pichlmair *et al.*, 2006; Schlee *et al.*, 2009; Schmidt *et al.*, 2009).

Recently, several structures have been reported that shed light on the working mechanism of RIG-I. The first structural information was obtained for the RIG-I RD alone (Cui *et al.*, 2008; Takahasi *et al.*, 2008) and in complex with 5′-triphosphate and blunt-end dsRNA (Kowalinski *et al.*, 2011; Lu *et al.*, 2010; Wang *et al.*, 2010), supporting the model of the RD being responsible for the specificity of RIG-I by binding to 5′-triphosphate RNA (Vela *et al.*, 2012). Further studies reported the structures of the RIG-I SF2 domain alone, RNA complexes of RIG-I SF2 and RIG-I SF2-RD, and the structure of RIG-I CARD-SF2 (Civril *et al.*, 2011; Ferrage *et al.*, 2012; Jiang *et al.*, 2011; Kohlway *et al.*, 2013; Kowalinski *et al.*, 2011; Luo *et al.*, 2011, 2012).

doi:10.1107/S2053230X14012230

Table 1
Crystal parameters and data-collection and refinement statistics for RIG-I SF2.

Values in parentheses are for the highest resolution shell.

Crystal data	
Space group	$P2_12_12_1$
Molecules per asymmetric unit	2
Unit-cell parameters (Å, °)	a = 98.9, b = 113.0, c = 124.3, $\alpha = \beta = \gamma = 90$
Solvent content (%)	53.2
Matthews coefficient (Å^3 Da^{-1})	2.63
Data-collection statistics	
Diffraction source	SLS beamline X06SA
Wavelength (Å)	0.99987
Detector	PSI Pilatus 6M pixel
Diffraction protocol	Single wavelength
Data-processing statistics	
Resolution range (Å)	50–2.70
No. of observed reflections	140661
No. of unique reflections	38287 (6016)
Completeness (%)	99.0 (98.4)
Multiplicity	3.67 (3.59)
$\langle I/\sigma(I)\rangle$	20.1 (2.2)
$R_{\rm meas}$ (%)	5.2 (77.4)
Refinement	
Resolution (Å)	49.06–2.71
No. of reflections used	38264
$R_{\rm work}$ (%)	21.9 (31.5)
$R_{\rm free}$ (%)	25.8 (36.8)
No. of atoms	
Total	8183
Protein	8126
Ligand	21
Water	36
Wilson B factor	68.53
Average B factors (Å^2)	
Protein	71.03
Ligand	67.69
Water	68.52
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.907
Ramachandran plot analysis	
Favoured (%)	96.0
Allowed (%)	4.0
Disallowed (%)	0
PDB code	4on9

The available structures allow the formulation of a mechanism for RIG-I activation. In the inactive state the second CARD domain contacts the insertion domain of the RIG-I SF2 (Hel2i), leading to an auto-inhibited state. RNA ligands bind to the RIG-I RD and SF2 domain and compete with CARDs at the helical insertion domain (Hel2i). In the resulting clamp-like structure of RIG-I SF2 around the dsRNA substrate, ATP binding leads to further compaction. This clamp formation results in the release of the two CARD domains from the auto-inhibited state (Ferrage *et al.*, 2012; Jiang *et al.*, 2011; Kohlway *et al.*, 2013; Kowalinski *et al.*, 2011; Leung & Amarasinghe, 2012; Luo *et al.*, 2011, 2012). The exposed CARDs interact with ubiquitin and CARDs of a downstream adaptor (Peisley *et al.*, 2014).

Here, we report the crystal structure of the human RIG-I SF2 in an unusual elongated conformation. Additional SAXS (small-angle X-ray scattering) studies support the conformational flexibility of this domain, which could play an important role in the formation of higher order structures or in initial substrate recognition.

2. Materials and methods

2.1. Cloning, expression and purification of Hs-RIG-I HD

Human RIG-I SF2 (hsRIG-I SF2; comprising residues 230–793 of the full-length protein) was cloned into pETDuet vector (Novagen) *via Eco*RI and *Not*I restriction sites. Transformed *Escherichia coli* Rosetta (DE3) cells were grown to an OD$_{600}$ of 0.8 before induction with IPTG (0.2 m*M* final concentration). Expression was carried out overnight at 18°C. The harvested cells were resuspended in lysis buffer (50 m*M* Tris pH 7.3, 500 m*M* NaCl, 10 m*M* imidazole, 4 m*M* DTT) and lysed by sonication. The N-terminally His$_6$-tagged hsRIG-I SF2 was first purified by Ni-affinity chromatography and the hsRIG-I SF2-containing elution fractions were dialyzed against 20 m*M* Tris pH 7.3, 100 m*M* NaCl, 4 m*M* DTT. In a second step this protein solution was applied onto a combination of anion- and cation-exchange resins (HiTrap Q/SP FF, GE Healthcare) and the flow-through was collected as the construct does not bind to these columns under the buffer conditions used. As a final purification step, size-exclusion chromatography was performed (Superdex S200 prep grade, GE Healthcare, running buffer 20 m*M* Tris pH 7.3, 100 m*M* NaCl, 4 m*M* DTT). Purified hsRIG-I SF2 could be concentrated to 60 mg ml^{-1} and aliquots were flash-frozen in liquid nitrogen.

2.2. Crystallization and data collection

HsRIG-I SF2 crystals were grown in hanging-drop vapour-diffusion geometry at 15°C. Crystals suitable for data collection were obtained by mixing 1 µl hsRIG-I SF2 solution with 1 µl reservoir solution [100 m*M* HEPES pH 6.8, 100 m*M* ammonium sulfate, 6.7%(*v*/*v*) 2-propanol, 9%(*v*/*v*) PEG 4000] with a total reservoir volume of 250 µl in the well. Crystals appeared after 4 d. For cryoprotection, the crystals were soaked in reservoir solution containing 25% ethylene glycol and flash-cooled in liquid nitrogen. Diffraction data were collected to 2.7 Å resolution on the X06SA beamline of the Swiss Light Source, Villigen, Switzerland. Diffraction data were indexed and integrated using *XDS* and scaled with *XSCALE* (Kabsch, 2010*a*,*b*). Statistics of the collected data set are shown in Table 1.

2.3. Structure solution and refinement

The structure of hsRIG-I SF2 was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007) within the *CCP*4 program suite (Winn *et al.*, 2011). Calculation of the Matthews coefficient with the space-group parameters (Table 1) and the molecular weight of hsRIG-I SF2 suggested two molecules per asymmetric unit, with a solvent content of 53.2% (Matthews, 1968). Molecular replacement was performed using an approach with the three isolated domains of the mouse RIG-I SF2 (PDB entry 3tbk; Civril *et al.*, 2011) as search models. For refinement of the initial maps a randomly chosen 5% of the reflections were used as a test set for cross-validation (Brünger, 1992). Refinement was performed in cyclic rounds of manual model building using *Coot* (Emsley *et al.*, 2010) and refinement using *PHENIX* (Afonine *et al.*, 2012). As a last step, water molecules were added automatically using *PHENIX* and subsequently manually checked and refined again. The statistics of the refined model are shown in Table 1.

2.4. Validation and deposition

The stereochemistry of the hsRIG-I SF2 domain was analyzed by *PROCHECK* (Laskowski *et al.*, 1996). All residues lie within the favoured and allowed regions of the Ramachandran plot (see Table 1). The atomic coordinates and structure factors of hsRIG-I SF2 have been deposited in the Protein Data Bank as entry 4on9.

2.5. Small-angle X-ray scattering (SAXS)

Small-angle X-ray scattering experiments were conducted at the X33 beamline at the European Molecular Biology Laboratory/

Deutsches Elektronen-Synchrotron, Hamburg, Germany. Samples were measured using a wavelength of 1.5 Å and with a q range from 0.06 to 6 nm (detector distance d = 2.7 m). Scattering data were recorded with a Pilatus 1M detector. Prior to measurement, samples were purified by size-exclusion chromatography and additionally centrifuged. All samples were monodisperse as could be judged from size-exclusion chromatography and dynamic light-scattering (Malvern-Viscotek 802 DLS) measurements. Scattering of the running buffer of this size-exclusion column (20 mM Tris pH 7.3, 100 mM NaCl, 4 mM DTT) was used for buffer correction of the protein sample measurements. Scattering data were recorded for hsRIG-I SF2 samples with concentrations of 2.5, 5 and 10 mg ml^{-1} (determined spectrophotometrically using $A_{280\,\mathrm{nm}}$ and the molar extinction coefficients). The samples did not show signs of radiation

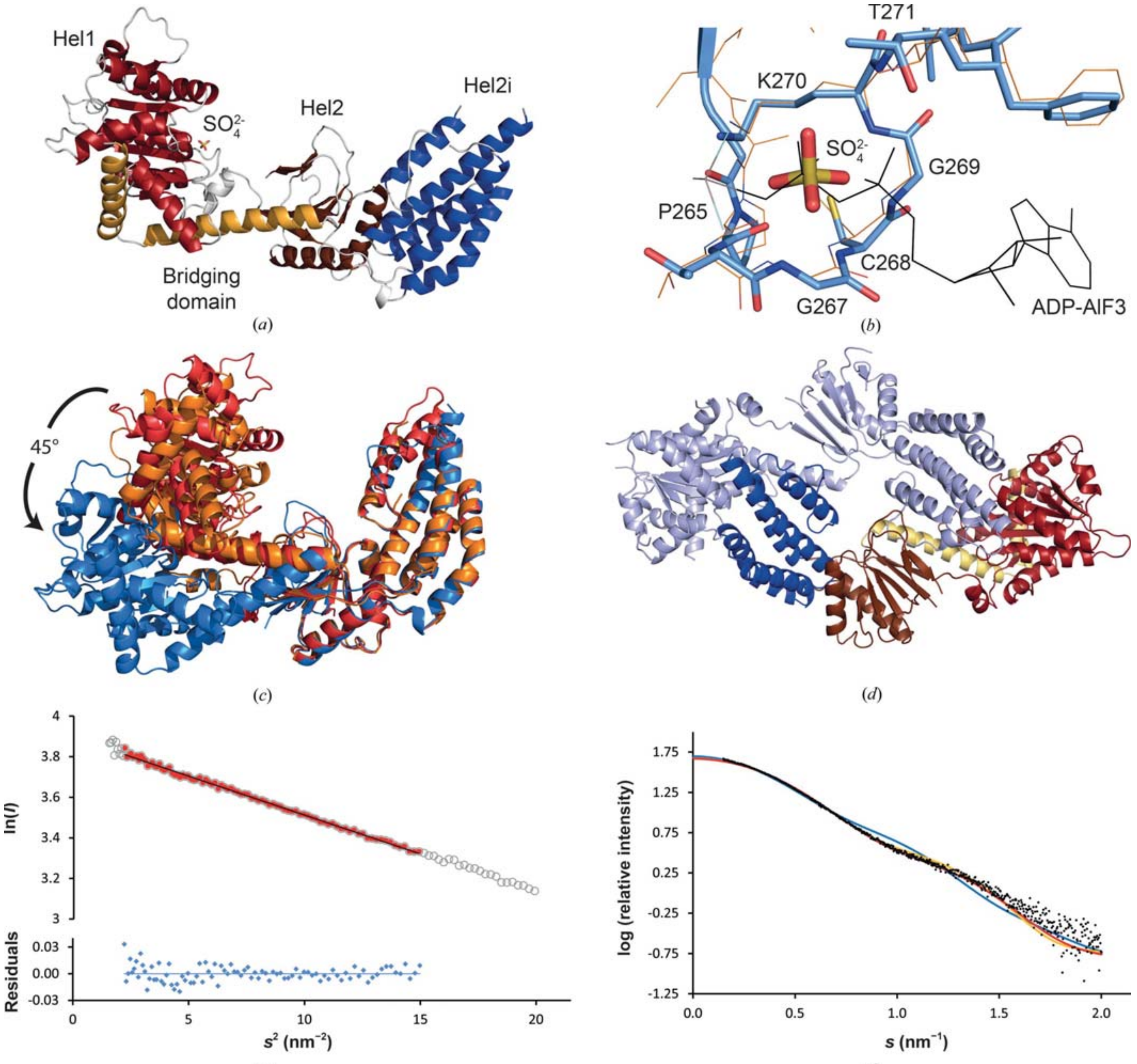


Figure 1

(*a*) Crystal structure of human RIG-I SF2. Subdomains are represented in different colours: Hel1, red; Hel2, brown; Hel2i, blue; bridging domain, orange. (*b*) Superposition of the ATP-binding site of the RIG-SF2 ADP-AlF$_3$-bound structure (PDB entry 4a36, shown as lines; Kowalinski *et al.*, 2011) and the corresponding region of the structure reported in this study (shown as sticks). The well coordinated sulfate ion occupies the space of the β-phosphate of the nucleotide-bound state. (*c*) Overlay of human (blue, this study), mouse (red, PDB entry 3tbk; Civril *et al.*, 2011) and duck (orange, PDB entry 4a2p; Kowalinski *et al.*, 2011) RIG-I SF2 crystal structures. (*d*) Cartoon representation of the crystallographic dimer with one monomer shown as in (*a*) and the second one in light blue. (*e*) Guinier plot [ln $I(s)$ *versus* s^2] of hsRIG-I SF2 SAXS data with linear regression (for points shown in red) used for determination of radius of gyration ($R_g = 3.37 \pm 0.01$ nm). The concentration of the sample was 5 mg ml^{-1}. The lower plot shows the residuals of the linear regression, indicating that the sample does not show aggregation or repulsive effects. (*f*) Comparison of measured hsRIG-I SF2 SAXS data (5 mg ml^{-1}) with theoretical scattering curves of RIG-I SF2 crystal structures calculated using *CRYSOL* (Svergun *et al.*, 1995). HsRIG-I SF2 (this study, PDB entry 4on9, blue curve), χ^2 = 15.0; mouse (PDB entry 3tbk, red curve), χ^2 = 2.6; duck (PDB entry 4a2p, orange curve), χ^2 = 2.9.

damage, which was assessed by automatic and manual comparison of eight consecutive 15 s exposure frames. Data were processed using *PRIMUS* from the *ATSAS* package (Konarev *et al.*, 2006). The protein molecular weight in solution was determined from the zero-angle scattering intensity $I(0)$ obtained from extrapolation of the Guinier plot [ln $I(s)$ *versus* s^2] region obeying the Guinier approximation for globular proteins ($s \times R_g < 1.3$) and from the Porod volume. BSA (66 kDa) was used as an $I(0)$ reference. Theoretical scattering curves were calculated using *CRYSOL* (Svergun *et al.*, 1995).

2.6. Size-exclusion chromatography coupled right-angle laser light scattering (SEC–RALS)

SEC–RALS data were measured using a Malvern/Viscotek TD270 right-angle laser light scattering detector and a VE3580 refractive-index detector connected to an analytical ÄKTAmicro chromatography system equipped with a 24 ml 10/300 GL Superdex S200 column (GE Healthcare). A BSA standard (66 kDa) was used to calibrate the system. For data evaluation the program package equipped with the instrument was used (*OmniSEC*, Malvern/Viskotek). The RIG-SF2 construct (c = 3 mg ml^{-1}) eluted as expected as a sharp single peak without any signs of aggregation in the SEC run.

3. Results and discussion

Human RIG-I SF2 crystallizes in the primitive orthorhombic space group $P2_12_12_1$ with two molecules per asymmetric unit. Traceable electron density could be found for the core hsRIG-I SF2 residues Phe241–Asp792 in chain *A* and Lys242–Ser793 in chain *B* and 36 solvent molecules interpreted as water, four sulfate ions and one chloride ion. The resolution of the structure is 2.7 Å and the model could be refined to an $R_{\rm work}$ and $R_{\rm free}$ of 0.219 and 0.258, respectively. As described previously (Civril *et al.*, 2011; Jiang *et al.*, 2011; Kowalinski *et al.*, 2011; Luo *et al.*, 2011), the structure comprises the two RecA-like domains Hel1 (242–444) and Hel2 (455–466 and 609–742) as well as the insertion domain Hel2i (271–603) and the bridging domain (745–793) (Fig. 1*a*). One sulfate ion coordinated by the P-loop (266–271) of the Hel1 domain occupies the space of the β-phosphate of ADP (Fig. 1*b* and Supplementary Fig. S1[1]).

In contrast to the C-shaped conformation of the SF2 domain observed previously, hsRIG-I SF2 crystallizes in an unusual elongated conformation in which Hel1 and the bridging domain are bent away from the rigid Hel2i-Hel2 module with an angular span of about 45° (Fig. 1*c*). The α-helical linker (445–454) and the short N-terminal stretch at the beginning of the bridging domain (743–745) serve as hinges. Conformational flexibility of RIG-I SF2 crystal structures has been observed before (Kowalinski *et al.*, 2011), but the structure described here is substantially more elongated than the previously described structures.

Two RIG-I SF2 domains build up a crystallographic dimer (Fig. 1*d*), with the Hel2i domain contacting the Hel1 domain of the neighbouring molecule and *vice versa*. The molecular weight determined by SEC–RALS ($M_{\rm w}^{\rm SEC-RALS}$ = 64 kDa; Supplementary Fig. S2) and SAXS data indicate that hsRIG-SF2 is monomeric in solution ($M_{\rm w}^{\rm theoretical}$ = 66.0 kDa), but it is possible that in RNA-activated RIG-I oligomers some contacts are mediated by the SF2 domains in addition to the CARD–CARD interactions.

To address solution conformations of the hsRIG-I SF2, we measured SAXS of the hsRIG-I SF2 construct. We could not observe signs of aggregation as judged from the linear Guinier plot (Fig. 1*e*). The molecular weight derived from SAXS data in the concentration range c = 2.5–10 mg ml^{-1} could only be explained by a monomeric hsRIG-I SF2 species in solution; even the highest concentration did not show signs of higher oligomers [$M_{\rm w}^{I(0);\ c=10\ {\rm mg\ ml}^{-1}}$ = 60.1 kDa; $M_{\rm w}^{{\rm Porod};\ c=10\ {\rm mg\ ml}^{-1}}$ = 55.5 kDa]. We then compared the experimental scattering data with the theoretical scattering curve of our crystallographic model. The obvious differences between the theoretical scattering curve of our model and the solution scattering curve (Fig. 1*f*) suggest that the crystallographically observed conformation is influenced by crystal packing (Supplementary Fig. S3) and does not represent a major conformation in solution. Notably, the solution scattering of the human RIG-I SF2 construct is very similar to the theoretical scattering curves of mouse and duck RIG-I SF2 crystal structures (Civril *et al.*, 2011).

In summary, the data support the current model that RIG-I SF2 possesses a high degree of conformational flexibility necessary for the formation of the different observed clamp conformations upon RNA and ATP binding and ATP-hydrolysis state, respectively (Kohlway *et al.*, 2013; Luo *et al.*, 2011, 2012). Therefore, the structure described here, even though not populated to a major degree in solution, expands the conformational spectrum of RIG-I SF2 that is likely to be needed to recognize specific MAMPs and form oligomeric structures for type I interferon activation.

We thank the staff of beamlines X06SA at the Swiss Light Source and X33 at EMBL/DESY Hamburg for excellent on-site support and beamtime allocation, and Filiz Civril and members of the Hopfner group for valuable discussions. This work was supported by DFG grant GRK1721 to GW and K-PH, DFG 3717/2-1 to GW and the National Institutes of Health grant U19AI083025 to K-PH. TD is supported by GRK1721.

References

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Civril, F., Bennett, M., Moldt, M., Deimling, T., Witte, G., Schiesser, S., Carell, T. & Hopfner, K. P. (2011). *EMBO Rep.* **12**, 1127–1134.

Cui, S., Eisenächer, K., Kirchhofer, A., Brzózka, K., Lammens, A., Lammens, K., Fujita, T., Conzelmann, K. K., Krug, A. & Hopfner, K. P. (2008). *Mol. Cell*, **29**, 169–179.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Fairman-Williams, M. E., Guenther, U. P. & Jankowsky, E. (2010). *Curr. Opin. Struct. Biol.* **20**, 313–324.

Ferrage, F., Dutta, K., Nistal-Villán, E., Patel, J. R., Sánchez-Aparicio, M. T., De Ioannes, P., Buku, A., Aseguinolaza, G. G., García-Sastre, A. & Aggarwal, A. K. (2012). *Structure*, **20**, 2048–2061.

Gack, M. U., Shin, Y. C., Joo, C.-H., Urano, T., Liang, C., Sun, L., Takeuchi, O., Akira, S., Chen, Z., Inoue, S. & Jung, J. U. (2007). *Nature (London)*, **446**, 916–920.

Gitlin, L., Barchet, W., Gilfillan, S., Cella, M., Beutler, B., Flavell, R. A., Diamond, M. S. & Colonna, M. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 8459–8464.

Goubau, D., Deddouche, S. & Reis, E. S. C. (2013). *Immunity*, **38**, 855–869.

Hornung, V., Ellegast, J., Kim, S., Brzózka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K. K., Schlee, M., Endres, S. & Hartmann, G. (2006). *Science*, **314**, 994–997.

Jiang, F., Ramanathan, A., Miller, M. T., Tang, G.-Q., Gale Jr, M., Patel, S. S. & Marcotrigiano, J. (2011). *Nature (London)*, **479**, 423–427.

Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.

Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.

Kato, H. *et al.* (2006). *Nature (London)*, **441**, 101–105.

---

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: TB5069).

Kohlway, A., Luo, D., Rawling, D. C., Ding, S. C. & Pyle, A. M. (2013). *EMBO Rep.* **14**, 772–779.

Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. (2006). *J. Appl. Cryst.* **39**, 277–286.

Kowalinski, E., Lunardi, T., McCarthy, A. A., Louber, J., Brunel, J., Grigorov, B., Gerlier, D. & Cusack, S. (2011). *Cell*, **147**, 423–435.

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). *J. Biomol. NMR*, **8**, 477–486.

Leung, D. W. & Amarasinghe, G. K. (2012). *Curr. Opin. Struct. Biol.* **22**, 297–303.

Loo, Y.-M., Fornek, J., Crochet, N., Bajwa, G., Perwitasari, O., Martinez-Sobrido, L., Akira, S., Gill, M. A., García-Sastre, A., Katze, M. G. & Gale Jr, M. (2008). *J. Virol.* **82**, 335–345.

Lu, C., Xu, H., Ranjith-Kumar, C. T., Brooks, M. T., Hou, T. Y., Hu, F., Herr, A. B., Strong, R. K., Kao, C. C. & Li, P. (2010). *Structure*, **18**, 1032–1043.

Luo, D., Ding, S. C., Vela, A., Kohlway, A., Lindenbach, B. D. & Pyle, A. M. (2011). *Cell*, **147**, 409–422.

Luo, D., Kohlway, A., Vela, A. & Pyle, A. M. (2012). *Structure*, **20**, 1983–1988.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Meylan, E., Curran, J., Hofmann, K., Moradpour, D., Binder, M., Bartenschlager, R. & Tschopp, J. (2005). *Nature (London)*, **437**, 1167–1172.

Peisley, A., Wu, B., Xu, H., Chen, Z. J. & Hur, S. (2014). *Nature* (London), **509**, 110–114.

Pichlmair, A., Schulz, O., Tan, C. P., Näslund, T. I., Liljeström, P., Weber, F. & Reis e Sousa, C. (2006). *Science*, **314**, 997–1001.

Saito, T., Hirai, R., Loo, Y.-M., Owen, D., Johnson, C. L., Sinha, S. C., Akira, S., Fujita, T. & Gale Jr, M. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 582–587.

Saito, T., Owen, D. M., Jiang, F., Marcotrigiano, J. & Gale Jr, M. (2008). *Nature (London)*, **454**, 523–527.

Schlee, M. *et al.* (2009). *Immunity*, **31**, 25–34.

Schmidt, A., Schwerd, T., Hamm, W., Hellmuth, J. C., Cui, S., Wenzel, M., Hoffmann, F. S., Michallet, M. C., Besch, R., Hopfner, K. P., Endres, S. & Rothenfusser, S. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 12067–12072.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.

Takahasi, K., Yoneyama, M., Nishihori, T., Hirai, R., Kumeta, H., Narita, R., Gale Jr, M., Inagaki, F. & Fujita, T. (2008). *Mol. Cell*, **29**, 428–440.

Vela, A., Fedorova, O., Ding, S. C. & Pyle, A. M. (2012). *J. Biol. Chem.* **287**, 42564–42573.

Wang, Y., Ludwig, J., Schuberth, C., Goldeck, M., Schlee, M., Li, H., Juranek, S., Sheng, G., Micura, R., Tuschl, T., Hartmann, G. & Patel, D. J. (2010). *Nature Struct. Mol. Biol.* **17**, 781–787.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Zeng, W., Sun, L., Jiang, X., Chen, X., Hou, F., Adhikari, A., Xu, M. & Chen, Z. J. (2010). *Cell*, **141**, 315–330.

3. References

1. Janeway, C.A., et al., *Immunobiology 5th edition: The immune System in Health and Disease*. 5th ed 2001, New York: Garland Science.
2. Murphy, K., *Janeway's Immunobiology*. 8th ed 2012, London and New York: Garland Science.
3. Medzhitov, R. and C.A. Janeway, Jr., *Innate immunity: the virtues of a nonclonal system of recognition.* Cell, 1997. **91**(3): p. 295-8.
4. Janeway, C.A., Jr. and R. Medzhitov, *Innate immune recognition.* Annu Rev Immunol, 2002. **20**: p. 197-216.
5. Poltorak, A., et al., *Physical contact between lipopolysaccharide and toll-like receptor 4 revealed by genetic complementation.* Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2163-7.
6. Lien, E. and D. Zipris, *The role of Toll-like receptor pathways in the mechanism of type 1 diabetes.* Curr Mol Med, 2009. **9**(1): p. 52-68.
7. Hugot, J.P., et al., *Mapping of a susceptibility locus for Crohn's disease on chromosome 16.* Nature, 1996. **379**(6568): p. 821-3.
8. Ogura, Y., et al., *A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease.* Nature, 2001. **411**(6837): p. 603-6.
9. Hugot, J.P., et al., *Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease.* Nature, 2001. **411**(6837): p. 599-603.
10. Wang, H., et al., *HMG-1 as a late mediator of endotoxin lethality in mice.* Science, 1999. **285**(5425): p. 248-51.
11. Scaffidi, P., T. Misteli, and M.E. Bianchi, *Release of chromatin protein HMGB1 by necrotic cells triggers inflammation.* Nature, 2002. **418**(6894): p. 191-5.
12. Hori, O., et al., *The receptor for advanced glycation end products (RAGE) is a cellular binding site for amphoterin. Mediation of neurite outgrowth and co-expression of rage and amphoterin in the developing nervous system.* J Biol Chem, 1995. **270**(43): p. 25752-61.
13. Anderson, K.V., L. Bokla, and C. Nusslein-Volhard, *Establishment of dorsal-ventral polarity in the Drosophila embryo: the induction of polarity by the Toll gene product.* Cell, 1985. **42**(3): p. 791-8.
14. Anderson, K.V., G. Jurgens, and C. Nusslein-Volhard, *Establishment of dorsal-ventral polarity in the Drosophila embryo: genetic studies on the role of the Toll gene product.* Cell, 1985. **42**(3): p. 779-89.
15. Newton, K. and V.M. Dixit, *Signaling in innate immunity and inflammation.* Cold Spring Harb Perspect Biol, 2012. **4**(3).
16. Kumar, H., T. Kawai, and S. Akira, *Pathogen recognition by the innate immune system.* Int Rev Immunol, 2011. **30**(1): p. 16-34.
17. Kawai, T. and S. Akira, *The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors.* Nat Immunol, 2010. **11**(5): p. 373-84.
18. Antosz, H. and M. Osiak, *NOD1 and NOD2 receptors: integral members of the innate and adaptive immunity system.* Acta Biochim Pol, 2013. **60**(3): p. 351-60.
19. Carneiro, L.A., et al., *Nod-like proteins in inflammation and disease.* J Pathol, 2008. **214**(2): p. 136-48.
20. Benko, S., D.J. Philpott, and S.E. Girardin, *The microbial and danger signals that activate Nod-like receptors.* Cytokine, 2008. **43**(3): p. 368-73.
21. Isaacs, A., R.A. Cox, and Z. Rotem, *Foreign nucleic acids as the stimulus to make interferon.* Lancet, 1963. **2**(7299): p. 113-6.
22. Rotem, Z., R.A. Cox, and A. Isaacs, *Inhibition of virus multiplication by foreign nucleic acid.* Nature, 1963. **197**: p. 564-6.
23. Alexopoulou, L., et al., *Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3.* Nature, 2001. **413**(6857): p. 732-8.

24. Kumagai, Y. and S. Akira, *Identification and functions of pattern-recognition receptors.* J Allergy Clin Immunol, 2010. **125**(5): p. 985-92.
25. Diebold, S.S., et al., *Viral infection switches non-plasmacytoid dendritic cells into high interferon producers.* Nature, 2003. **424**(6946): p. 324-8.
26. Wu, J. and Z.J. Chen, *Innate immune sensing and signaling of cytosolic nucleic acids.* Annu Rev Immunol, 2014. **32**: p. 461-88.
27. Sun, Y.W., *RIG-I, a human homolog gene of RNA helicase, is induced by retinoic acid during the differentiation of acute promyelocytic leukemia cell*, in *Shanghai Institute of Hematology, Rui-Jin Hospital, Shanghai Second Medical University,* 1997.
28. Kang, D.C., et al., *mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties.* Proc Natl Acad Sci U S A, 2002. **99**(2): p. 637-42.
29. 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. 730-7.
30. Andrejeva, J., et al., *The V proteins of paramyxoviruses bind the IFN-inducible RNA helicase, mda-5, and inhibit its activation of the IFN-beta promoter.* Proc Natl Acad Sci U S A, 2004. **101**(49): p. 17264-9.
31. Goubau, D., S. Deddouche, and E.S.C. Reis, *Cytosolic sensing of viruses.* Immunity, 2013. **38**(5): p. 855-69.
32. Takeuchi, O. and S. Akira, *Pattern recognition receptors and inflammation.* Cell, 2010. **140**(6): p. 805-20.
33. Yoo, J.S., H. Kato, and T. Fujita, *Sensing viral invasion by RIG-I like receptors.* Curr Opin Microbiol, 2014. **20C**: p. 131-138.
34. Yoneyama, M., et al., *Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity.* J Immunol, 2005. **175**(5): p. 2851-8.
35. Satoh, T., et al., *LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses.* Proc Natl Acad Sci U S A, 2010. **107**(4): p. 1512-7.
36. Venkataraman, T., et al., *Loss of DExD/H box RNA helicase LGP2 manifests disparate antiviral responses.* J Immunol, 2007. **178**(10): p. 6444-55.
37. Kato, H., et al., *Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5.* J Exp Med, 2008. **205**(7): p. 1601-10.
38. Baum, A., R. Sachidanandam, and A. Garcia-Sastre, *Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing.* Proc Natl Acad Sci U S A, 2010. **107**(37): p. 16303-8.
39. Hornung, V., et al., *5'-Triphosphate RNA is the ligand for RIG-I.* Science, 2006. **314**(5801): p. 994-7.
40. Pichlmair, A., et al., *RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates.* Science, 2006. **314**(5801): p. 997-1001.
41. 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**(1): p. 25-34.
42. Schmidt, A., et al., *5'-triphosphate RNA requires base-paired structures to activate antiviral signaling via RIG-I.* Proc Natl Acad Sci U S A, 2009. **106**(29): p. 12067-72.
43. Panda, D., et al., *Induction of interferon and interferon signaling pathways by replication of defective interfering particle RNA in cells constitutively expressing vesicular stomatitis virus replication proteins.* J Virol, 2010. **84**(9): p. 4826-31.
44. 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. 336-46.
45. Runge, S., et al., *In vivo ligands of MDA5 and RIG-I in measles virus-infected cells.* PLoS Pathog, 2014. **10**(4): p. e1004081.

46. Malathi, K., et al., *Small self-RNA generated by RNase L amplifies antiviral innate immunity.* Nature, 2007. **448**(7155): p. 816-9.
47. Malathi, K., et al., *RNase L releases a small RNA from HCV RNA that refolds into a potent PAMP.* RNA, 2010. **16**(11): p. 2108-19.
48. Pichlmair, A., et al., *Activation of MDA5 requires higher-order RNA structures generated during virus infection.* J Virol, 2009. **83**(20): p. 10761-9.
49. Myong, S., et al., *Cytosolic viral sensor RIG-I is a 5'-triphosphate-dependent translocase on double-stranded RNA.* Science, 2009. **323**(5917): p. 1070-4.
50. Patel, J.R., et al., *ATPase-driven oligomerization of RIG-I on RNA allows optimal activation of type-I interferon.* EMBO Rep, 2013. **14**(9): p. 780-7.
51. 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.
52. Jiang, X., et al., *Ubiquitin-induced oligomerization of the RNA sensors RIG-I and MDA5 activates antiviral innate immune response.* Immunity, 2012. **36**(6): p. 959-73.
53. Oshiumi, H., et al., *A distinct role of Riplet-mediated K63-Linked polyubiquitination of the RIG-I repressor domain in human antiviral innate immune responses.* PLoS Pathog, 2013. **9**(8): p. e1003533.
54. 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. 315-30.
55. Peisley, A., et al., *Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I.* Nature, 2014.
56. Wu, B., et al., *Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5.* Cell, 2013. **152**(1-2): p. 276-89.
57. Peisley, A., et al., *Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition.* Proc Natl Acad Sci U S A, 2011. **108**(52): p. 21010-5.
58. Peisley, A., et al., *Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments.* Proc Natl Acad Sci U S A, 2012. **109**(49): p. E3340-9.
59. Berke, I.C., et al., *MDA5 assembles into a polar helical filament on dsRNA.* Proc Natl Acad Sci U S A, 2012. **109**(45): p. 18437-41.
60. Berke, I.C. and Y. Modis, *MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA.* EMBO J, 2012. **31**(7): p. 1714-26.
61. Xu, L.G., et al., *VISA is an adapter protein required for virus-triggered IFN-beta signaling.* Mol Cell, 2005. **19**(6): p. 727-40.
62. Seth, R.B., et al., *Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3.* Cell, 2005. **122**(5): p. 669-82.
63. 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. 1167-72.
64. Kawai, T., et al., *IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction.* Nat Immunol, 2005. **6**(10): p. 981-8.
65. Sun, Q., et al., *The specific and essential role of MAVS in antiviral innate immune responses.* Immunity, 2006. **24**(5): p. 633-42.
66. Kumar, H., et al., *Essential role of IPS-1 in innate immune responses against RNA viruses.* J Exp Med, 2006. **203**(7): p. 1795-803.
67. Hou, F., et al., *MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response.* Cell, 2011. **146**(3): p. 448-61.
68. Chiang, J.J., M.E. Davis, and M.U. Gack, *Regulation of RIG-I-like receptor signaling by host and viral proteins.* Cytokine Growth Factor Rev, 2014.
69. Cui, S., et al., *The C-terminal regulatory domain is the RNA 5'-triphosphate sensor of RIG-I.* Mol Cell, 2008. **29**(2): p. 169-79.
70. Takahasi, K., et al., *Nonself RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses.* Mol Cell, 2008. **29**(4): p. 428-40.

71. Lu, C., et al., *Crystal structure of RIG-I C-terminal domain bound to blunt-ended double-strand RNA without 5' triphosphate.* Nucleic Acids Res, 2011. **39**(4): p. 1565-75.
72. Lu, C., et al., *The structural basis of 5' triphosphate double-stranded RNA recognition by RIG-I C-terminal domain.* Structure, 2010. **18**(8): p. 1032-43.
73. 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. 781-7.
74. Caruthers, J.M. and D.B. McKay, *Helicase structure and mechanism.* Curr Opin Struct Biol, 2002. **12**(1): p. 123-33.
75. Deimling, T., et al., *Crystal and solution structure of the human RIG-I SF2 domain.* Acta Crystallogr F Struct Biol Commun, 2014. **70**(Pt 8): p. 1027-31.
76. Civril, F., et al., *The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling.* EMBO Rep, 2011. **12**(11): p. 1127-34.
77. Jiang, F., et al., *Structural basis of RNA recognition and activation by innate immune receptor RIG-I.* Nature, 2011. **479**(7373): p. 423-7.
78. Kowalinski, E., et al., *Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA.* Cell, 2011. **147**(2): p. 423-35.
79. Kohlway, A., et al., *Defining the functional determinants for RNA surveillance by RIG-I.* EMBO Rep, 2013. **14**(9): p. 772-9.
80. Luo, D., et al., *Structural insights into RNA recognition by RIG-I.* Cell, 2011. **147**(2): p. 409-22.
81. Luo, D., et al., *Visualizing the determinants of viral RNA recognition by innate immune sensor RIG-I.* Structure, 2012. **20**(11): p. 1983-8.
82. Leung, D.W. and G.K. Amarasinghe, *Structural insights into RNA recognition and activation of RIG-I-like receptors.* Curr Opin Struct Biol, 2012. **22**(3): p. 297-303.
83. Rawling, D.C. and A.M. Pyle, *Parts, assembly and operation of the RIG-I family of motors.* Curr Opin Struct Biol, 2014. **25**: p. 25-33.
84. Oshiumi, H., M. Matsumoto, and T. Seya, *Ubiquitin-mediated modulation of the cytoplasmic viral RNA sensor RIG-I.* J Biochem, 2012. **151**(1): p. 5-11.
85. Arimoto, K., et al., *Negative regulation of the RIG-I signaling by the ubiquitin ligase RNF125.* Proc Natl Acad Sci U S A, 2007. **104**(18): p. 7500-5.
86. Chen, W., et al., *Induction of Siglec-G by RNA viruses inhibits the innate immune response by promoting RIG-I degradation.* Cell, 2013. **152**(3): p. 467-78.
87. Wang, L., et al., *USP4 positively regulates RIG-I-mediated antiviral response through deubiquitination and stabilization of RIG-I.* J Virol, 2013. **87**(8): p. 4507-15.
88. Friedman, C.S., et al., *The tumour suppressor CYLD is a negative regulator of RIG-I-mediated antiviral response.* EMBO Rep, 2008. **9**(9): p. 930-6.
89. Nistal-Villan, E., et al., *Negative role of RIG-I serine 8 phosphorylation in the regulation of interferon-beta production.* J Biol Chem, 2010. **285**(26): p. 20252-61.
90. Gack, M.U., et al., *Phosphorylation-mediated negative regulation of RIG-I antiviral activity.* J Virol, 2010. **84**(7): p. 3220-9.
91. Wies, E., et al., *Dephosphorylation of the RNA sensors RIG-I and MDA5 by the phosphatase PP1 is essential for innate immune signaling.* Immunity, 2013. **38**(3): p. 437-49.
92. Hayakawa, S., et al., *ZAPS is a potent stimulator of signaling mediated by the RNA helicase RIG-I during antiviral responses.* Nat Immunol, 2011. **12**(1): p. 37-44.
93. Kok, K.H., et al., *The double-stranded RNA-binding protein PACT functions as a cellular activator of RIG-I to facilitate innate antiviral response.* Cell Host Microbe, 2011. **9**(4): p. 299-309.
94. Chen, H., et al., *RAVER1 is a coactivator of MDA5-mediated cellular antiviral response.* J Mol Cell Biol, 2013. **5**(2): p. 111-9.
95. Diao, F., et al., *Negative regulation of MDA5- but not RIG-I-mediated innate antiviral signaling by the dihydroxyacetone kinase.* Proc Natl Acad Sci U S A, 2007. **104**(28): p. 11706-11.

96. Habjan, M., et al., *Processing of genome 5' termini as a strategy of negative-strand RNA viruses to avoid RIG-I-dependent interferon induction.* PLoS One, 2008. **3**(4): p. e2032.

97. Schneider, U., et al., *Genome trimming by Borna disease viruses: viral replication control or escape from cellular surveillance?* Cell Mol Life Sci, 2007. **64**(9): p. 1038-42.

98. Wang, H., et al., *Old World hantaviruses do not produce detectable amounts of dsRNA in infected cells and the 5' termini of their genomic RNAs are monophosphorylated.* J Gen Virol, 2011. **92**(Pt 5): p. 1199-204.

99. Kinpe, D.M. and P.M. Howley, *Fields' Virology*. 5th ed 2007, Philadelphia: Lippcinott Wiliams & Wilkins.

100. Gonzalez, C.M., L. Wang, and B. Damania, *Kaposi's sarcoma-associated herpesvirus encodes a viral deubiquitinase.* J Virol, 2009. **83**(19): p. 10224-33.

101. Inn, K.S., et al., *Inhibition of RIG-I-mediated signaling by Kaposi's sarcoma-associated herpesvirus-encoded deubiquitinase ORF64.* J Virol, 2011. **85**(20): p. 10899-904.

102. Davis, M.E., et al., *Antagonism of the Phosphatase PP1 by the Measles Virus V Protein Is Required for Innate Immune Escape of MDA5.* Cell Host Microbe, 2014. **16**(1): p. 19-30.

103. Horvath, C.M., *Silencing STATs: lessons from paramyxovirus interferon evasion.* Cytokine Growth Factor Rev, 2004. **15**(2-3): p. 117-27.

104. Motz, C., et al., *Paramyxovirus V proteins disrupt the fold of the RNA sensor MDA5 to inhibit antiviral signaling.* Science, 2013. **339**(6120): p. 690-3.

105. Barral, P.M., et al., *RIG-I is cleaved during picornavirus infection.* Virology, 2009. **391**(2): p. 171-6.

106. Kuo, R.L., et al., *MDA5 plays a crucial role in enterovirus 71 RNA-mediated IRF3 activation.* PLoS One, 2013. **8**(5): p. e63431.

107. Feng, Q., et al., *Enterovirus 2Apro targets MDA5 and MAVS in infected cells.* J Virol, 2014. **88**(6): p. 3369-78.

108. Kovacsovics, M., et al., *Overexpression of Helicard, a CARD-containing helicase cleaved during apoptosis, accelerates DNA degradation.* Curr Biol, 2002. **12**(10): p. 838-43.

109. Barral, P.M., et al., *MDA-5 is cleaved in poliovirus-infected cells.* J Virol, 2007. **81**(8): p. 3677-84.

110. Freundt, E.C., et al., *Molecular determinants for subcellular localization of the severe acute respiratory syndrome coronavirus open reading frame 3b protein.* J Virol, 2009. **83**(13): p. 6631-40.

111. Kopecky-Bromberg, S.A., et al., *Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists.* J Virol, 2007. **81**(2): p. 548-57.

112. Lifland, A.W., et al., *Human respiratory syncytial virus nucleoprotein and inclusion bodies antagonize the innate immune response mediated by MDA5 and MAVS.* J Virol, 2012. **86**(15): p. 8245-58.

113. Roberts, W.K., et al., *Interferon-mediated protein kinase and low-molecular-weight inhibitor of protein synthesis.* Nature, 1976. **264**(5585): p. 477-80.

114. Williams, B.R., *Signal integration via PKR.* Sci STKE, 2001. **2001**(89): p. re2.

115. Yang, Y.L., et al., *Deficient signaling in mice devoid of double-stranded RNA-dependent protein kinase.* EMBO J, 1995. **14**(24): p. 6095-106.

116. Kuchta, K., et al., *Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human.* Nucleic Acids Res, 2009. **37**(22): p. 7701-14.

117. Chakrabarti, A., B.K. Jha, and R.H. Silverman, *New insights into the role of RNase L in innate immunity.* J Interferon Cytokine Res, 2011. **31**(1): p. 49-57.

118. 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. 537-40.

119. Kristiansen, H., et al., *The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities.* J Interferon Cytokine Res, 2011. **31**(1): p. 41-7.
120. Silverman, R.H., *Viral encounters with 2',5'-oligoadenylate synthetase and RNase L during the interferon antiviral response.* J Virol, 2007. **81**(23): p. 12720-9.
121. Hemmi, H., et al., *A Toll-like receptor recognizes bacterial DNA.* Nature, 2000. **408**(6813): p. 740-5.
122. Ishii, K.J., et al., *A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA.* Nat Immunol, 2006. **7**(1): p. 40-8.
123. Stetson, D.B. and R. Medzhitov, *Recognition of cytosolic DNA activates an IRF3-dependent innate immune response.* Immunity, 2006. **24**(1): p. 93-103.
124. Paludan, S.R. and A.G. Bowie, *Immune sensing of DNA.* Immunity, 2013. **38**(5): p. 870-80.
125. Jin, L., et al., *MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex class II and mediates transduction of apoptotic signals.* Mol Cell Biol, 2008. **28**(16): p. 5014-26.
126. Ishikawa, H. and G.N. Barber, *STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling.* Nature, 2008. **455**(7213): p. 674-8.
127. Sun, W., et al., *ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization.* Proc Natl Acad Sci U S A, 2009. **106**(21): p. 8653-8.
128. Zhong, B., et al., *The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation.* Immunity, 2008. **29**(4): p. 538-50.
129. Ishikawa, H., Z. Ma, and G.N. Barber, *STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity.* Nature, 2009. **461**(7265): p. 788-92.
130. Hayashi, T., et al., *MAM: more than just a housekeeper.* Trends Cell Biol, 2009. **19**(2): p. 81-8.
131. 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.* Infect Immun, 2011. **79**(2): p. 688-94.
132. Chen, H., et al., *Activation of STAT6 by STING is critical for antiviral innate immunity.* Cell, 2011. **147**(2): p. 436-46.
133. Huang, Y.H., et al., *The structural basis for the sensing and binding of cyclic di-GMP by STING.* Nat Struct Mol Biol, 2012. **19**(7): p. 728-30.
134. 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. 1073-86.
135. Shang, G., et al., *Crystal structures of STING protein reveal basis for recognition of cyclic di-GMP.* Nat Struct Mol Biol, 2012. **19**(7): p. 725-7.
136. Shu, C., et al., *Structure of STING bound to cyclic di-GMP reveals the mechanism of cyclic dinucleotide recognition by the immune system.* Nat Struct Mol Biol, 2012. **19**(7): p. 722-4.
137. Yin, Q., et al., *Cyclic di-GMP sensing via the innate immune signaling protein STING.* Mol Cell, 2012. **46**(6): p. 735-45.
138. Su, Y.C., et al., *Crystallization studies of the murine c-di-GMP sensor protein STING.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2012. **68**(Pt 8): p. 906-10.
139. Chin, K.H., et al., *Novel c-di-GMP recognition modes of the mouse innate immune adaptor protein STING.* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 3): p. 352-66.
140. Saitoh, T., et al., *Atg9a controls dsDNA-driven dynamic translocation of STING and the innate immune response.* Proc Natl Acad Sci U S A, 2009. **106**(49): p. 20842-6.
141. Tanaka, Y. and Z.J. Chen, *STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway.* Sci Signal, 2012. **5**(214): p. ra20.
142. Romling, U., M.Y. Galperin, and M. Gomelsky, *Cyclic di-GMP: the first 25 years of a universal bacterial second messenger.* Microbiol Mol Biol Rev, 2013. **77**(1): p. 1-52.
143. Witte, G., et al., *Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates.* Mol Cell, 2008. **30**(2): p. 167-78.

144. Davies, B.W., et al., *Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for V. cholerae virulence.* Cell, 2012. **149**(2): p. 358-70.

145. 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. 1703-5.

146. McWhirter, S.M., et al., *A host type I interferon response is induced by cytosolic sensing of the bacterial second messenger cyclic-di-GMP.* J Exp Med, 2009. **206**(9): p. 1899-911.

147. Monroe, K.M., S.M. McWhirter, and R.E. Vance, *Induction of type I interferons by bacteria.* Cell Microbiol, 2010. **12**(7): p. 881-90.

148. Jin, L., et al., *MPYS is required for IFN response factor 3 activation and type I IFN production in the response of cultured phagocytes to bacterial second messengers cyclic-di-AMP and cyclic-di-GMP.* J Immunol, 2011. **187**(5): p. 2595-601.

149. Burdette, D.L., et al., *STING is a direct innate immune sensor of cyclic di-GMP.* Nature, 2011. **478**(7370): p. 515-8.

150. 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. 748-62.

151. Schoggins, J.W., et al., *A diverse range of gene products are effectors of the type I interferon antiviral response.* Nature, 2011. **472**(7344): p. 481-5.

152. 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. 786-91.

153. Wu, J., et al., *Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA.* Science, 2013. **339**(6121): p. 826-30.

154. Ablasser, A., et al., *cGAS produces a 2'-5'-linked cyclic dinucleotide second messenger that activates STING.* Nature, 2013. **498**(7454): p. 380-4.

155. Diner, E.J., et al., *The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING.* Cell Rep, 2013. **3**(5): p. 1355-61.

156. 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. 1094-107.

157. Li, X.D., et al., *Pivotal roles of cGAS-cGAMP signaling in antiviral defense and immune adjuvant effects.* Science, 2013. **341**(6152): p. 1390-4.

158. Ablasser, A., et al., *Cell intrinsic immunity spreads to bystander cells via the intercellular transfer of cGAMP.* Nature, 2013. **503**(7477): p. 530-4.

159. Civril, F., et al., *Structural mechanism of cytosolic DNA sensing by cGAS.* Nature, 2013. **498**(7454): p. 332-7.

160. Kato, K., et al., *Structural and functional analyses of DNA-sensing and immune activation by human cGAS.* PLoS One, 2013. **8**(10): p. e76983.

161. Kranzusch, P.J., et al., *Structure of human cGAS reveals a conserved family of second-messenger enzymes in innate immunity.* Cell Rep, 2013. **3**(5): p. 1362-8.

162. Li, X., et al., *Cyclic GMP-AMP synthase is activated by double-stranded DNA-induced oligomerization.* Immunity, 2013. **39**(6): p. 1019-31.

163. 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**(3): p. 421-30.

164. Hornung, V., et al., *OAS proteins and cGAS: unifying concepts in sensing and responding to cytosolic nucleic acids.* Nat Rev Immunol, 2014. **14**(8): p. 521-8.

165. Shu, C., X. Li, and P. Li, *The mechanism of double-stranded DNA sensing through the cGAS-STING pathway.* Cytokine Growth Factor Rev, 2014.

166. Zhang, X., et al., *Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING.* Mol Cell, 2013. **51**(2): p. 226-35.

167. Jin, L., et al., *Identification and characterization of a loss-of-function human MPYS variant.* Genes Immun, 2011. **12**(4): p. 263-9.

168. Tsuchida, T., et al., *The ubiquitin ligase TRIM56 regulates innate immune responses to intracellular double-stranded DNA.* Immunity, 2010. **33**(5): p. 765-76.
169. Zhang, J., et al., *TRIM32 protein modulates type I interferon induction and cellular antiviral response by targeting MITA/STING protein for K63-linked ubiquitination.* J Biol Chem, 2012. **287**(34): p. 28646-55.
170. 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.
171. Konno, H., K. Konno, and G.N. Barber, *Cyclic dinucleotides trigger ULK1 (ATG1) phosphorylation of STING to prevent sustained innate immune signaling.* Cell, 2013. **155**(3): p. 688-98.
172. 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. 329-41.
173. Chen, H., et al., *An alternative splicing isoform of MITA antagonizes MITA-mediated induction of type I IFNs.* J Immunol, 2014. **192**(3): p. 1162-70.
174. Deretic, V., *Autophagy as an innate immunity paradigm: expanding the scope and repertoire of pattern recognition receptors.* Curr Opin Immunol, 2012. **24**(1): p. 21-31.
175. Deretic, V., T. Saitoh, and S. Akira, *Autophagy in infection, inflammation and immunity.* Nat Rev Immunol, 2013. **13**(10): p. 722-37.
176. 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 stimulator of IFN genes.* J Immunol, 2011. **187**(10): p. 5268-76.
177. McFarlane, S., et al., *Early induction of autophagy in human fibroblasts after infection with human cytomegalovirus or herpes simplex virus 1.* J Virol, 2011. **85**(9): p. 4212-21.
178. 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. 803-15.
179. Manzanillo, P.S., et al., *Mycobacterium tuberculosis activates the DNA-dependent cytosolic surveillance pathway within macrophages.* Cell Host Microbe, 2012. **11**(5): p. 469-80.
180. 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. 228-38.
181. 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. 691-5.
182. Aguirre, S., et al., *DENV inhibits type I IFN production in infected cells by cleaving human STING.* PLoS Pathog, 2012. **8**(10): p. e1002934.
183. Yu, C.Y., et al., *Dengue virus targets the adaptor protein MITA to subvert host innate immunity.* PLoS Pathog, 2012. **8**(6): p. e1002780.
184. 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.
185. Ding, Q., et al., *Hepatitis C virus NS4B blocks the interaction of STING and TBK1 to evade host innate immunity.* J Hepatol, 2013. **59**(1): p. 52-8.
186. 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.
187. Crow, Y.J., et al., *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus.* Nat Genet, 2006. **38**(8): p. 917-20.
188. Gall, A., et al., *Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease.* Immunity, 2012. **36**(1): p. 120-31.
189. Stetson, D.B., et al., *Trex1 prevents cell-intrinsic initiation of autoimmunity.* Cell, 2008. **134**(4): p. 587-98.

190. Ablasser, A., et al., *TREX1 Deficiency Triggers Cell-Autonomous Immunity in a cGAS-Dependent Manner.* J Immunol, 2014. **192**(12): p. 5993-7.

191. Ablasser, A., et al., *Nucleic acid driven sterile inflammation.* Clin Immunol, 2013. **147**(3): p. 207-15.

192. Nazmi, A., et al., *STING mediates neuronal innate immune response following Japanese encephalitis virus infection.* Sci Rep, 2012. **2**: p. 347.

193. 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. 369-81.

194. 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. 422-34.

195. 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. 1132-42.

196. Gao, D., et al., *Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses.* Science, 2013. **341**(6148): p. 903-6.

197. 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.* Proc Natl Acad Sci U S A, 2014. **111**(5): p. E611-7.

198. Horan, K.A., et al., *Proteasomal degradation of herpes simplex virus capsids in macrophages releases DNA to the cytosol for recognition by DNA sensors.* J Immunol, 2013. **190**(5): p. 2311-9.

199. Dai, P., et al., *Modified vaccinia virus Ankara triggers type I IFN production in murine conventional dendritic cells via a cGAS/STING-mediated cytosolic DNA-sensing pathway.* PLoS Pathog, 2014. **10**(4): p. e1003989.

200. Sunthamala, N., et al., *E2 proteins of high risk human papillomaviruses down-modulate STING and IFN-kappa transcription in keratinocytes.* PLoS One, 2014. **9**(3): p. e91473.

201. Lam, E., S. Stein, and E. Falck-Pedersen, *Adenovirus detection by the cGAS/STING/TBK1 DNA sensing cascade.* J Virol, 2014. **88**(2): p. 974-81.

202. Wang, X., et al., *Hepatitis B virus X protein suppresses virus-triggered IRF3 activation and IFN-beta induction by disrupting the VISA-associated complex.* Cell Mol Immunol, 2010. **7**(5): p. 341-8.

203. Archer, K.A., J. Durack, and D.A. Portnoy, *STING-dependent type I IFN production inhibits cell-mediated immunity to Listeria monocytogenes.* PLoS Pathog, 2014. **10**(1): p. e1003861.

204. Jin, L., et al., *STING/MPYS mediates host defense against Listeria monocytogenes infection by regulating Ly6C(hi) monocyte migration.* J Immunol, 2013. **190**(6): p. 2835-43.

205. Koppe, U., et al., *Streptococcus pneumoniae stimulates a STING- and IFN regulatory factor 3-dependent type I IFN production in macrophages, which regulates RANTES production in macrophages, cocultured alveolar epithelial cells, and mouse lungs.* J Immunol, 2012. **188**(2): p. 811-7.

206. Parker, D., et al., *Streptococcus pneumoniae DNA initiates type I interferon signaling in the respiratory tract.* MBio, 2011. **2**(3): p. e00016-11.

207. de Almeida, L.A., et al., *MyD88 and STING signaling pathways are required for IRF3-mediated IFN-beta induction in response to Brucella abortus infection.* PLoS One, 2011. **6**(8): p. e23135.

208. 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.

209. Prantner, D., T. Darville, and U.M. Nagarajan, *Stimulator of IFN gene is critical for induction of IFN-beta during Chlamydia muridarum infection.* J Immunol, 2010. **184**(5): p. 2551-60.

210. Jones, J.W., et al., *Absent in melanoma 2 is required for innate immune recognition of Francisella tularensis.* Proc Natl Acad Sci U S A, 2010. **107**(21): p. 9771-6.

211. Lippmann, J., et al., *Dissection of a type I interferon pathway in controlling bacterial intracellular infection in mice.* Cell Microbiol, 2011. **13**(11): p. 1668-82.
212. Sharma, S., et al., *Innate immune recognition of an AT-rich stem-loop DNA motif in the Plasmodium falciparum genome.* Immunity, 2011. **35**(2): p. 194-207.
213. Jameson, M.B., et al., *Clinical aspects of a phase I trial of 5,6-dimethylxanthenone-4-acetic acid (DMXAA), a novel antivascular agent.* Br J Cancer, 2003. **88**(12): p. 1844-50.
214. Lara, P.N., Jr., et al., *Randomized phase III placebo-controlled trial of carboplatin and paclitaxel with or without the vascular disrupting agent vadimezan (ASA404) in advanced non-small-cell lung cancer.* J Clin Oncol, 2011. **29**(22): p. 2965-71.
215. Perera, P.Y., et al., *Activation of LPS-inducible genes by the antitumor agent 5,6-dimethylxanthenone-4-acetic acid in primary murine macrophages. Dissection of signaling pathways leading to gene induction and tyrosine phosphorylation.* J Immunol, 1994. **153**(10): p. 4684-93.
216. Prantner, D., et al., *5,6-Dimethylxanthenone-4-acetic acid (DMXAA) activates stimulator of interferon gene (STING)-dependent innate immune pathways and is regulated by mitochondrial membrane potential.* J Biol Chem, 2012. **287**(47): p. 39776-88.
217. 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.* J Immunol, 2013. **190**(10): p. 5216-25.
218. Kramer, M.J., R. Cleeland, and E. Grunberg, *Antiviral activity of 10-carboxymethyl-9-acridanone.* Antimicrob Agents Chemother, 1976. **9**(2): p. 233-8.
219. Taylor, J.L., C.K. Schoenherr, and S.E. Grossberg, *High-yield interferon induction by 10-carboxymethyl-9-acridanone in mice and hamsters.* Antimicrob Agents Chemother, 1980. **18**(1): p. 20-6.
220. Cavlar, T., et al., *Species-specific detection of the antiviral small-molecule compound CMA by STING.* EMBO J, 2013. **32**(10): p. 1440-50.
221. 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. 576-91.
222. 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. 1065-72.
223. Fernandes-Alnemri, T., et al., *AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA.* Nature, 2009. **458**(7237): p. 509-13.
224. Hornung, V., et al., *AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC.* Nature, 2009. **458**(7237): p. 514-8.
225. Muruve, D.A., et al., *The inflammasome recognizes cytosolic microbial and host DNA and triggers an innate immune response.* Nature, 2008. **452**(7183): p. 103-7.
226. Roberts, T.L., et al., *HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA.* Science, 2009. **323**(5917): p. 1057-60.
227. 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. 501-5.
228. Unterholzner, L., et al., *IFI16 is an innate immune sensor for intracellular DNA.* Nat Immunol, 2010. **11**(11): p. 997-1004.
229. Zhang, Z., et al., *The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells.* Nat Immunol, 2011. **12**(10): p. 959-65.
230. Ferguson, B.J., et al., *DNA-PK is a DNA sensor for IRF-3-dependent innate immunity.* Elife, 2012. **1**: p. e00047.
231. 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**(8): p. 4541-5.
232. Kondo, T., et al., *DNA damage sensor MRE11 recognizes cytosolic double-stranded DNA and induces type I interferon by regulating STING trafficking.* Proc Natl Acad Sci U S A, 2013. **110**(8): p. 2969-74.

Supplementary data

(attached on CD-ROM)

RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling

Filiz Civril, Matthew Bennett, Manuela Moldt, Tobias Deimling, Gregor Witte, Stefan Schiesser, Thomas Carell & Karl-Peter Hopfner

EMBO rep. 2011 Oct 28; **12**(11): 1127 - 34

Supplementary Information

Supplementary Materials and Methods

Constructs and cloning

The sequence encoding full-length *Mus musculus* RIG-I was amplified from total mouse cDNA (courtesy of Dr. Stefan Bauersachs) and cloned into pET28 (Novagen). ΔRD RIG-I was subcloned in a modified pET21a (Novagen) to express 6xHis-MBP-TEV-RIG-I (1-797). The mutants (T348E, E374Q, K380E, double mutant Q508AQ512A, R547E, and R731A,) were generated by site directed mutagenesis with PfuUltra (Stratagene).

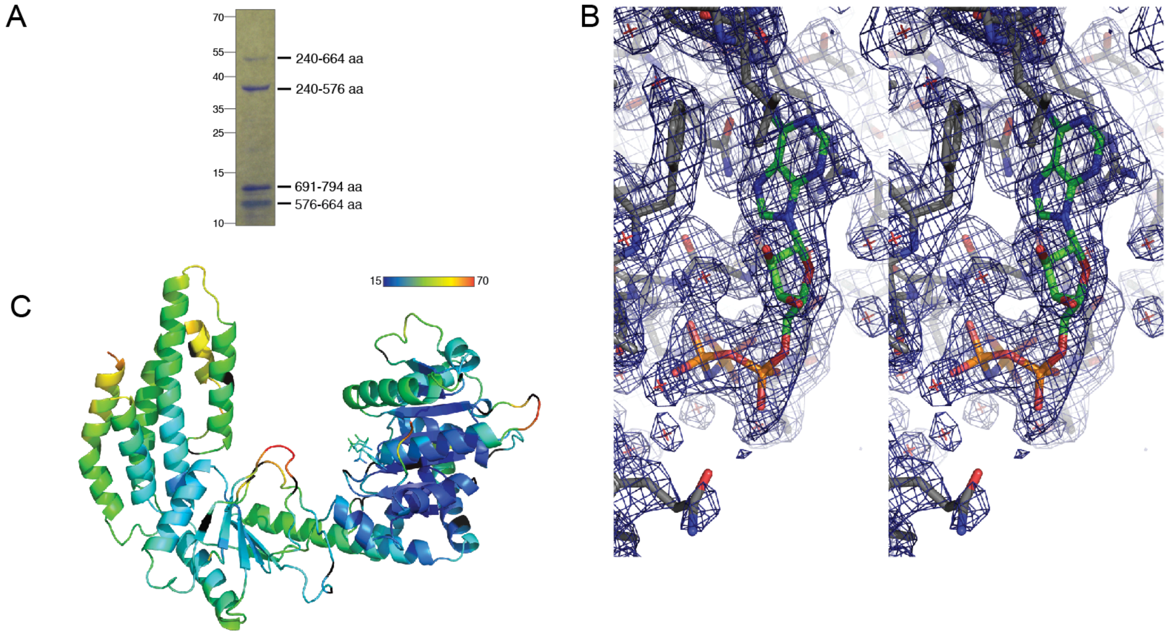
Expression and protein purification

All constructs were expressed in *E.coli* Rosetta (DE3) or B834 (DE3) strains. Native protein was grown in LB media and selenomethionine labeled protein in modified M9 media. Bacteria were grown until an OD_{600} of 0.6 to 0.8 and induced at 18°C for 16 to 18 hrs with 0.1mM IPTG. Proteins were purified by Ni-NTA agarose resin and when applicable incubated with TEV protease (ratio 1:50) at 4°C for 20 to 22 hrs to remove the 6xHis-MBP-tag. The proteins were further purified by ion exchange followed by size exclusion chromatography using a Superdex 200 column (GE Healthcare) equilibrated in 20mM TRIS pH 7.5, 150mM NaCl and 1mM DTT. Purified RIG-I ΔRD was concentrated to 27.5 mg/ml for crystallization. All purified proteins were frozen in liquid N_2 and stored at -80°C.

Crystallization

Selenomethionine labeled RIG-I^{SF2} was crystallized using an in-drop proteolysis approach. To purified RIG-I (27.5 mg/ml) 1:10 (v/v) of 50 mM of ATP derivative adenosine 5′-(β,γ-imido)triphosphate (AMP-PNP) and 1:500 (w/w) 1mg/ml subtilisin were added and the mix was crystallized by sitting drop vapor diffusion method in crystallization solution [100mM BIS-TRIS pH 6.6 and 22% (w/v) PEG3350]. The crystals appeared in 1 day at 20°C and were soaked in cryoprotectant solution (15% ethane-1,2-diol in crystallization solution) and flash frozen. SDS-PAGE analysis of the crystals (Suppl. Fig. 1A) shows that the protease cleaved between the N-terminal CARDs and SF2 as well as removing two internal loops of SF2.

Supplementary Figures

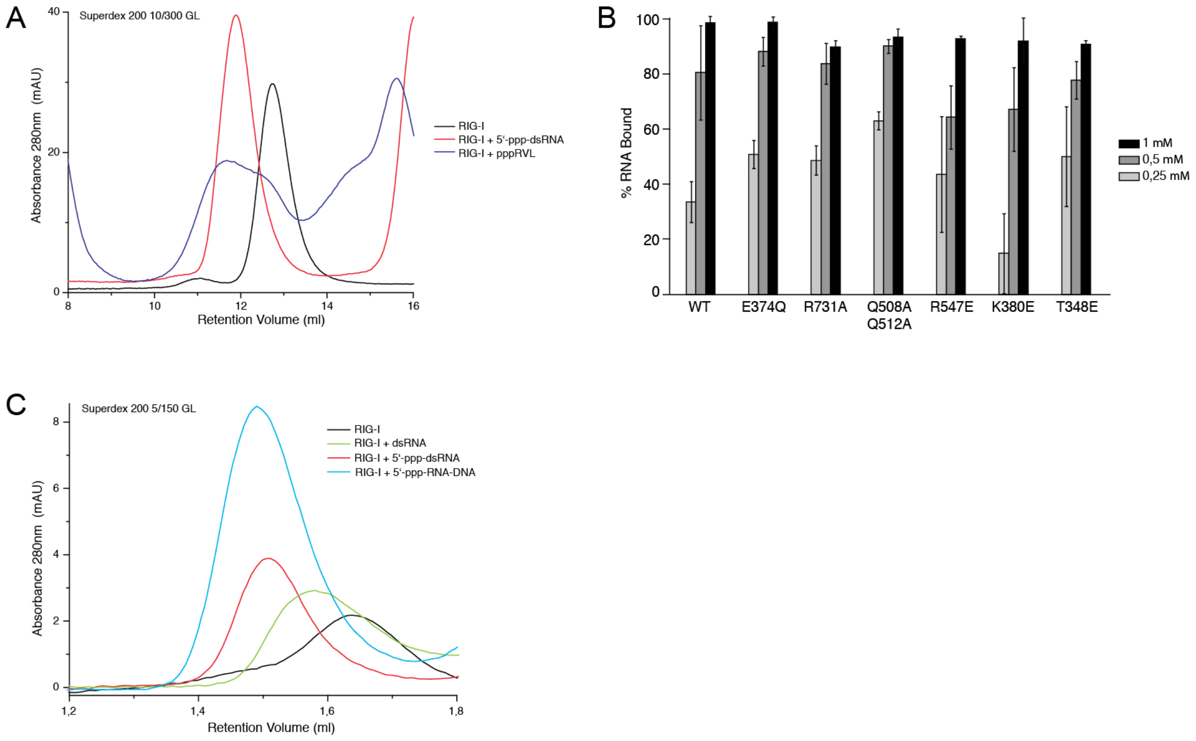


Supplementary Figure 1: Structure of RIG-I SF2 Domain

A) Crystals of RIG-I SF2 domain were collected and washed three times with reservoir solution and loaded on SDS-PAGE. The bands were assigned the indicated regions based on the ordered parts in crystal structure.

B) 2Fo-Fc electron density overlaid with the final model around the AMP-PNP molecule. The $2F_0$-F_c map is contoured at 1 σ.

C) View of the RIG-I SF2 domain structure in the orientation of Fig 1A, colored according to the Cα B-factors. Black residues are glycines.

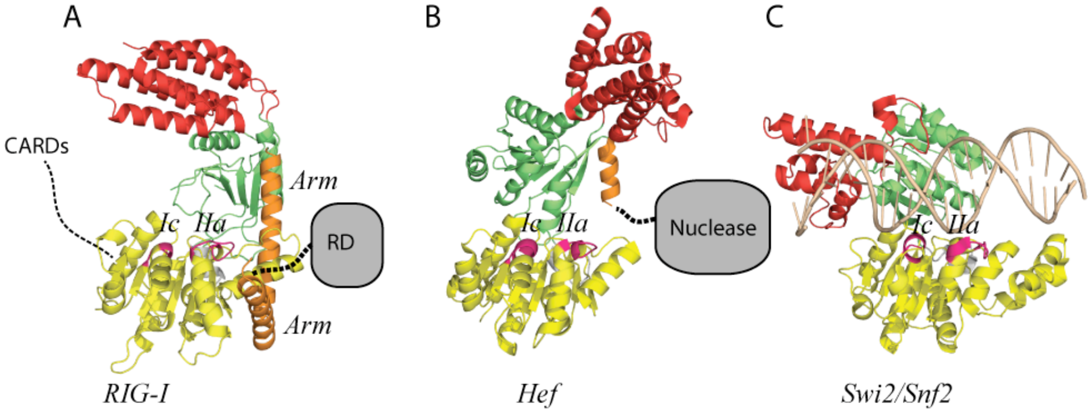


Supplementary Figure 2: RNA dependent activation of RIG-I

A) Size exclusion chromatography analysis of RIG-I (black) preincubated with either chemically synthesized 5’-ppp-dsRNA (25-mer) (red) or *in vitro* transcribed 5’-ppp containing rabies virus leader RNA 58-mer (pppRVL) (blue). Upon 5’-ppp-dsRNA binding RIG-I dimerizes as in the case of pppRVL (Cui et al, 2008).

B) Electron mobility shift analysis of RIG-I mutants bound to 25-mer dsRNA. Plotted bars: mean±sd (n=3)

C) Size exclusion chromatography analysis of RIG-I (black) preincubated with 5’-ppp-dsRNA (25-mer) (red) or dsRNA (25mer) (green) or 5’-ppp-RNA-DNA hybrid (cyan). Hybrid ppp-nucleic acid dimerizes RIG-I like ppp-dsRNA



Supplementary Figure 3: Comparison to Hef and Swi/Snf2 dsDNA binding proteins

Comparison of crystal structures of the SF2 domains of RIG-I (A, this work), Hef (B, (Nishino et al, 2005)) and Swi/Snf2 in complex with dsDNA (C, (Durr et al, 2005)) RIG-I and Hef have a related domain structure, although the orientation of domains 1 (yellow) and 2 (green) are more canonical in Hef, while oriented in a rather non-canonical oritentation in RIG-I by the RLR specific helical arm. RD and CARDs are attached to SF2 by linkers, while Hef contains a linked nuclease domain. Swi/Snf2 binds dsDNA at motifs Ic and IIa, which we identify by mutational analysis in RIG-I. The overall similarity of domains 1a in all three structures suggests a related mode of double-stranded nucleic acid binding.

Supplementary references

Cui S, Eisenacher K, Kirchhofer A, Brzozka K, Lammens A, Lammens K, Fujita T, Conzelmann KK, Krug A, Hopfner KP (2008) The C-terminal regulatory domain is the RNA 5'-triphosphate sensor of RIG-I. *Mol Cell* **29**(2)**:** 169-179

Durr H, Korner C, Muller M, Hickmann V, Hopfner KP (2005) X-ray structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. *Cell* **121**(3)**:** 363-373

Nishino T, Komori K, Tsuchiya D, Ishino Y, Morikawa K (2005) Crystal structure and functional implications of Pyrococcus furiosus hef helicase domain involved in branched DNA processing. *Structure* **13**(1)**:** 143-153

Species-specific detection of the antiviral small-molecule compound CMA by STING

Taner Cavlar*, Tobias Deimling*, Andrea Ablasser, Karl-Peter Hopfner & Veit Hornung

EMBO J. 2013 March 18; **32**(10): 1440 – 50

Supplementary figure legends

Figure S1. Dose-dependent antiviral activity of CMA

A, Antiviral activity of CMA in macrophages was studied using a FFLuc encoding VSV replicon. B, Macrophages were stimulated with decreasing concentrations of CMA (500μg/ml – 31,25μg/ml in two-fold dilutions) and subsequently infected with the VSV replicon particles. 18h after stimulation luciferase activity was assessed (left panel) and cell viability was determined (right panel). C, pIFNβ-firefly-luciferase macrophages were treated with decreasing concentrations of CMA (500μg/ml – 31,25μg/ml in two-fold dilutions). After 18h luciferase activity was assessed (left panel) and cell viability was determined (right panel). Representative results out of 3 independent experiments are depicted.

Figure S2. CMA-triggered type I IFN response is TLR- and MAVS-independent and mediated via TBK1/IKKe

A, Wild type and TBK1/IKKε double-deficient MEFs were transfected with poly(I:C), pppRNA, plasmid DNA or stimulated with CMA (500μg/ml) for 2h. Cells were collected and assessed for phospho-IRF3. Bone marrow-derived wild type (B), MyD88$^{-/-}$ (C) and Trif$^{-/-}$ (D) macrophages (MΦ) were stimulated as described above. After 18h supernatants were collected and IP-10 production was determined by ELISA. In addition phospho-IRF3 was analyzed 2h after stimulation. (E) Bone marrow-derived wild type or (F) MAVS$^{-/-}$ macrophages were stimulated with pppRNA or CMA. ELISA and western blot were performed as described for A-D. Representative results out of 2 independent experiments are depicted.

Figure S3. Antiviral activity of CMA is STING-dependent

A, Supernatants of pretreated macrophages were assessed for antiviral properties using the VSV-FFLuc replicon in supernatant transfer experiments. B, bone marrow-derived wild type or STING-deficient macrophages were left untreated, transfected with pppRNA or stimulated with CMA (500μg/ml). After 18h the supernatants were harvested and diluted 1:20. The diluted supernatants (grey bars) were then transferred onto macrophages that were subsequently infected with VSV replicon particles. As a control infected macrophages were left untreated or stimulated with the equivalent concentration of CMA (25μg/ml; 1:20 dilution of 500 μg/ml) in the diluted supernatants (hatched bars). Representative results out of 2 independent experiments are depicted.

Figure S4. HEK293T stably expressing murine STING respond to CMA

A-B, HEK293T cells stably expressing murine STING were transiently transfected with a total amount of 200ng DNA per 96-well, whereas 12.5ng of pIFNβ-GLuc (A) or pELAM-GLuc (B) reporter plasmid were included. An empty pCI vector served as a stuffer. After 24h cells were transfected with poly(I:C), pppRNA, ISD and c-diGMP or stimulated with LPS or decreasing concentrations of CMA (from 500μg/ml to 15,62μg/ml). Luciferase activity was measured after an additional period of 24h in the supernatant (upper panel). Cell viability was determined by CellTiter-Blue assay (lower panel). Representative results out of 3 independent experiments are depicted.

Figure S5. R231A dissociates c-diGMP- from CMA-mediated STING activation

A-C, 293T cells were transiently transfected with the indicated STING constructs (25, 12.5, 6.25 and 0ng), whereas 12.5ng of pIFNβ-GLuc reporter plasmid were included. For titrations, an empty pCI vector served as a stuffer to obtain 200ng total plasmid DNA. After 24h cells were stimulated with c-diGMP or CMA (500μg/ml and 125μg/ml). Luciferase activity was measured after an additional period of 24h in the supernatant and data were normalized to the condition without STING overexpression. Plasmids coding for full-length murine STING (A), murine STING-I199N (B) and murine STING-R231A (C) were tested. D, Expression of the above-described constructs was studied in 293T cells 24h after transfection (200ng per 96-well) using western blot, whereas β-Actin served as a loading control. Representative results out of 3 independent experiments are depicted.

Figure S6. The CTT is not required for CMA binding

The interaction of mSTING-LBD-ΔCTT with c-diGMP, c-diAMP and CMA detected by thermal shift assay. Purified mSTING LBD lacking the CTT (138-341) (i) was analyzed with different concentrations of c-diGMP/c-diAMP/CMA. (ii) Thermal shifts of (iii) fluorescence intensity vs. temperature are shown. Representative results out of 2 experiments are depicted.

Figure S7. Comparison of mouse STING:CMA and human STING:c-diGMP

Superposition of mouse STING, bound to two CMA molecules (brown ribbon model and with magenta CMA stick models), with human STING, bound to c-diGMP (PDB accession code 4F5D, cyan ribbon model with green c-diGMP stick model). The superposition shows a similar dimer structure, shared ligand binding site and overall similar folding of the lid region, suggesting that CMA activates murine STING by a similar structural mechanism than c-diGMP.

Table S1. Data Collection and Refinement Statistics

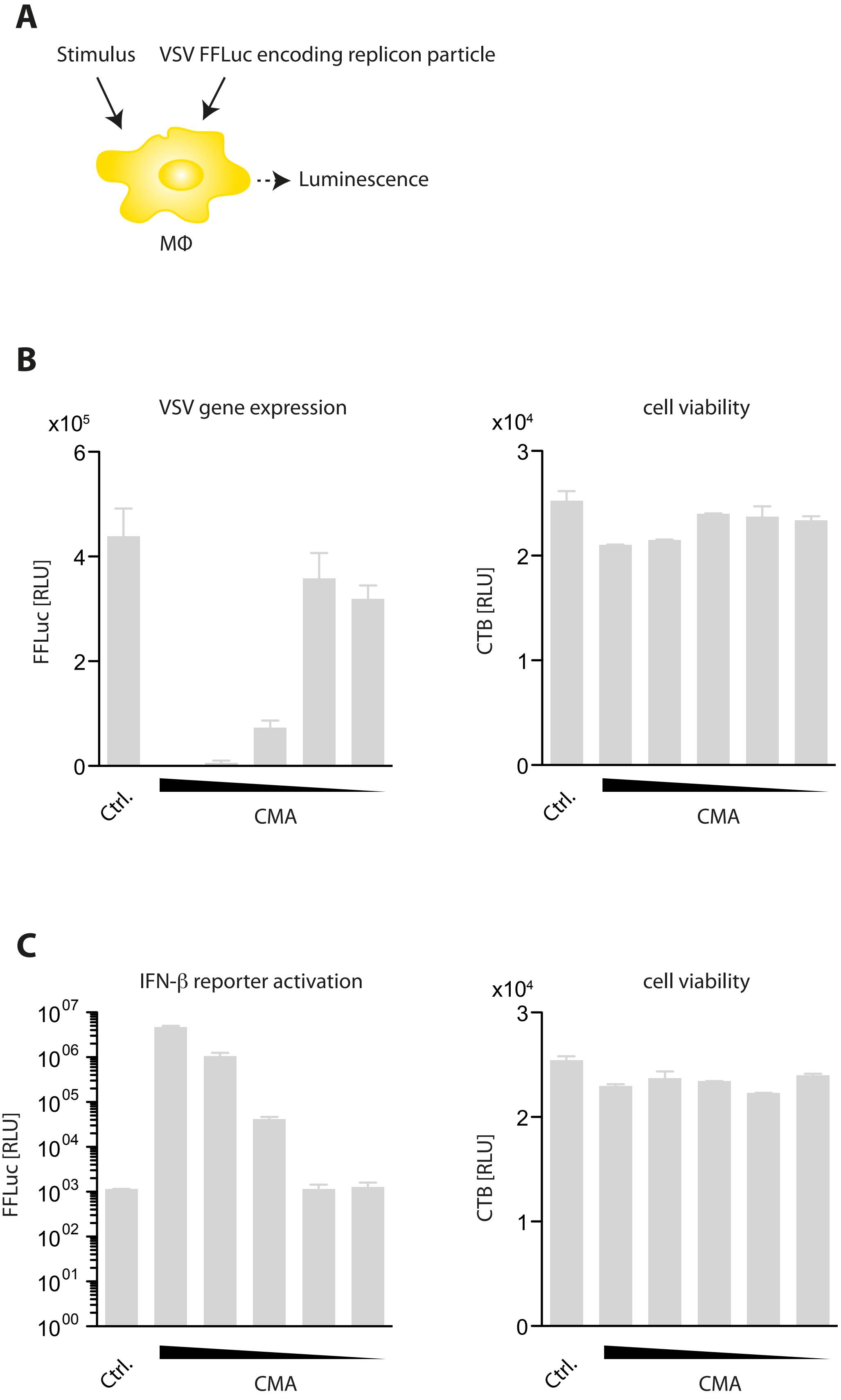

A
Stimulus
VSV FFLuc encoding replicon particle
Luminescence
MΦ
B
VSV gene expression
x10^{5}
6
4
2
0
FFLuc [RLU]
Ctrl.
CMA
cell viability
x10^{4}
3
2
1
0
CTB [RLU]
Ctrl.
CMA
C
IFN-β reporter activation
10^{07}
10^{06}
10^{05}
10^{04}
10^{03}
10^{02}
10^{01}
10^{00}
FFLuc [RLU]
Ctrl.
CMA
cell viability
x10^{4}
3
2
1
0
CTB [RLU]
Ctrl.
CMA

Figure S1

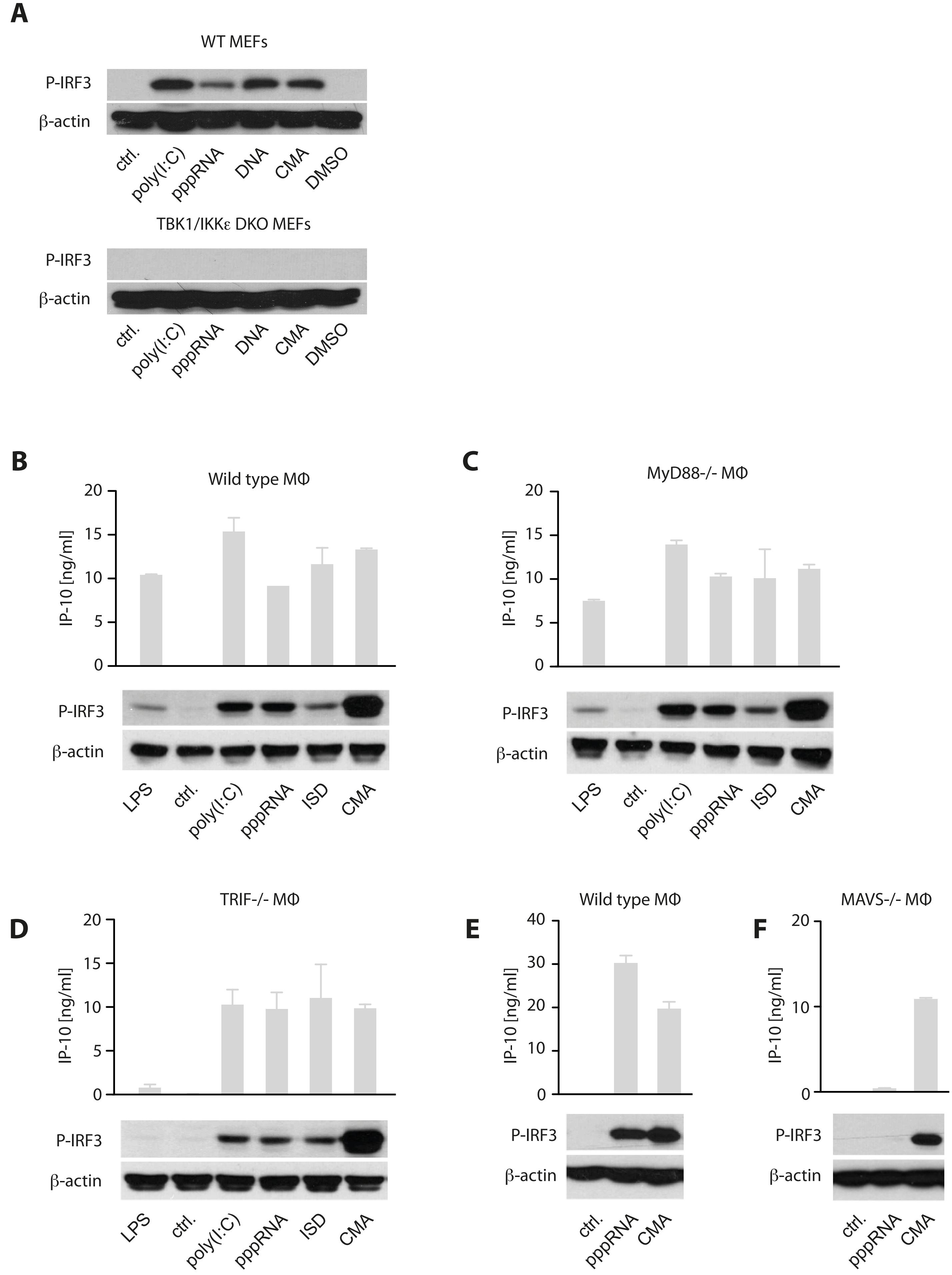

A
WT MEFs
P-IRF3
β-actin
ctrl.
poly(I:C)
pppRNA
DNA
CMA
DMSO
TBK1/IKKε DKO MEFs
P-IRF3
β-actin
ctrl.
poly(I:C)
pppRNA
DNA
CMA
DMSO
B
Wild type MΦ
IP-10 [ng/ml]
20
15
10
5
0
P-IRF3
β-actin
LPS
ctrl.
poly(I:C)
pppRNA
ISD
CMA
C
MyD88-/- MΦ
IP-10 [ng/ml]
20
15
10
5
0
P-IRF3
β-actin
LPS
ctrl.
poly(I:C)
pppRNA
ISD
CMA
D
TRIF-/- MΦ
IP-10 [ng/ml]
20
15
10
5
0
P-IRF3
β-actin
LPS
ctrl.
poly(I:C)
pppRNA
ISD
CMA
E
Wild type MΦ
IP-10 [ng/ml]
40
30
20
10
0
P-IRF3
β-actin
ctrl.
pppRNA
CMA
F
MAVS-/- MΦ
IP-10 [ng/ml]
20
10
0
P-IRF3
β-actin
ctrl.
pppRNA
CMA

Figure S2

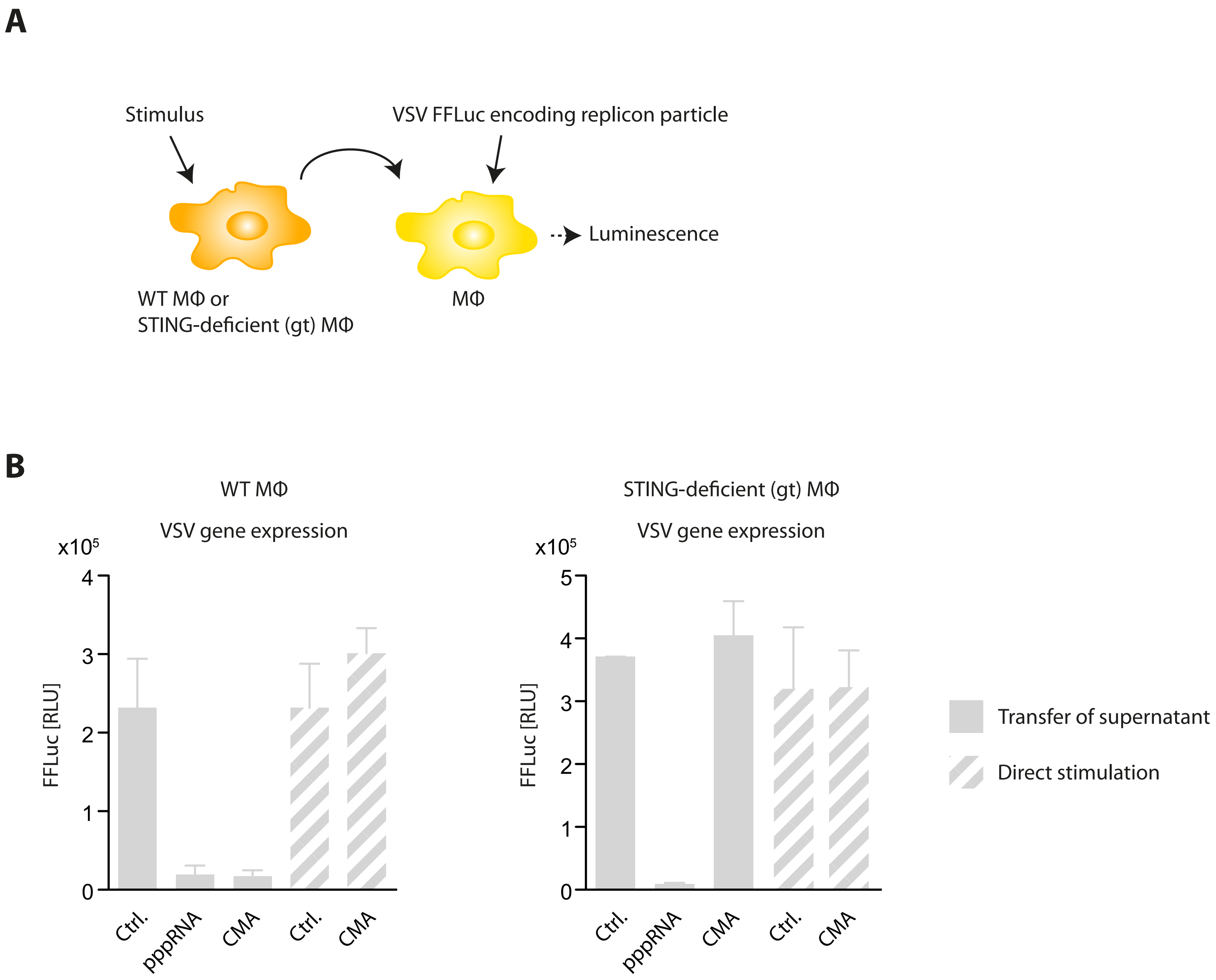

A
Stimulus
VSV FFLuc encoding replicon particle
Luminescence
WT MΦ or
STING-deficient (gt) MΦ
MΦ
B
WT MΦ
VSV gene expression
x10⁵
FFLuc [RLU]
0
1
2
3
4
Ctrl.
pppRNA
CMA
Ctrl.
CMA
STING-deficient (gt) MΦ
VSV gene expression
x10⁵
FFLuc [RLU]
0
1
2
3
4
5
Ctrl.
pppRNA
CMA
Ctrl.
CMA
Transfer of supernatant
Direct stimulation

Figure S3

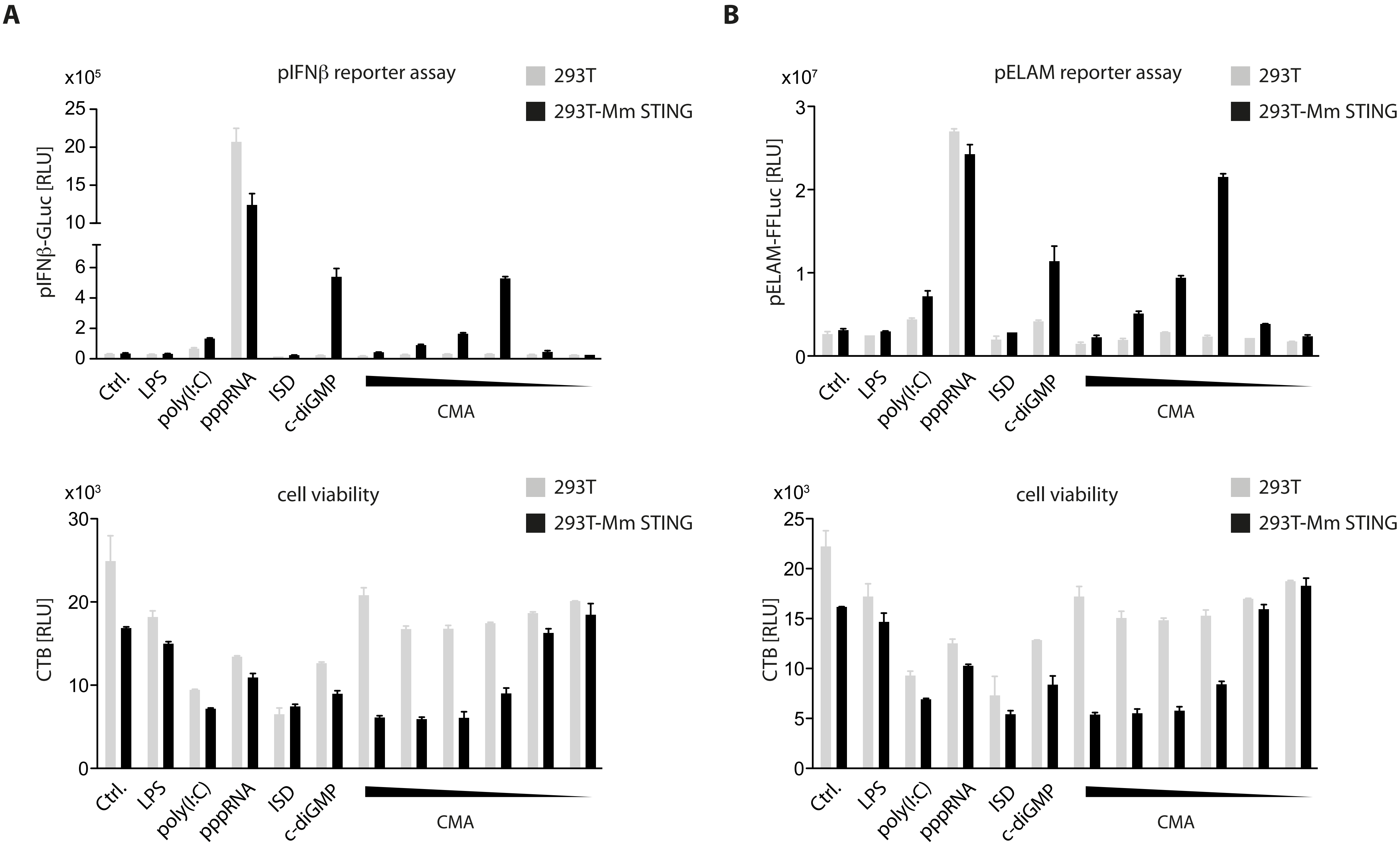

A
pIFNβ reporter assay
x10⁵
pIFNβ-GLuc [RLU]
25
20
15
10
6
4
2
0
293T
293T-Mm STING
Ctrl.
LPS
poly(I:C)
pppRNA
ISD
c-diGMP
CMA
B
pELAM reporter assay
x10⁷
pELAM-FFLuc [RLU]
3
2
1
0
293T
293T-Mm STING
Ctrl.
LPS
poly(I:C)
pppRNA
ISD
c-diGMP
CMA
cell viability
x10³
CTB [RLU]
30
20
10
0
293T
293T-Mm STING
Ctrl.
LPS
poly(I:C)
pppRNA
ISD
c-diGMP
CMA
cell viability
x10³
CTB [RLU]
25
20
15
10
5
0
293T
293T-Mm STING
Ctrl.
LPS
poly(I:C)
pppRNA
ISD
c-diGMP
CMA

Figure S4

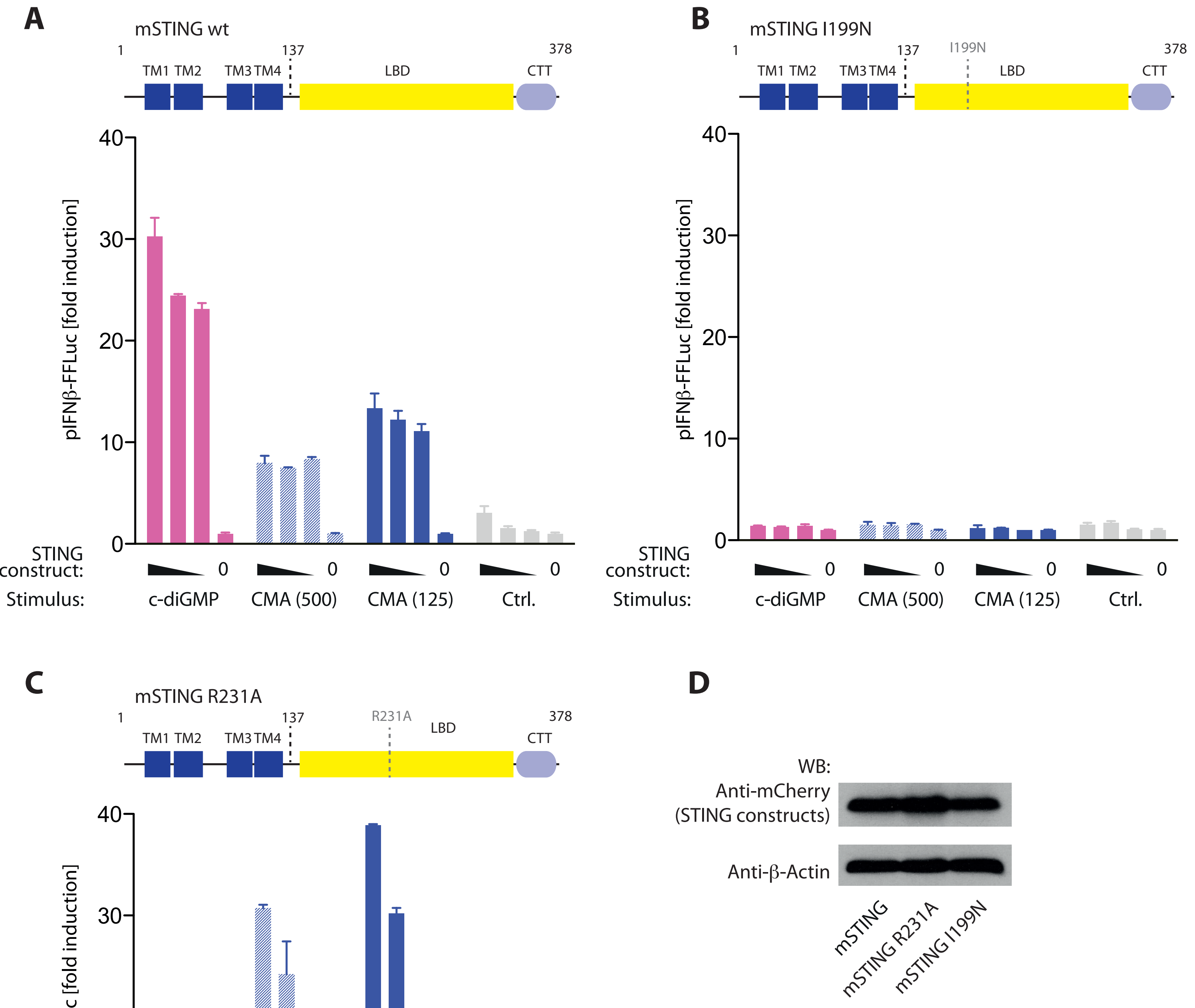

A
mSTING wt
1
137
378
TM1
TM2
TM3
TM4
LBD
CTT
pIFNβ-FFLuc [fold induction]
40
30
20
10
0
STING construct:
Stimulus:
0
0
0
0
c-diGMP
CMA (500)
CMA (125)
Ctrl.
B
mSTING I199N
1
137
I199N
378
TM1
TM2
TM3
TM4
LBD
CTT
pIFNβ-FFLuc [fold induction]
40
30
20
10
0
STING construct:
Stimulus:
0
0
0
0
c-diGMP
CMA (500)
CMA (125)
Ctrl.
C
mSTING R231A
1
137
R231A
LBD
378
TM1
TM2
TM3
TM4
CTT
pIFNβ-FFLuc [fold induction]
40
30
20
10
0
STING construct:
Stimulus:
0
0
0
0
c-diGMP
CMA (500)
CMA (125)
Ctrl.
D
WB:
Anti-mCherry
(STING constructs)
Anti-β-Actin
mSTING
mSTING R231A
mSTING I199N

Figure S5

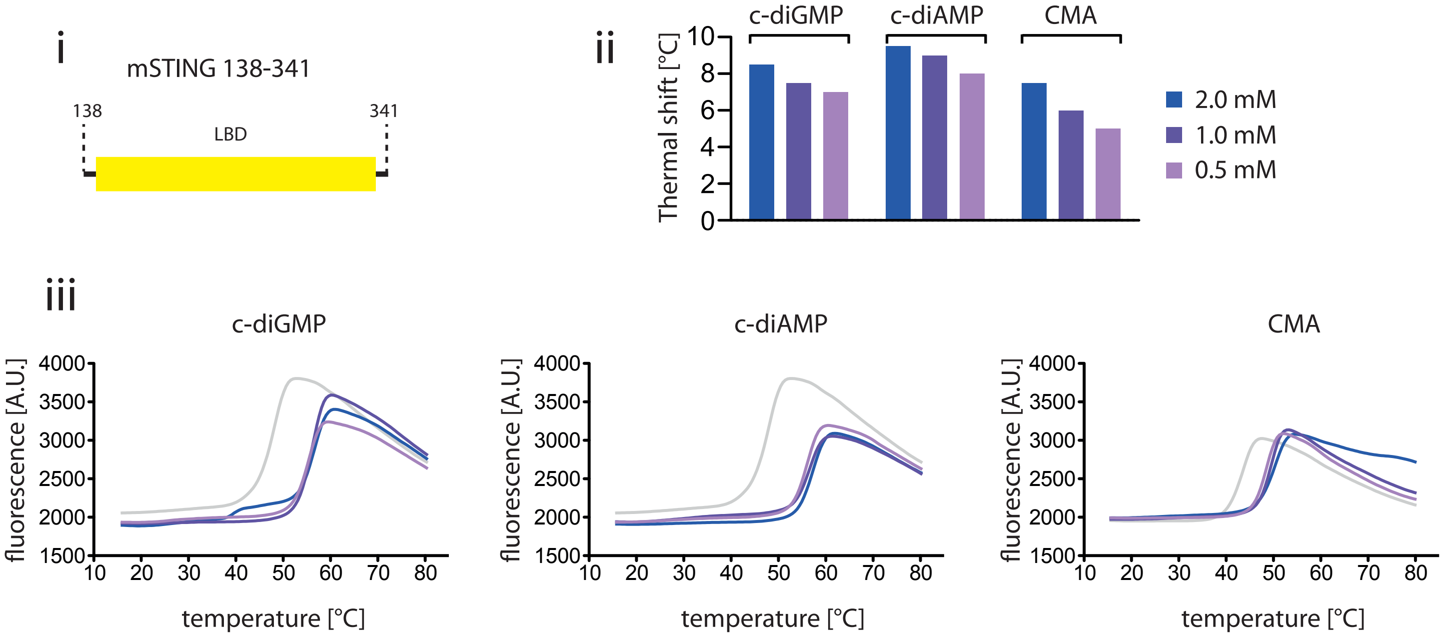

i
mSTING 138-341
138
341
LBD
ii
c-diGMP
c-diAMP
CMA
Thermal shift [°C]
0
2
4
6
8
10
2.0 mM
1.0 mM
0.5 mM
iii
c-diGMP
c-diAMP
CMA
fluorescence [A.U.]
1500
2000
2500
3000
3500
4000
10
20
30
40
50
60
70
80
temperature [°C]

Figure S6

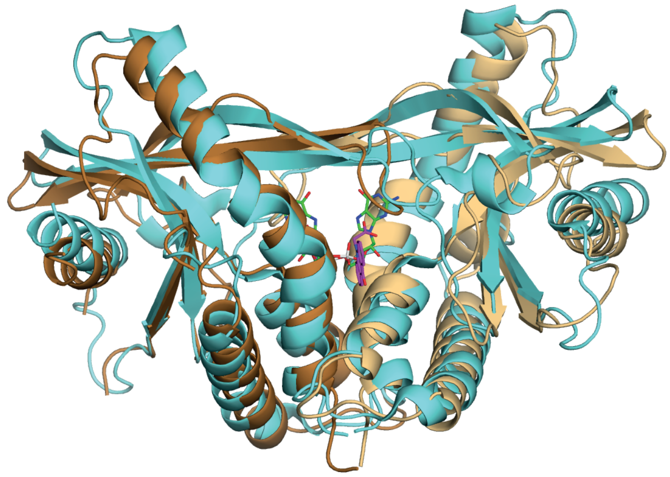


Figure S7

Data collection	
Space group	R3
Cell dimensions	
a, *b*, *c* (Å)	108,54 108.54, 101.83
a, b, g (°)	90.0, 90.0, 120.0
Resolution (Å)	50.0 – 2.75 (2.91 – 2.75) *
R_{meas} *(%)*	5.1 (77.2)
$I / \sigma I$	19.17 (1.97)
Completeness (%)	99.3 (98.3)
Redundancy	4.0 (3.9)
Refinement	
Resolution (Å)	34.38 – 2.75
No. reflections	11444
R_{work} / R_{free} (%)	21.0 / 23.7
No. atoms	
Protein	2989
Ligand/ion	38
Water	32
B-factors	
Protein	73,5
Ligand/ion	39,2
Water	62,5
R.m.s deviations	
Bond lengths (Å)	0,007
Bond angles (°)	1,159
Ramachandran values	
Favored	343
Allowed	19
Outliers	0
PDB Accession code	4JC5

* Values in parentheses are for highest resolution shell

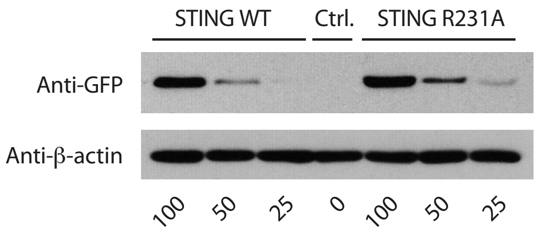
Table S1. Data Collection and Refinement Statistics

cGAS produces a 2’-5’-linked cyclic dinucleotide second messenger that activates STING

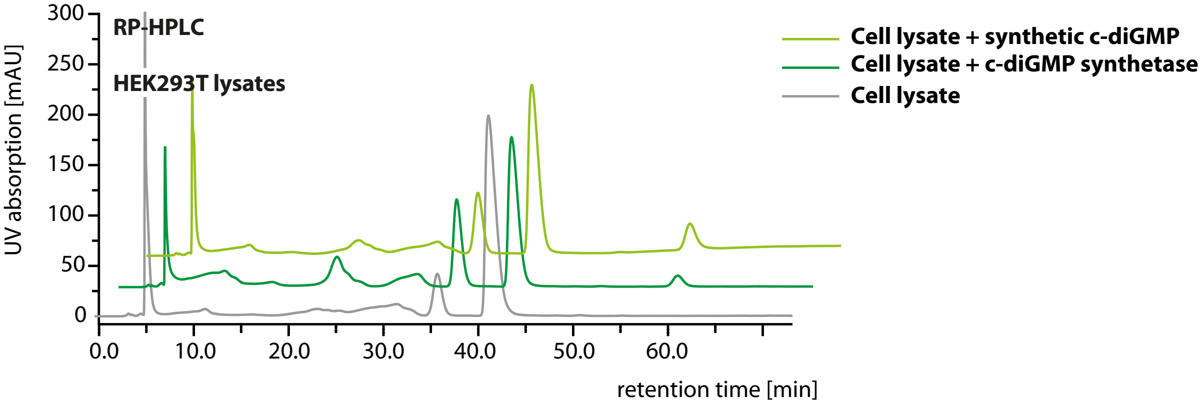
Andrea Ablasser, Marion Goldeck, Taner Cavlar, Tobias Deimling, Gregor Witte,
Ingo Röhl, Karl-Perter Hopfner, Janos Ludwig & Veit Hornung

Nature 2013 June 20; **498**(7454): 380 - 4

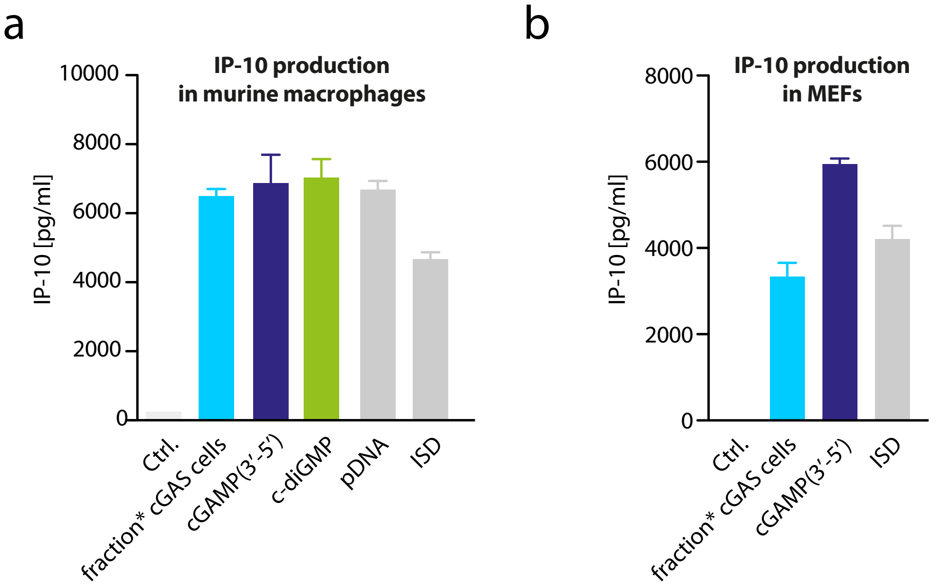
doi:10.1038/nature12306



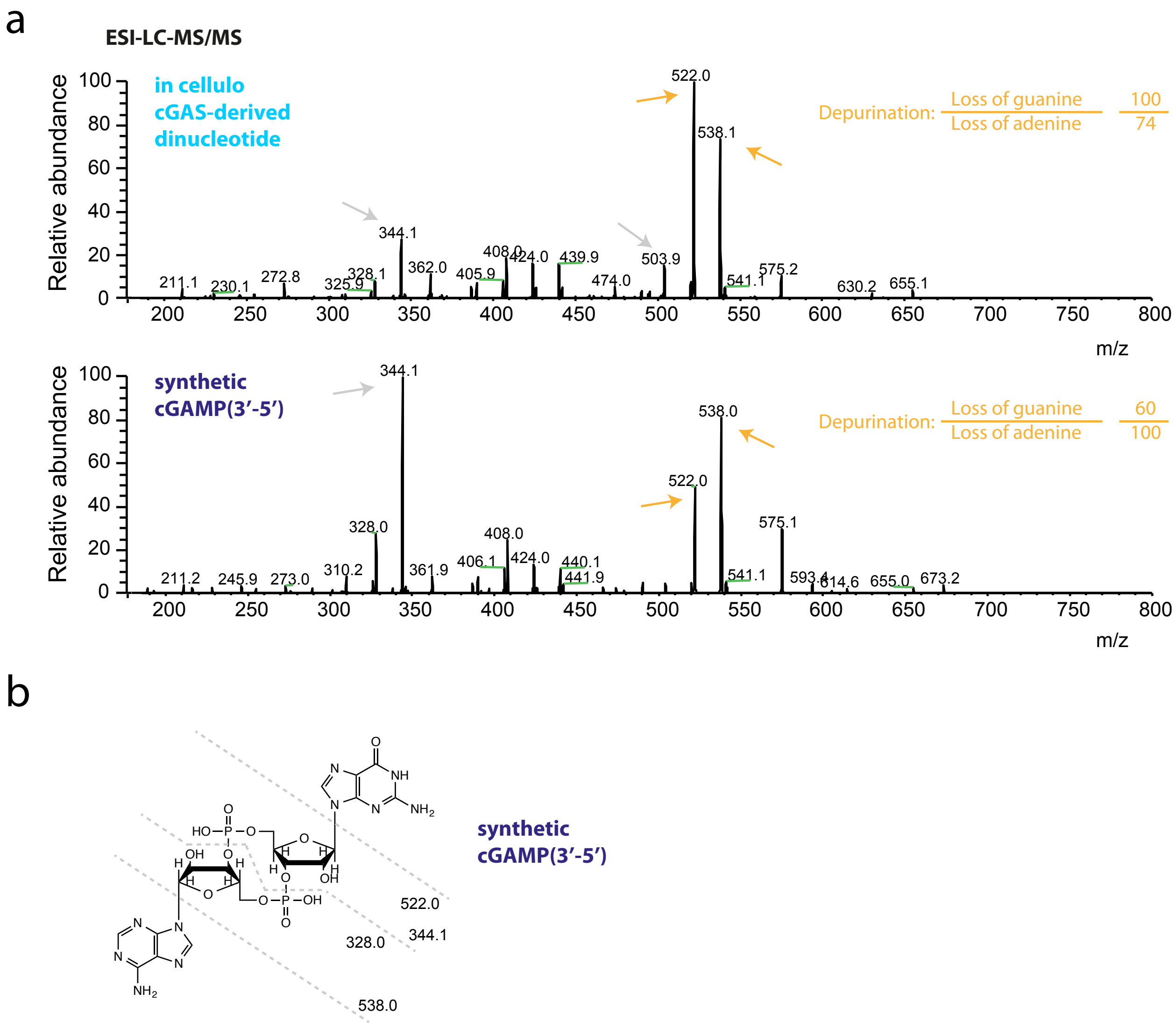
Supplementary Figure 1 | Expression analysis of wild type murine STING and the R231A mutant. Expression analysis of wild type murine STING and its mutant R231A (both harboring a GFP-Tag) by immunoblotting in HEK293T cells 16h after transfection. Immunoblot for β-Actin served as a control. Representative data out of two independent experiments are shown.



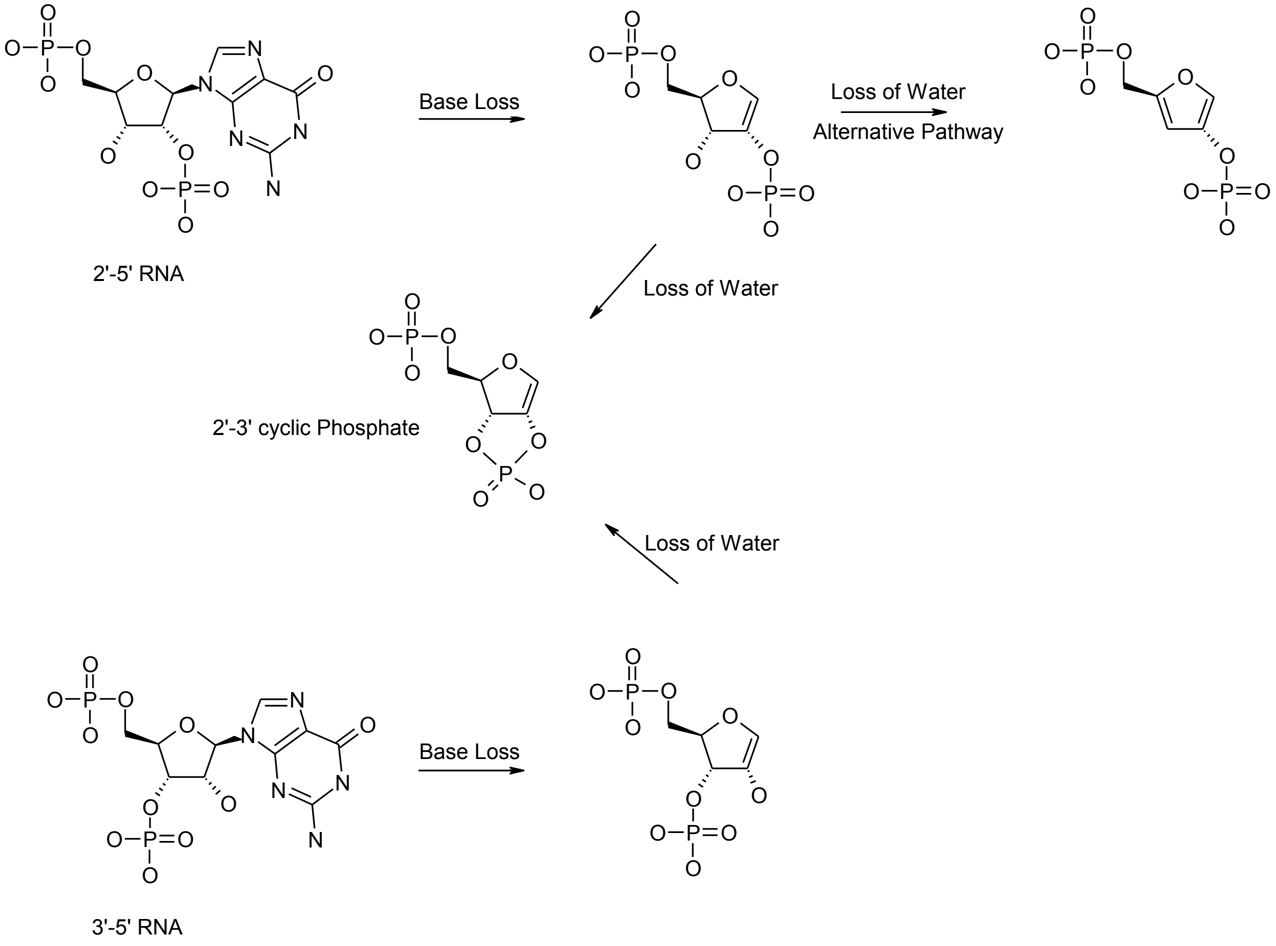
Supplementary Figure 2 | Overexpression of cyclic di-GMP synthetase induces endogenous expression of cyclic di-GMP in HEK293T cells. Chromatograms from purified lysates of untreated HEK293T cells (grey) and cyclic di-GMP synthetase overexpressing HEK293T cells (dark green) are shown together with synthetic cyclic di-GMP spiked into untreated HEK293T lysate (light green). Representative data out of two independent experiments are shown.



Supplementary Figure 3 | The endogenous cGAS product is a potent trigger of type I IFN response in BMDMs and MEFs. (**a-b**) BMDMs (**a**) and primary MEFs (**b**) were transfected with purified cGAS product extracted from HEK293T cells overexpressing cGAS, synthetic cGAMP(3'-5'), cyclic di-GMP, plasmid DNA and ISD. After 14h supernatants were collected and the production of IP-10 was assessed by ELISA. Representative data out of three independent experiments are shown as mean values + SEM.



Supplementary Figure 4 | Tandem mass spectrometry of the cGAS-derived product suggests the presence of a 2'-5' phosphodiester linkage. (**a**) Tandem MS spectra from in cellulo synthesized cGAS product (light blue) and synthetic cGAMP (3´-5´) (dark blue) is shown. Grey arrows highlight an ion product with m/z 344.1, indicative of a nucleotide with 2'-3'-cyclic phosphates. Orange arrows highlight signals consistent with depurination of the dinucleotides (m/z 522.0 and m/z 538.1). Representative data out of three independent experiments are shown. (**b**) Expected fragmentation pathway of synthetic cGAMP(3'-5') is shown.



Supplementary Figure 5 | Tandem mass spectrometry fragmentation of 2´-5´ vs. 3´-5´linked RNA.

Supplementary Note 1 | Analysis of tandem MS spectra suggests the presence of a 2'-5' phosphodiester bond within the cGAS-derived dinucleotide product.

The tandem MS^2 spectra of synthetic cGAMP(3'-5') and the cGAS-derived product isolated from cell culture lysates showed significant differences, when analyzed by tandem MS in negative ion mode. Among other differences, the cGAS-derived dinucleotide showed a higher propensity for depurination of the guanosine, indicative of a lower stability of the N-glycosidic bond at the ribose ring in comparison to the synthetic cGAMP(3'-5') molecule (Supplementary Figure 4). A likely scenario for this change in stability could be ascribed to a 2'-instead of 3'-phosphodiester bond at the ribose ring of the guanosine.

The fragmentation of 3'-5'-RNA mainly leads to the y- and c-fragments induced by a cleavage of a 5'-P-O bond [1,2]. For a cyclic dinucleotide this fragmentation pathway will open the cyclic dinucleotide structure by forming a 2'-3'-cyclic phosphate, but it is mass neutral in the first step (Supplementary Figure 5). In a second fragmentation step the same fragmentation pathway will dissociate the dinucleotide in two nucleotides with 2'-3'-cyclic phosphates, with m/z = 344 (G) and m/z = 328 (A). The product ion with m/z = 344 was detected with highest abundance as the main fragmentation product of synthetic cGAMP with two 3'-5'-phosphodiester groups. This result is in sharp contrast to the spectrum obtained for the cGAS product isolated from cell lysate, with just 30% relative abundance for this product ion. These findings suggested a different configuration at the G-nucleotide in the synthetic cGAMP(3′-5′) versus the *in cellulo* produced cGAS product.

Another well-described fragmentation pathway of RNA is the neutral loss of a base that leads to the signals at m/z = 538 and m/z = 522, by forming a C1 to C2 double bond. Remarkably, for the synthetic cGAMP(3′-5′) adenine is eliminated with a higher relative abundance compared to guanine, whereas the ratio is reversed in the dinucleotide isolated from cGAS overexpressing cells. It is known, that the 2'-hydroxy group has a stabilizing effect on the N-glycosidic bond in RNA [3,4]. But in the case the 2'-hydroxy group is changed to 2'-phosphoester bond the stabilizing effect becomes lost and the neutral base loss of guanine should significantly increase.

Additional evidence for a guanosine 2'-phosphodiester is the strongly increased neutral loss of water in the cGAS-derived substance after elimination of the guanine base (m/z = 522 to m/z = 504 in the tandem MS^3 spectrum of m/z = 522; data not shown). The significantly higher loss of water in the natural compound over the synthetic reference was also observed by Wu et al. [5], but not attributed to different structures. For a 2'-5'-phosphodiester nucleotide bond two different pathways are possible for the loss of a water molecule after base elimination. The first pathway is similar between 2'-5' and 3'-5' RNA, where water is eliminated by forming a cyclic phosphate between the free 2'- or 3'-hydroxy and the adjacent phosphodiester. The alternative pathway is unique to the 2'-5' RNA, where water can eliminated from C3/C4 by forming another double bond and generating a thermodynamically stable aromatic furan ring (Supplementary Figure 4). No differences were found for the elimination of water after neutral loss of adenine, which supports a similar configuration in both compounds at the adenosine nucleotide.

All in all, these mass spectrometry findings strongly support the concept of a cyclic dinucleotide structure with guanosine connected 2’ via a phosphodiester to 5’-adenosine and adenosine connected 3’ via a phosphodiester to 5’-guanosine.

References

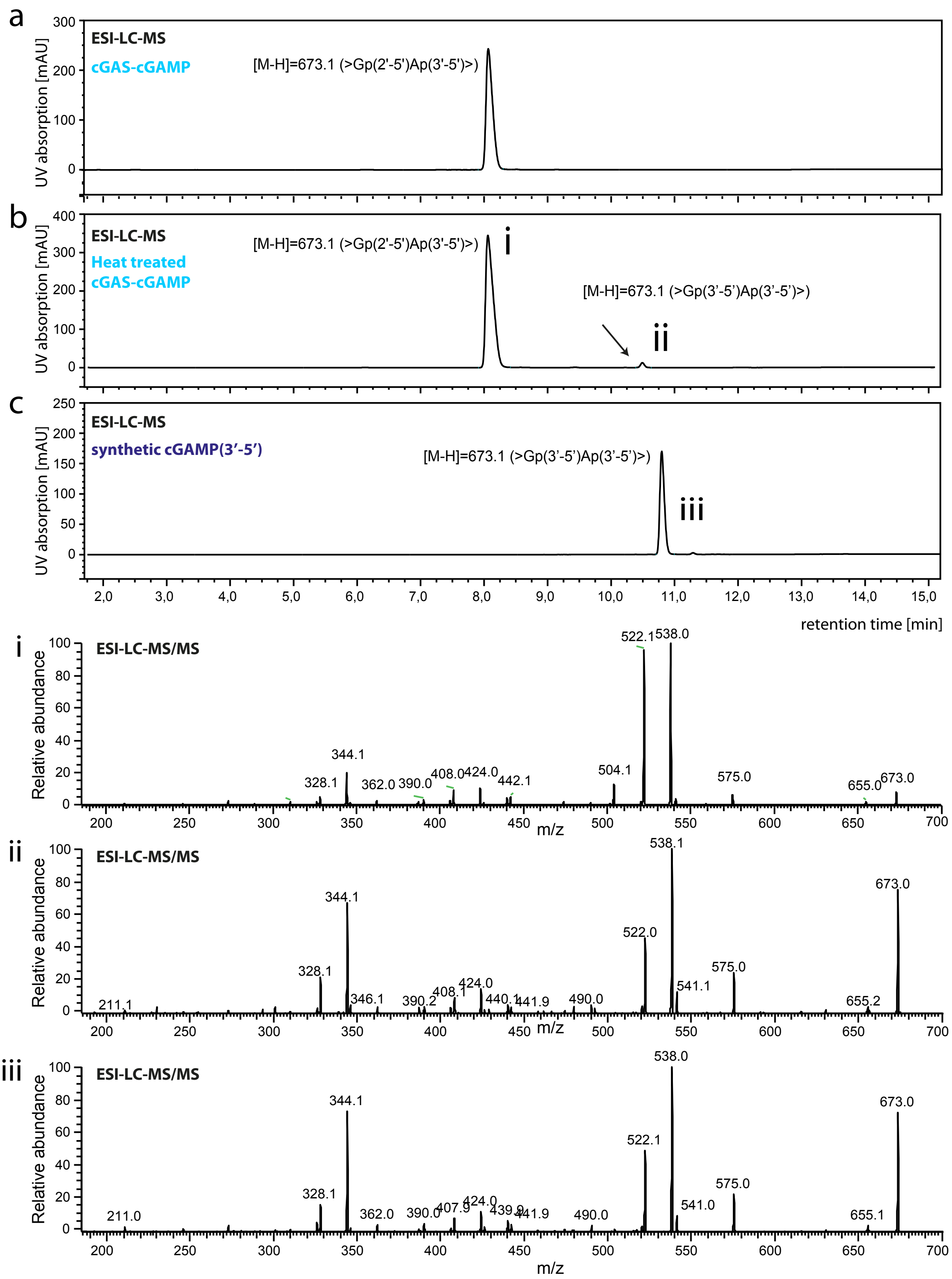
1 Andersen, T. E., Kirpekar, F. & Haselmann, K. F. RNA fragmentation in MALDI mass spectrometry studied by H/D-exchange: mechanisms of general applicability to nucleic acids. *J Am Soc Mass Spectrom* **17**, 1353-1368, doi:10.1016/j.jasms.2006.05.018 (2006).

2 Schurch, S., Bernal-Mendez, E. & Leumann, C. J. Electrospray tandem mass spectrometry of mixed-sequence RNA/DNA oligonucleotides. *J Am Soc Mass Spectrom* **13**, 936-945, doi:10.1016/S1044-0305(02)00413-0 (2002).

3 Tang, W., Zhu, L. & Smith, L. M. Controlling DNA Fragmentation in MALDI-MS by Chemical Modification. *Anal Chem* **69**, 302-312, doi:10.1021/ac960865o (1997).

4 Griffey, R. H., Greig, M. J., An, H., Sasmor, H. & Manalili, S. Targeted Site-Specific Gas-Phase Cleavage of Oligoribonucleotides. Application in Mass Spectrometry-Based Identification of Ligand Binding Sites. *Journal of the American Chemical Society* **121**, 474-475, doi:10.1021/ja983131x (1998).

5 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, doi:10.1126/science.1232458 (2013).



Supplementary Figure 6 | The endogenous cGAS product isomerizes at high temperature to cGAMP(3'-5'). (**a-c**) Chromatograms from in cellulo synthesized cGAS product (**a**), heat-treated in cellulo synthesized cGAS product (**b**) and synthetic cGAMP(3'-5') (**c**) are shown. Tandem MS spectra of elution peaks from (**b**) (i and ii) and (**c**) (iii) are depicted. Representative data out of two independent experiments are shown.

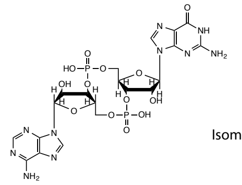
Supplementary Note 2 | The cGAS-derived dinucleotide product can isomerize into cGAMP(3’-5’) at high temperature

Interestingly, upon prolonged incubation at high temperature (90°), a small, but consistent fraction of the cGAS-derived dinucleotide product converted into a molecule with identical retention time in RP-HPLC as well as an identical MS/MS fragmentation pattern as the synthetic cGAMP(3’-5’) (Supplementary Figure 6). These data strongly suggested that the cGAS derived molecule and synthetic cGAMP(3’-5’) were structural isomers with different phosphodiester linkages, being consistent with the idea of the cGAS-derived product harboring a 2’-5’ phosphodiester linkage that could isomerize into a 3’-5’ phosphodiester bond [1].

References

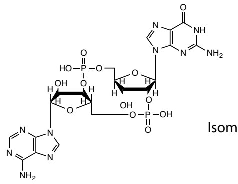
1 Seiffert, S. *et al.* Characterization of side reactions during the annealing of small interfering RNAs. *Anal Biochem* **414**, 47-57, doi:10.1016/j.ab.2011.02.040 (2011).

1 >Gp(3′-5′)Ap(3′-5′)>



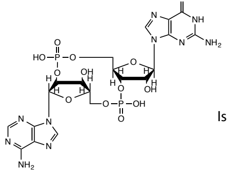
RNase T1	ApGp
RNase T2	Ap, Gp
S1 Nuclease	pA, pG
SVPDE	None
Isomerization to cGAMP	N/A
Periodate reaction	negative

2 >Gp(2′-5′)Ap(3′-5′)>



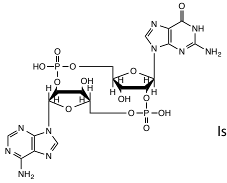
RNase T1	None
RNase T2	GpAp
S1 Nuclease	pGpA
SVPDE	pG, pA
Isomerization to cGAMP	yes
Periodate reaction	negative

3 >Gp(3′-5′)Ap(2′-5′)>



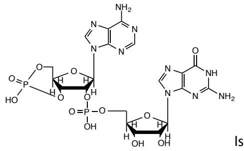
RNase T1	ApGp
RNase T2	ApGp
S1 Nuclease	pApG
SVPDE	pG, pA
Isomerization to cGAMP	yes
Periodate reaction	negative

4 >Gp(2′-5′)Ap(2′-5′)>



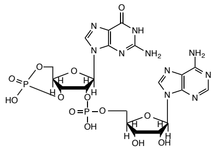
RNase T1	None
RNase T2	None
S1 Nuclease	None
SVPDE	pG, pA
Isomerization to cGAMP	yes
Periodate reaction	negative

5 cAMPp(2′-5′)G



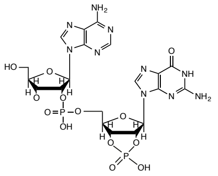
RNase T1	None
RNase T2	None
S1 Nuclease	None
SVPDE	cAMP, pG
Isomerization to cGAMP	no
Periodate reaction	positive

6 cGMPp(2′-5′)A



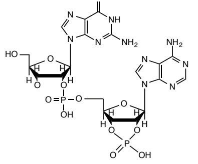
RNase T1	None
RNase T2	None
S1 Nuclease	None
SVPDE	cGMP, pA
Isomerization to cGAMP	no
Periodate reaction	positive

7 Ap(2′-5′)G>p



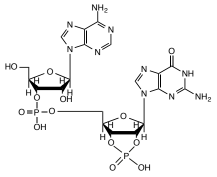
RNase T1	None
RNase T2	None
S1 Nuclease	None
SVPDE	None
Isomerization to cGAMP	no
Periodate reaction	negative

8 Gp(2′-5′)A>p



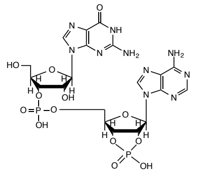
RNase T1	None
RNase T2	None
S1 Nuclease	None
SVPDE	None
Isomerization to cGAMP	no
Periodate reaction	negative

9 Ap(3′-5′)G>p



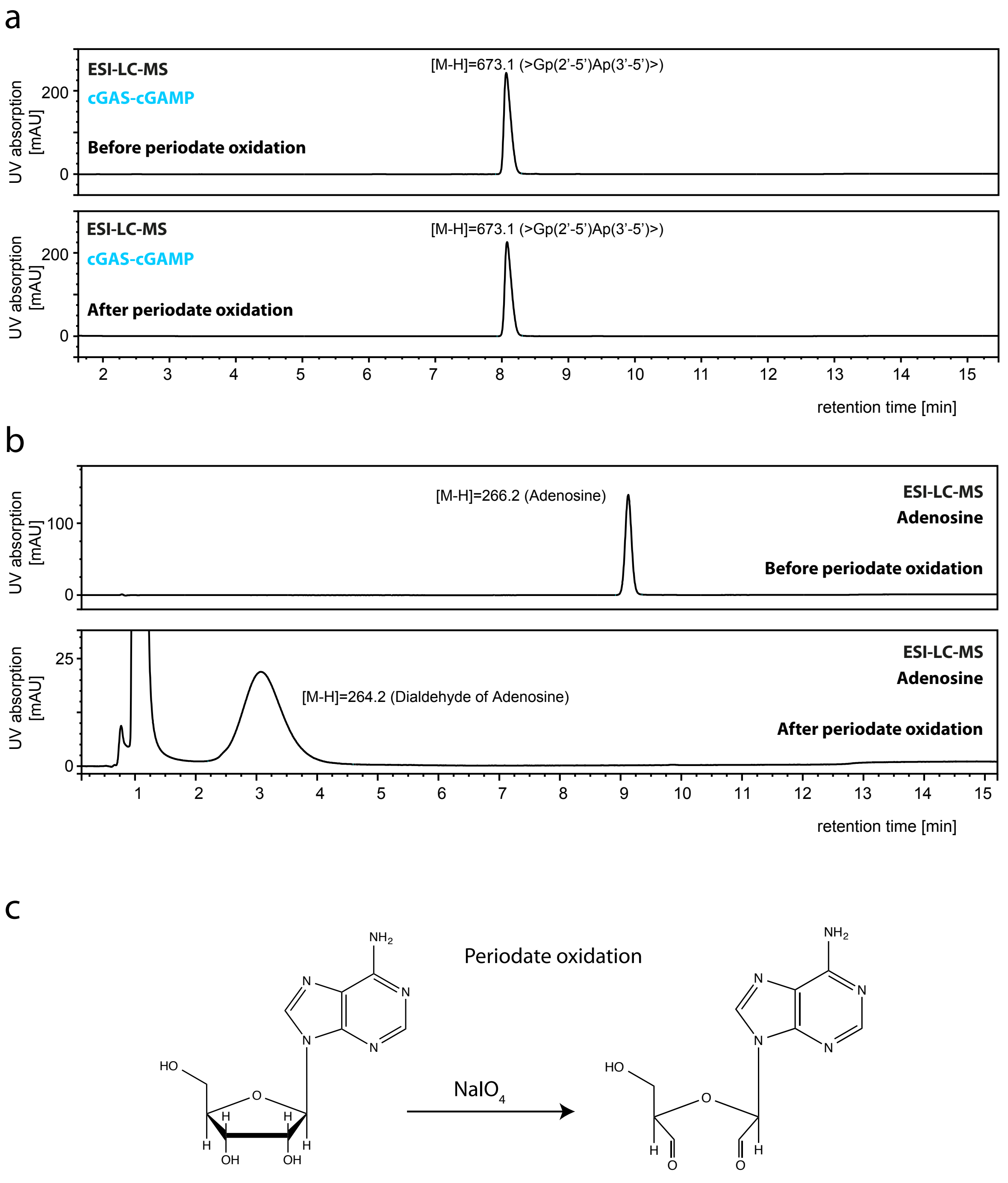
RNase T1	None
RNase T2	Ap, G>p
S1 Nuclease	A, pG>p
SVPDE	None
Isomerization to cGAMP	no
Periodate reaction	negative

10 Gp(3′-5′)A>p

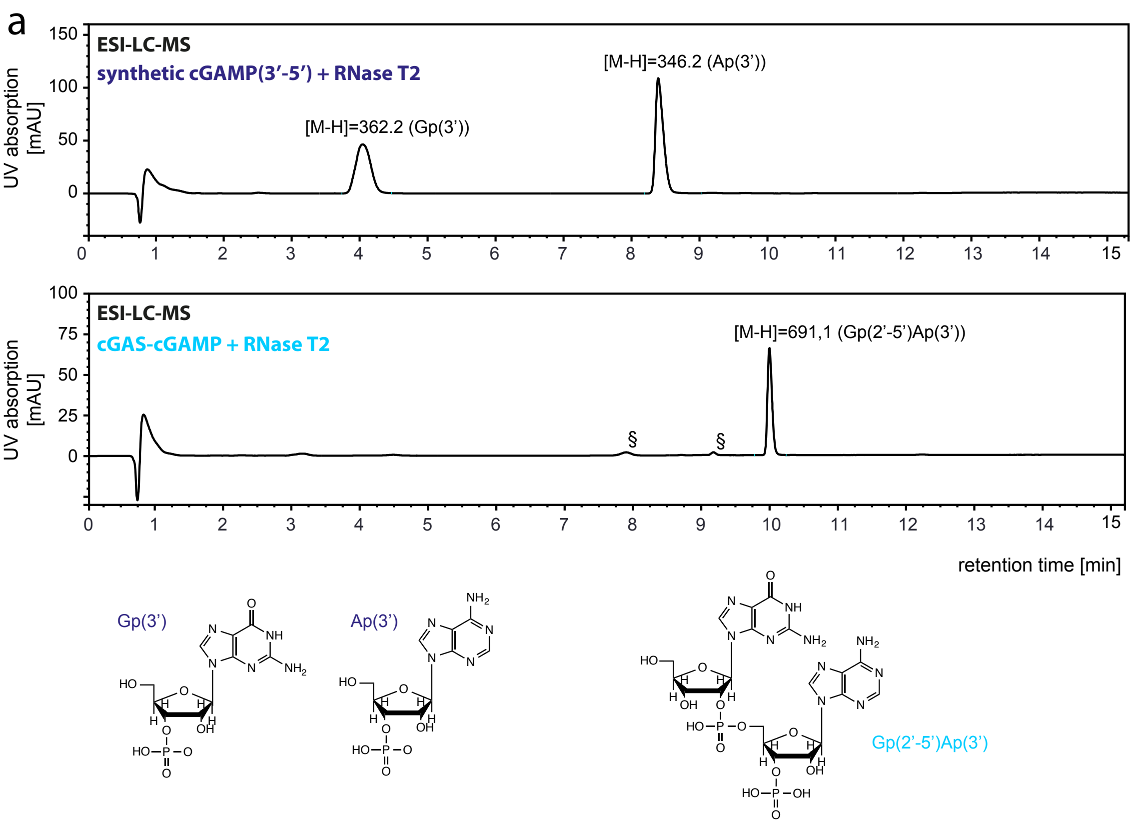


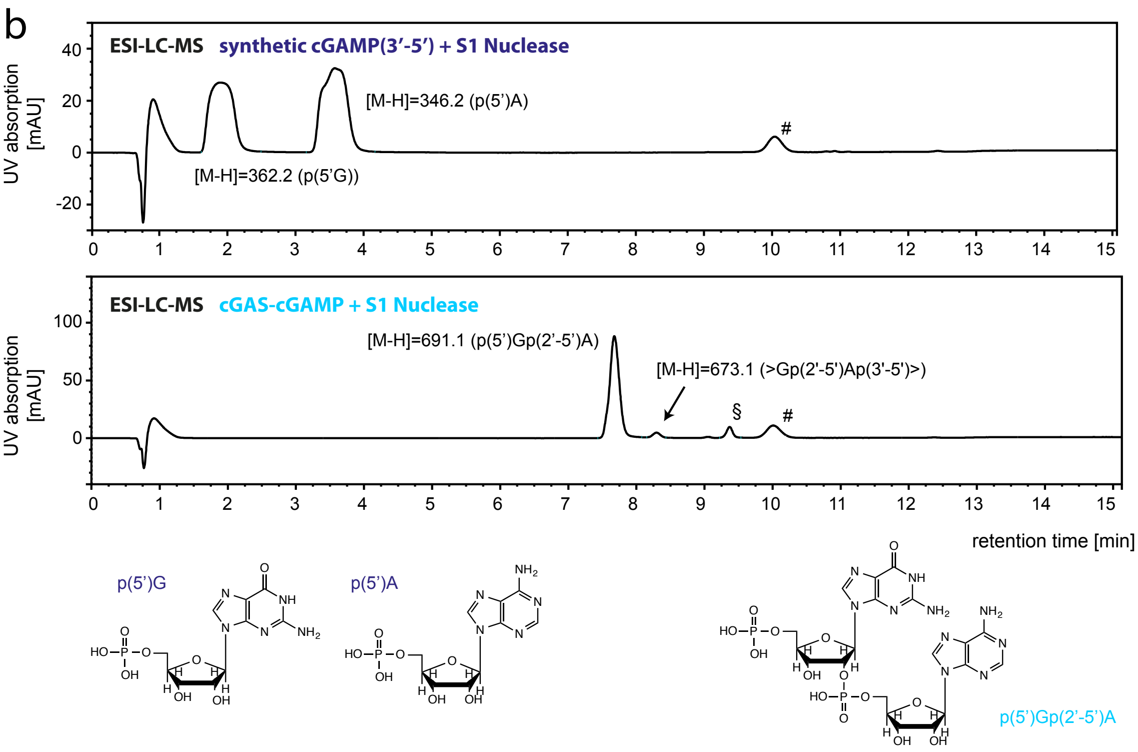
RNase T1	Gp, A>p
RNase T2	Gp, A>p
S1 Nuclease	G, pA>a
SVPDE	None
Isomerization to cGAMP	no
Periodate reaction	negative

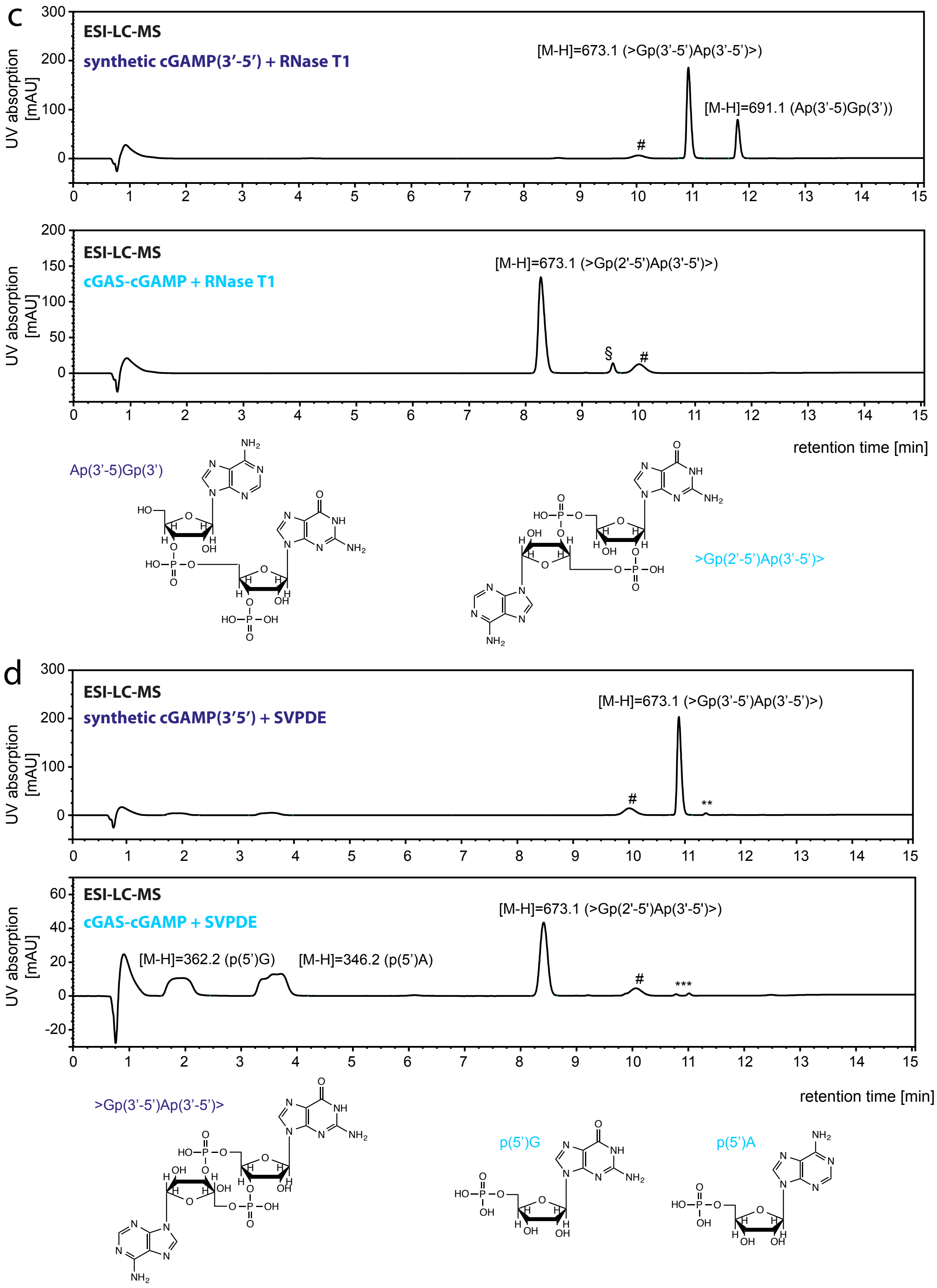
Supplementary Figure 7 | Candidate products of cGAS. A schematic overview of candidate molecules consistent with the LC-MS-determined molecular mass of the endogenous cGAS product is depicted. Expected products upon incubation with the indicated enzymes are listed. cGAMP(2’-5’) is highlighted in yellow.



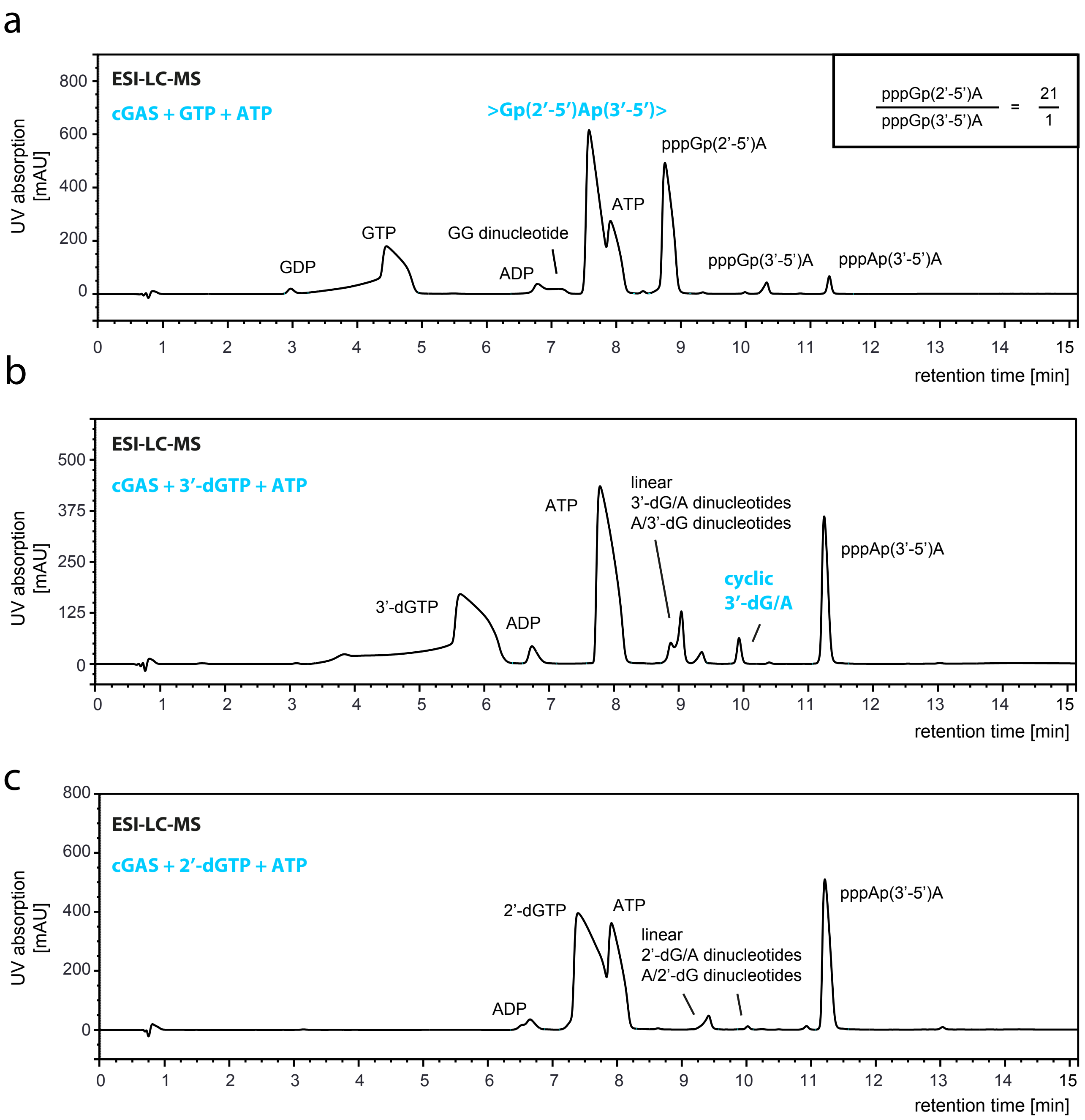
Supplementary Figure 8 | Negative periodate oxidation rules out the presence of free OH-groups at the 2´ and 3´terminal positions. (**a**) Chromatograms from in cellulo synthesized cGAS product before and after periodate oxidation are shown. (**b**) Chromatograms from adenosine before and after periodate oxidation are shown. (**c**) Schematic view of the reaction pathway of periodate oxidation is depicted. Representative data out of two independent experiments are shown.


a
ESI-LC-MS
synthetic cGAMP(3'-5') + RNase T2
[M-H]=362.2 (Gp(3'))
[M-H]=346.2 (Ap(3'))
UV absorption [mAU]
150
100
50
0
ESI-LC-MS
cGAS-cGAMP + RNase T2
[M-H]=691,1 (Gp(2'-5')Ap(3'))
§
§
100
75
50
25
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
retention time [min]
Gp(3')
Ap(3')
Gp(2'-5')Ap(3')

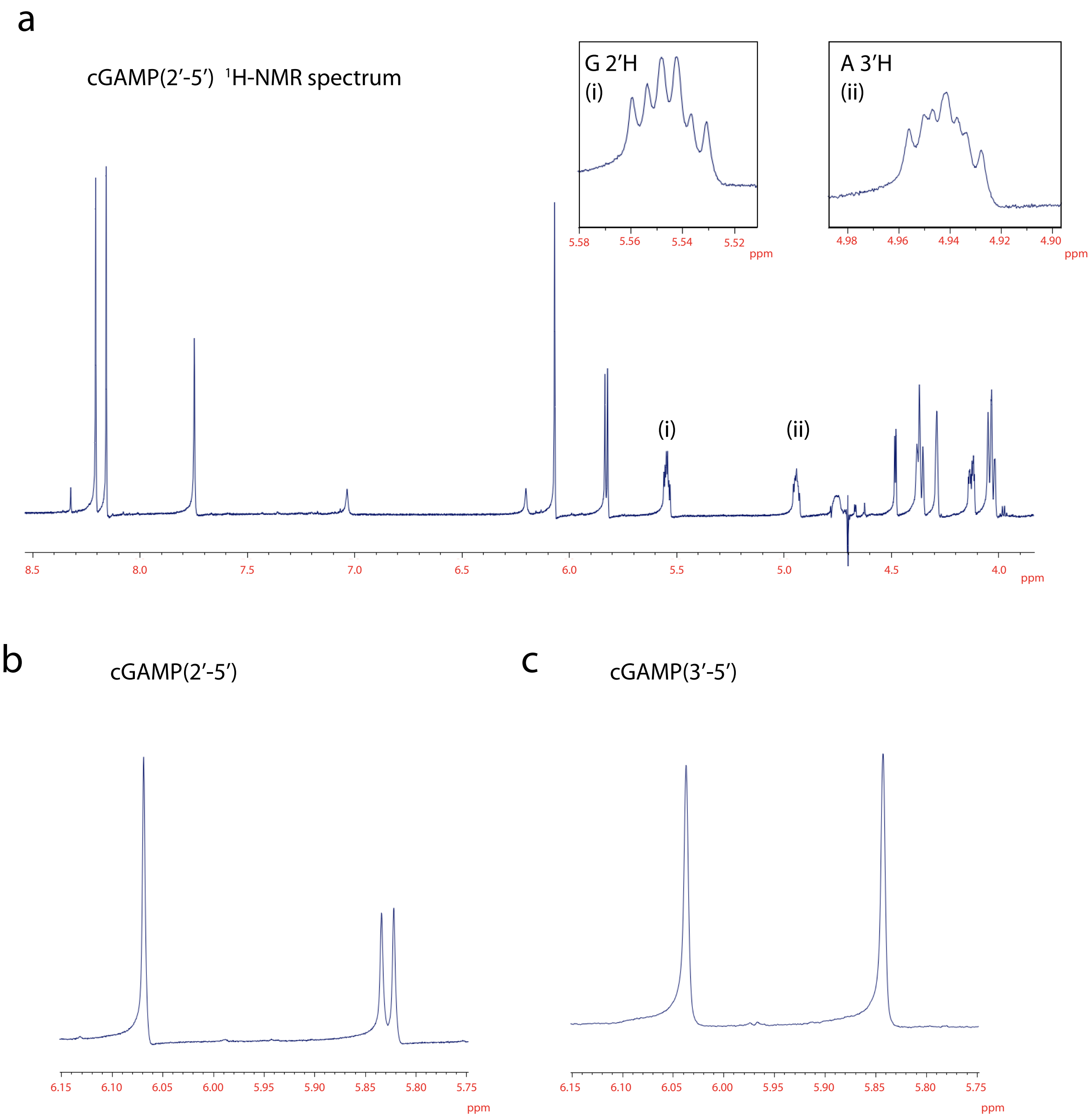

b
ESI-LC-MS synthetic cGAMP(3'-5') + S1 Nuclease
[M-H]=346.2 (p(5')A)
[M-H]=362.2 (p(5'G))
#
UV absorption [mAU]
40
20
0
-20
ESI-LC-MS cGAS-cGAMP + S1 Nuclease
[M-H]=691.1 (p(5')Gp(2'-5')A)
[M-H]=673.1 (>Gp(2'-5')Ap(3'-5')>)
§
#
100
50
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
retention time [min]
p(5')G
p(5')A
p(5')Gp(2'-5')A



Supplementary Figure 9 | Enzymatic assays reveal the endogenous cGAS product to be (>Gp(2'-5')Ap(3'-5')>). (**a**) Chromatograms and schematic view of the reaction products from (**a**) RNase T2, (**b**) S1 Nuclease, (**c**) RNase T1 and (**d**) SVPDE digestion of synthetic cGAMP(3'-5') (dark blue) and in cellulo synthesized cGAS product (light blue). Representative data out of two independent experiments are shown. Peaks marked with #, §, ** and *** represent phenol contamination, not assignable peaks (no ESI-MS signal), chemical synthesis by-products and isomerization products, respectively.



Supplementary Figure 10 | GTP, 3'-dGTP and 2'-dGTP lead to distinct cGAS reaction product profiles. (**a-c**) Chromatograms of in vitro cGAS enzymatic reaction products with ATP and GTP (**a**), 3'-dGTP (**b**) or 2'-dGTP (**c**) are shown. The assignment of peaks to mono- and dinucleotide species was assessed by LC-MS. In case of arguable isomers, characterization of peaks was complemented with enzymatic digestion analysis for the crucial components. The panel in the upper right of the cGAS/ATP/GTP reaction (**a**) illustrates the relative ratio of the abundance of pppGp(2'-5')A vs. pppGp(3'-5')A. Data are representative of n=2 experiments.



Supplementary Figure 11 | Characterization of cGAMP(2’-5’) and cGAMP(3’-5’) by ^{1}H-NMR spectroscopy. (**a**) The complete ^{1}H-NMR spectrum of the in vitro reaction product of cGAS (cGAMP(2’-5’)) is shown. The inserts in the top indicate the expanded G 2´H (i) and A 3´H (ii) regions. (**b** and **c**) Close-up view of the 1´H regions of the ^{1}H-NMR spectrum of cGAMP(2'-5') (**b**) and cGAMP(3'-5') (**c**).

Supplementary Note 3 | Characterization of cGAMP(2’-5’) and cGAMP(3’-5’) by ^{1}H-NMR spectroscopy

^{1}H-NMR spectrum of the cGAS synthase reaction product (Supplementary Figure 11a) was measured from a 0.1 mM solution in D2O at 700 MHz. Singlets of the base protons indicated that no multimerization occurs under such conditions [1]. The spectral parameters were determined as follows (700 MHz; D2O): 8.2 (1H, s, A-H8); 8.16 (1H, s, A-H2); 7.75 (1H, s, G-H8); 6.07 (1H, s, A-H1´, $J_{1'2'}$ = 0 Hz); 5.83 (1H, s, G-H1´, $J_{1'2'}$ = 8.49 Hz); 5.55 (1H, m, G-H2´, $J_{1'2'}$ = 8.49 Hz, $J_{2'P}$ = 9.0 Hz, $J_{2'3'}$ = 4.06Hz); 4.94 (1H, m, A-H3´, $J_{2'3'}$ = 4.06Hz, $J_{3'P}$ = 6.7 Hz, $J_{3'4'}$ = 10 Hz); 4.75* (s); 4.48 (1H, d, A-H2´, $J_{1'2'}$ = 0 Hz , $J_{2'3'}$ = 4.06 Hz); 4.37-4.07* (unresolved m, A and G 4´ H and 5´, 5´´ H).

Comparison of the H-1´ regions of the ^{1}H-NMR spectra of >Gp(2’-5’)Ap(3’-5’)> and >Gp(3’-5’)Ap(3’-5’)> revealed that the ribose of the 2’-5’ linked guanosine preferentially occupies a 2´-endo conformation ($J_{(1'H\text{-}2'H)}$ = 8.49 Hz) [2,3] (Supplementary Figure 11b and c), whereas the singlet of the other ribose 1´H indicates the presence of an 3´-endo sugar for adenosine, similarly to the sugar pucker observed in the 3´-5´,3´-5´ cGAMP isomer [4]. The distinction between adenosine and guanosine 1´H signals is based on the fact that 1´H signals of adenosine 5´, 3´and 2´ monophosphates appear always downfield from the corresponding 1´H positions of the corresponding guanosine monophosphates, with no overlap observed between A and G series [5,6]. Of note, this trend is also valid for the recently characterized cyclic 3´-5´ GA diphosphates [4].

The significantly different $J_{1'2'}$ coupling constants for G and A allows further distinctions between some neighboring atoms the following way: Two signals in the spectrum at 5.55 ppm and at 4.94 ppm show significant downfield shift compared to the standard positions of A and G 2´or 3´ protons. These values indicate the presence of O-phosphoryl groups at the corresponding positions. Based on literature data for of G and A 2´and 3´ monophosphates [5] the most downfield 5.55 ppm peak may result from H2´ signals of 2´-GMP or 2´-AMP type environments. 2´-AMP can be eliminated based on analysis of the fine structure of the 5.55 ppm peak. This spin system could be modeled with good precision using $J_{1'2'}$ = 8.49 Hz, $J_{2'P}$ = 9.0 Hz, $J_{2'3'}$ = 4.06Hz parameters i.e. the 5.55 peak could be assigned to the guanosine having a 1´H proton with $J_{1'2'}$ = 8.49 Hz. In this calculation literature $J_{2'P}$ values for nucleoside 2´ phosphate esters were employed [7].

The 4,94 ppm multiplet could be modeled with the following parameters: $J_{2'3'}$ = 4.06Hz, $J_{3'P}$ = 6.7 Hz, $J_{3'4'}$ = 10 Hz thus excluding 2´-H of adenosine, which shows no $J_{1'2'}$ coupling thus suggesting that this proton is attached to adenosine 3´C. $J_{(3,4)}$ was derived using the ($J_{3'4'}$ + $J_{1',2'}$)= 10 Hz relationship [3].

For the residual 4 - 4.8 ppm region containing signals for 4´H, 5´CH_2 and the remaining ribose protons, partial overlap with the HDO background signal and the complex unresolved multiplets for the 5´ and 5´´ CH_2 protons allowed only a tentative assignments (These peaks are marked with *).

All in all, this partial assignment confirms that the cGAS reaction product contains 2´-phosphorylated guanosine and 3´-phosphorylated adenosine substructures, thereby clearly supporting the result of enzymatic digestion experiments indicating the presence of >Gp(2’-5’)Ap(3’-5’)> linked ring structure.

Based on these observations, a 3D structure model of cGAMP(2’-5’) was generated by using the fragment based approach of COSMOS that relies on nucleotide structure parameters from crystallography databases [8]. The structure containing 2´-endo G ribose was selected based on the observed ^{1}H-NMR parameters for the conformation of the G ribose, and the adenine and guanine bases were rotated into trans orientation with Pymol to match the previously determined structure of cyclic-diGMP bound to STING by Yin et al. (4F9G.pdb) [9].

References

1 Gentner, M., Allan, M. G., Zaehringer, F., Schirmer, T. & Grzesiek, S. Oligomer formation of the bacterial second messenger c-di-GMP: reaction rates and equilibrium constants indicate a monomeric state at physiological concentrations. *J Am Chem Soc* **134**, 1019-1029, doi:10.1021/ja207742q (2012).

2 Moreau, C. *et al.* Synthesis of cyclic adenosine 5'-diphosphate ribose analogues: a C2'endo/syn "southern" ribose conformation underlies activity at the sea urchin cADPR receptor. *Org Biomol Chem* **9**, 278-290, doi:10.1039/c0ob00396d (2011).

3 Altona, C. & Sundaralingam, M. Conformational analysis of the sugar ring in nucleosides and nucleotides. Improved method for the interpretation of proton magnetic resonance coupling constants. *J Am Chem Soc* **95**, 2333-2344 (1973).

4 Kellenberger, C. A., Wilson, S. C., Sales-Lee, J. & Hammond, M. C. RNA-Based Fluorescent Biosensors for Live Cell Imaging of Second Messengers Cyclic di-GMP and Cyclic AMP-GMP. *J Am Chem Soc*, doi:10.1021/ja311960g (2013).

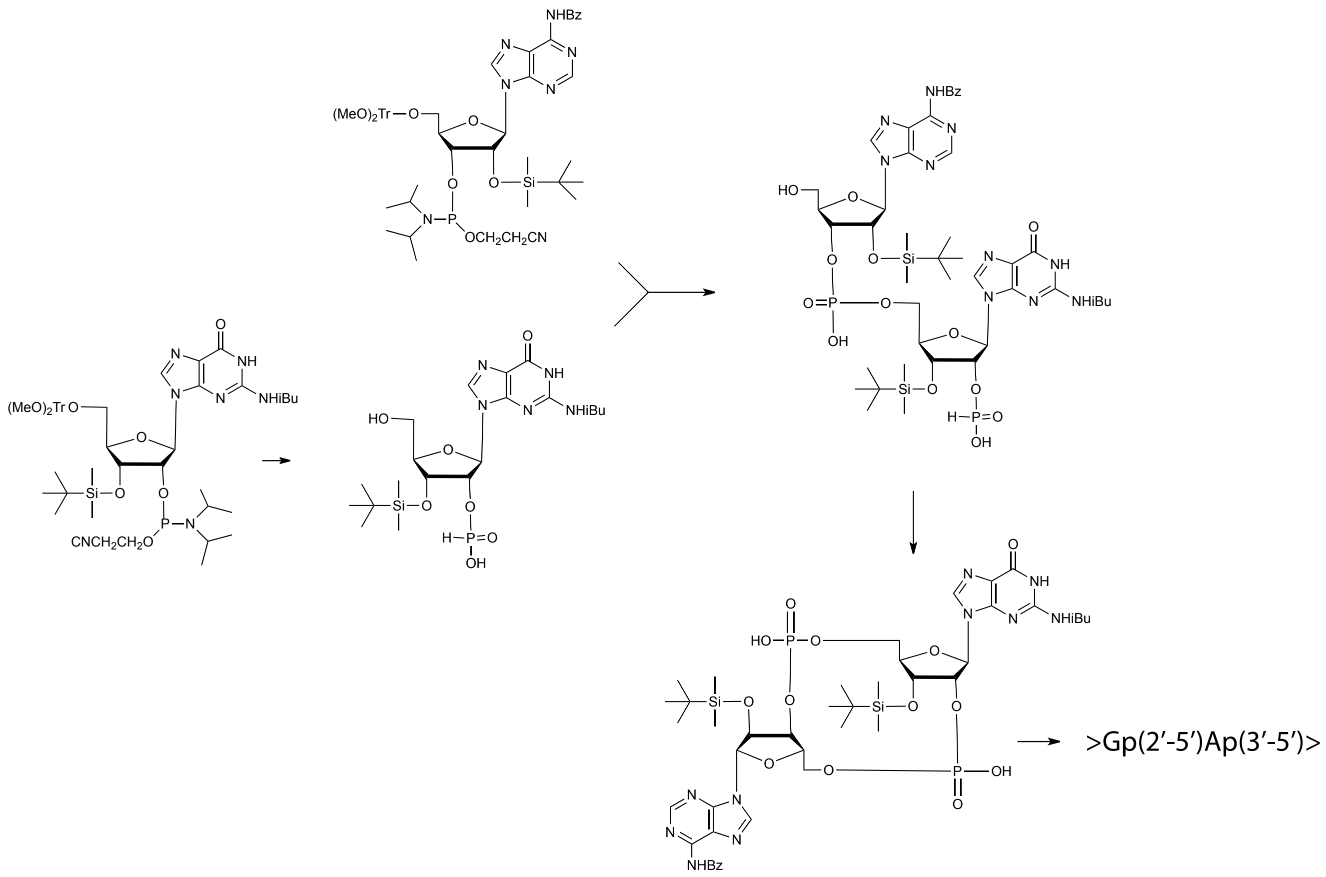
5 Davies, D. B. & Danyluk, S. S. Nuclear magnetic resonance studies of 2'- and 3'-ribonucleotide structures in solution. *Biochemistry* **14**, 543-554 (1975).

6 Davies, D. B. & Danyluk, S. S. Nuclear magnetic resonance studies of 5'-ribo- and deoxyribonucleotide structures in solution. *Biochemistry* **13**, 4417-4434 (1974).

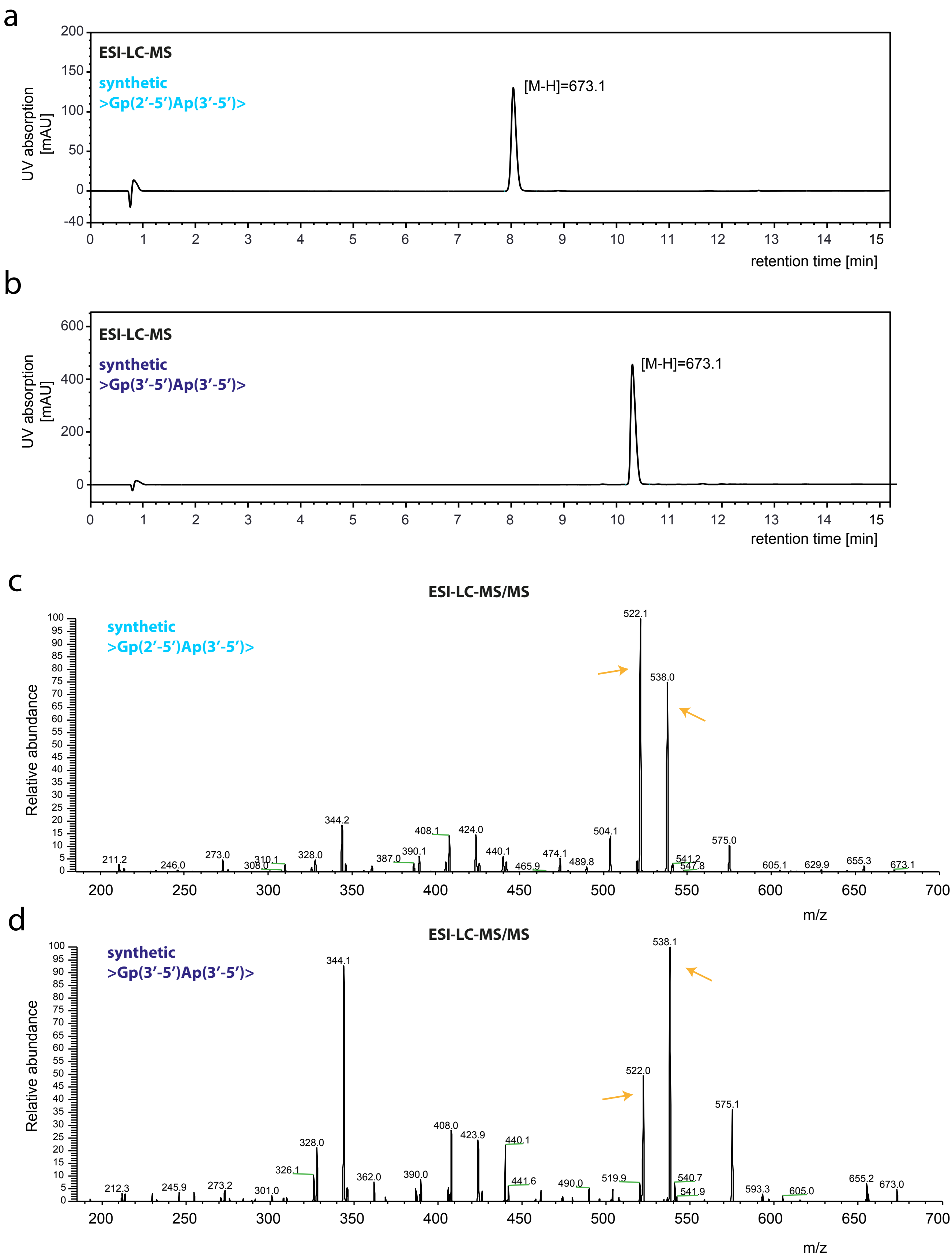
7 Sund, C., Agback, P., Koole, L. H., Sandstr√∂m, A. & Chattopadhyaya, J. Assessment of competing 2′»5′ versus 3′»5′ stackings in solution structure of branched-RNA by 1H- & 31 P-NMR spectroscopy. *Tetrahedron* **48**, 695-718, doi:10.1016/S0040-4020(01)88130-8 (1992).

8 Andronico, A., Randall, A., Benz, R. W. & Baldi, P. Data-driven high-throughput prediction of the 3-D structure of small molecules: review and progress. *J Chem Inf Model* **51**, 760-776, doi:10.1021/ci100223t (2011).

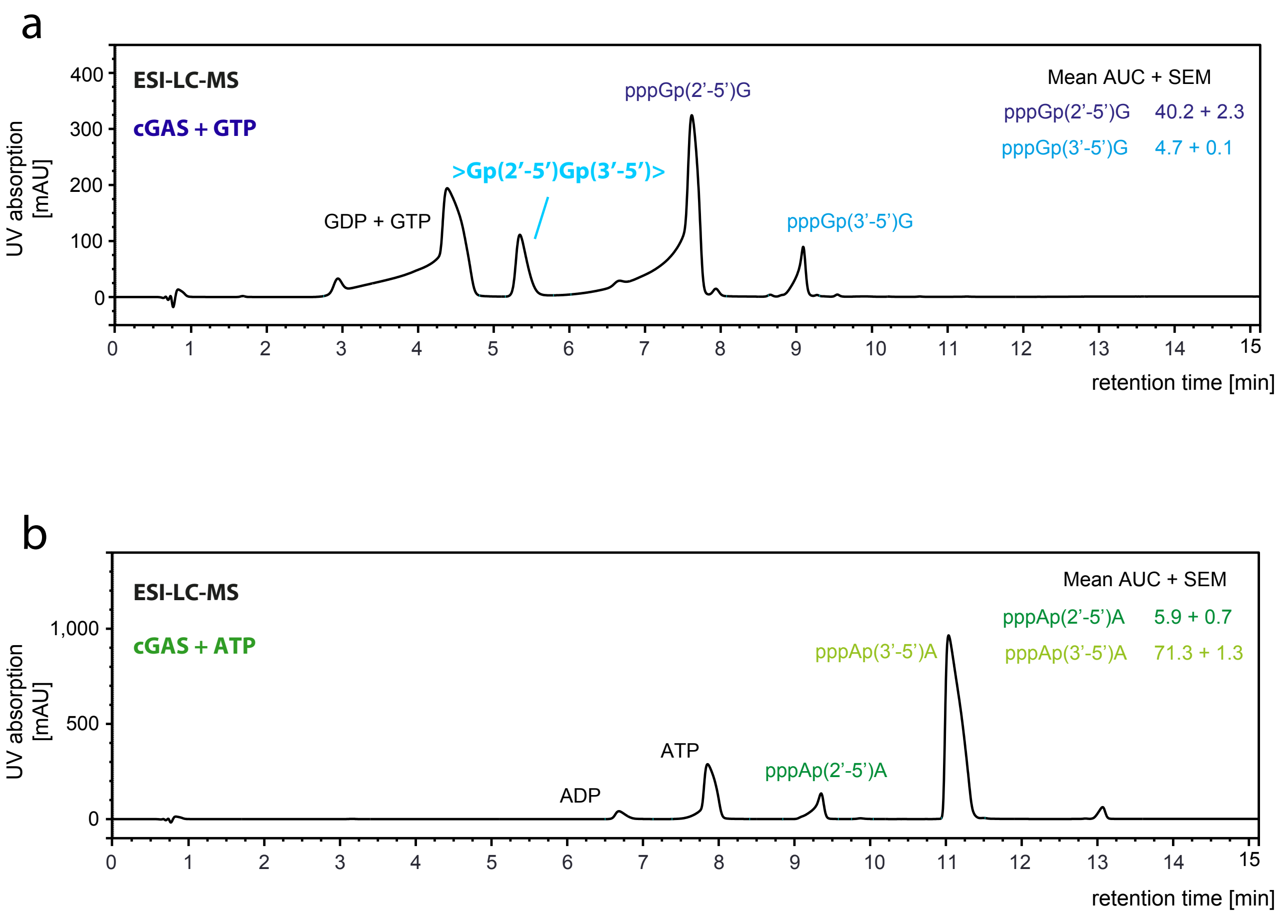
9 Yin, Q. *et al.* Cyclic di-GMP sensing via the innate immune signaling protein STING. *Mol Cell* **46**, 735-745, doi:10.1016/j.molcel.2012.05.029 (2012)



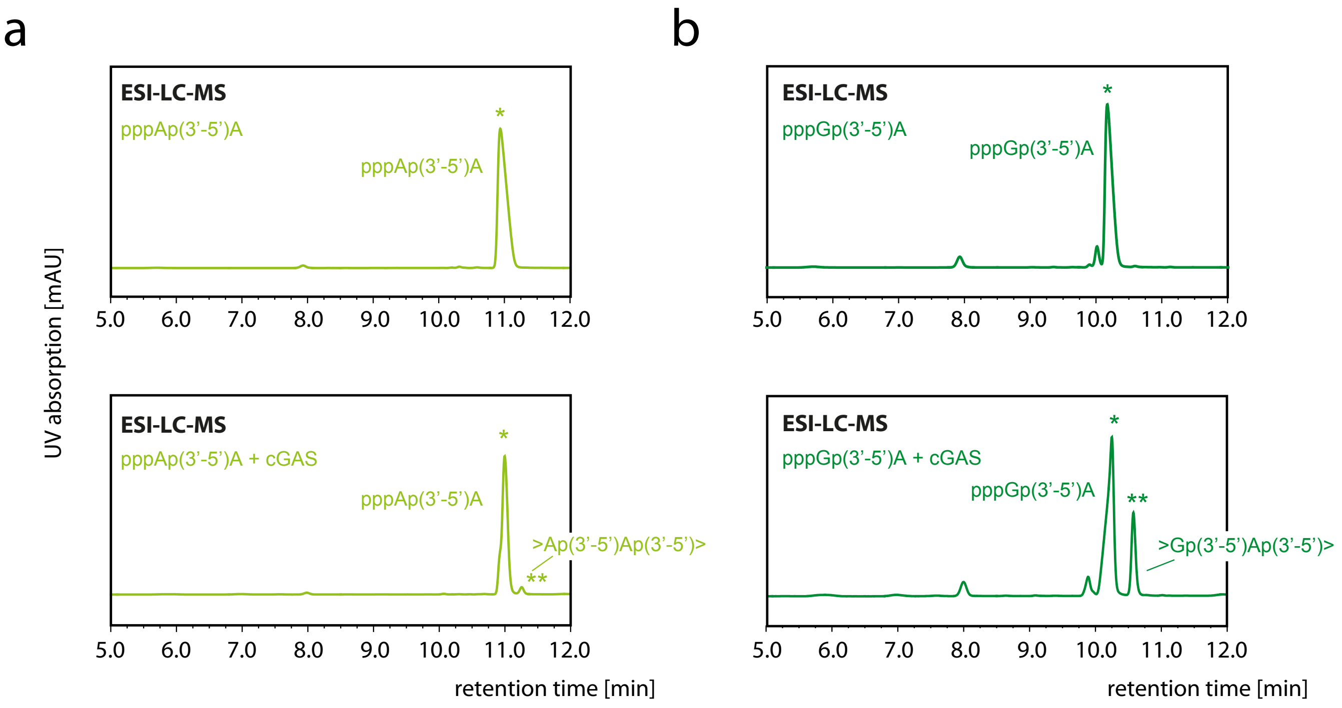
Supplementary Figure 12 | Chemical synthesis of cGAMP(2′-5′). Schematic overview of cGAMP(2’-5’) chemical synthesis. A detailed protocol is described in the material and methods section.



Supplementary Figure 13 | LC-MS analysis of chemically synthesized cGAMP(2′-5′) and cGAMP(3’-5’). (**a-b**) Chromatograms of chemically synthesized (**a**) cGAMP(2’-5’) and (**b**) cGAMP(3’-5’) are shown. (**c-d**) Tandem MS spectra of the product peaks from (**a**) and (**b**) are depicted.



Supplementary Figure 14 | LC-MS analysis of cGAS enzymatic reaction products. (**a-b**) Chromatograms of in vitro cGAS enzymatic reaction products with (**a**) ATP or (**b**) GTP as only substrates are shown. The assignment of peaks to mono- and dinucleotide species was determined by LC-MS. In case of arguable isomers, characterization of the peaks was complemented with enzymatic digestion analysis for the crucial components. The insert on the upper right depicts the relative abundance of the two observed dinucleotide species. Data are representative of n=2 experiments.



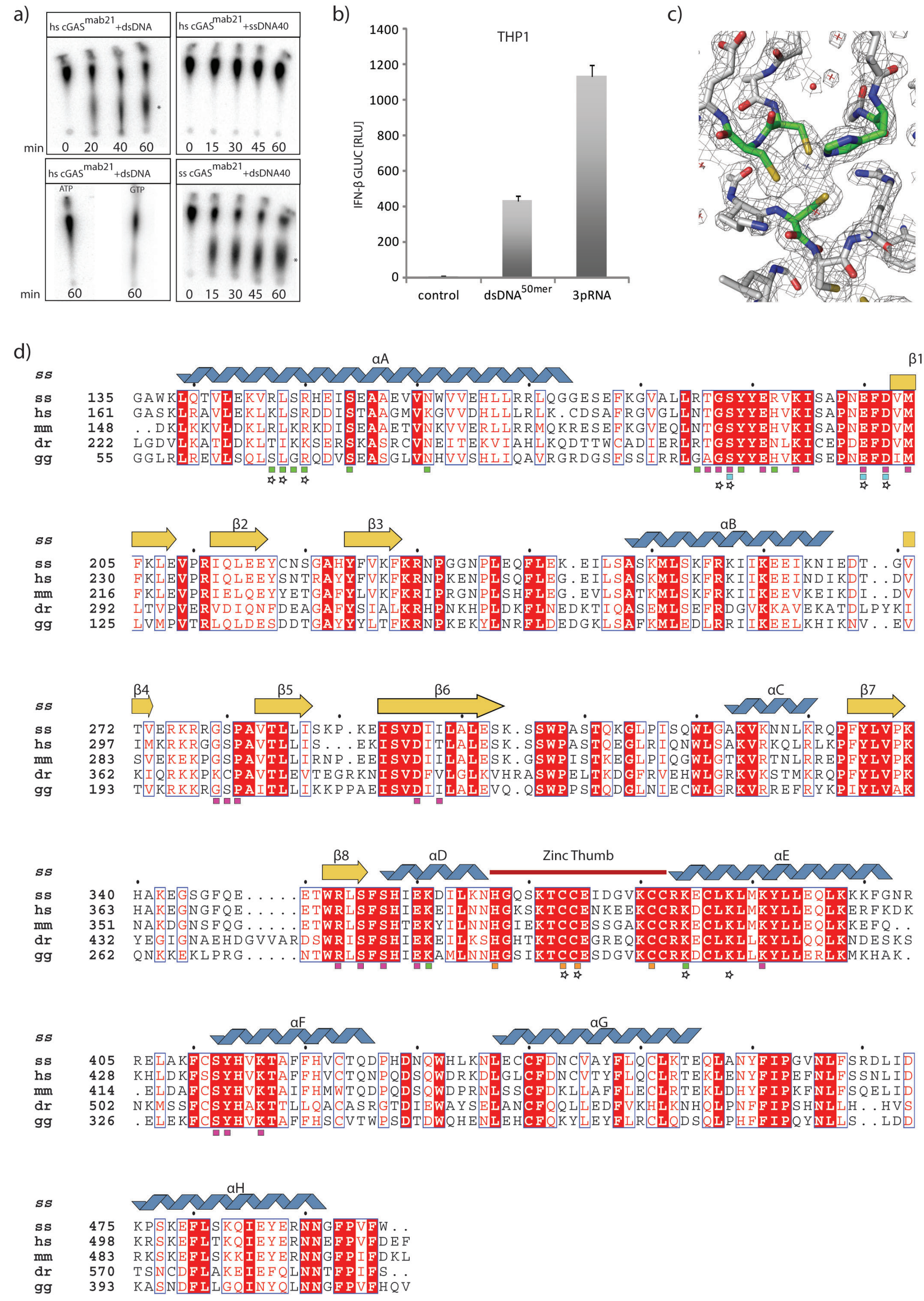
Supplementary Figure 15 | The linear 3´-5´ linked dinucleotides pppRp(3´-5´)A are scarcely or not at all processed by cGAS. (**a-b**) Chromatograms before (top) and after (bottom) incubation of pppAp(3'-5')A (**a**) and pppGp(3'-5')A (**b**) with cGAS in in vitro enzymatic reactions are shown. Asterisks indicate the position of the substrates (*) and the resulting products (**), respectively. Data are representative of n=2 experiments.

Structural mechanism of cytosolic DNA sensing by cGAS

Filiz Civril*, Tobias Deimling*, Carina C. De Oliveira Mann, Andrea Ablasser,
Manuela Moldt, Gregor Witte, Veit Hornung & Karl-Peter Hopfner

Nature 2013 June 20; **498**(7454): 332 - 7

doi:10.1038/nature12305



```
ss  135  GAWKLQTVLEKVRLSRHEISEAAEVVNWVVEHLLRRLQGGESEFKGVALLRTGSYYERVKISAPNEFDVM
hs  161  GASKLRAVLEKLKLSRDDISTAAGMVKGVVDHLLRLK.CDSAFRGVGLLNTGSYYEHVKISAPNEFDVM
mm  148  ..DKLKKVLDKLRLKRKDISEAAETVNKVVERLLRRMQKRESEFKGVEQLNTGSYYEHVKISAPNEFDVM
dr  222  LGDVLKATLDKLTIKKSERSKASRCVNEITEKVIAHLKQDTTWCADIERLRTGSYYENLKICEPDEFDVM
gg   55  GGLRLREVLSQLSLGRQDVSEASGLVNHVVSHLIQAVRGRDGSFSSIRRLGAGSYYEHVKISEPNEFDIM

ss  205  FKLEVPRIQLEEYCNSGAHYFVKFKRNPGGNPLEQFLEK.EILSASKMLSKFRKIIKEEIKNIEDT..GV
hs  230  FKLEVPRIQLEEYSNTRAYYFVKFKRNPKENPLSQFLEG.EILSASKMLSKFRKIIKEEINDIKDT..DV
mm  216  FKLEVPRIELQEYYETGAFYLVKFKRIPRGNPLSHFLEG.EVLSATKMLSKFRKIIKEEVKEIKDI..DV
dr  292  LTVPVERVDIQNFDEAGAFYSIALKRHPNKHPLDKFLNEDKTIQASEMLSEFRDGVKKAVEKATDLPYKI
gg  125  LVMPVTRLQLDESDDTGAYYYLTFKRNPKEKYLNRFLDEDGKLSAFKMLEDLRRIIKEELKHIKNV..EV

ss  272  TVERKRRGSPAVTLLISKP.KEISVDIILALESK.SSWPASTQKGLPISQWLGAKVKNNLKRQPFYLVPK
hs  297  IMKRKRGGSPAVTLLIS...EKISVDITLALESK.SSWPASTQEGLRIQNWLSAKVRKQLRLKPFYLVPK
mm  283  SVEKEKPGSPAVTLLIRNP.EEISVDIILALESK.GSWPISTKEGLPIQGWLGTKVRTNLRREPFYLVPK
dr  362  KIQRKKPKCPAVTLEVTEGRKNISVDFVLGLKVHRASWPELTKDGFRVEHWLGRKVKSTMKRQPFYLVPK
gg  193  TVKRKKRGSPAITLLIKKPPAEISVDIILALEVQ.QSWPPSTQDGLNIECWLGRKVRREFRYKPIYLVAK

ss  340  HAKEGSGFQE.....ETWRLSFSHIEKDILKNHGQSKTCCEIDGVKCCRKECLKLMKYLLEQLKKKFGNR
hs  363  HAKEGNGFQE.....ETWRLSFSHIEKEILNNHGKSKTCCENKEEKCCRKDCLKLMKYLLEQLKERFKDK
mm  351  NAKDGNSFQG.....ETWRLSFSHTEKYILNNHGIEKTCCESSGAKCCRKECLKLMKYLLEQLKKEFQ..
dr  432  YEGIGNAEHDGVVARDSWRISFSHIEKEILKSHGHTKTCCEGREQKCCRKECLKLLKYLLQQLKNDESKS
gg  262  QNKKEKLPRG.....NTWRLSFSHIEKAMLNNHGSIKTCCESDGVKCCRKDCLKLLKYLLERLKMKHAK.

ss  405  RELAKFCSYHVKTAFFHVCTQDPHDNQWHLKNLECCFDNCVAYFLQCLKTEQLANYFIPGVNLFSRDLID
hs  428  KHLDKFSSYHVKTAFFHVCTQNPQDSQWDRKDLGLCFDNCVTYFLQCLRTEKLENYFIPEFNLFSSNLID
mm  414  .ELDAFCSYHVKTAIFHMWTQDPQDSQWDPRNLSSCFDKLLAFFLECLRTEKLDHYFIPKFNLFSQELID
dr  502  NKMSSFCSYHAKTTLLQACASRGTDIEWAYSELANCFQQLLEDFVKHLKNHQLPNFFIPSHNLLH..HVS
gg  326  .ELEKFCSYHVKTAFFHSCVTWPSDTDWQHENLEHCFQKYLEYFLRCLQDSQLPHFFIPQYNLLS..LDD

ss  475  KPSKEFLSKQIEYERNNGFPVFW..
hs  498  KRSKEFLTKQIEYERNNEFPVFDEF
mm  483  RKSKEFLSKKIEYERNNGFPIFDKL
dr  570  TSNCDFLAKEIEFQLNNTFPIFS..
gg  393  KASNDFLLGQINYQLNNGFPVFHQV
```

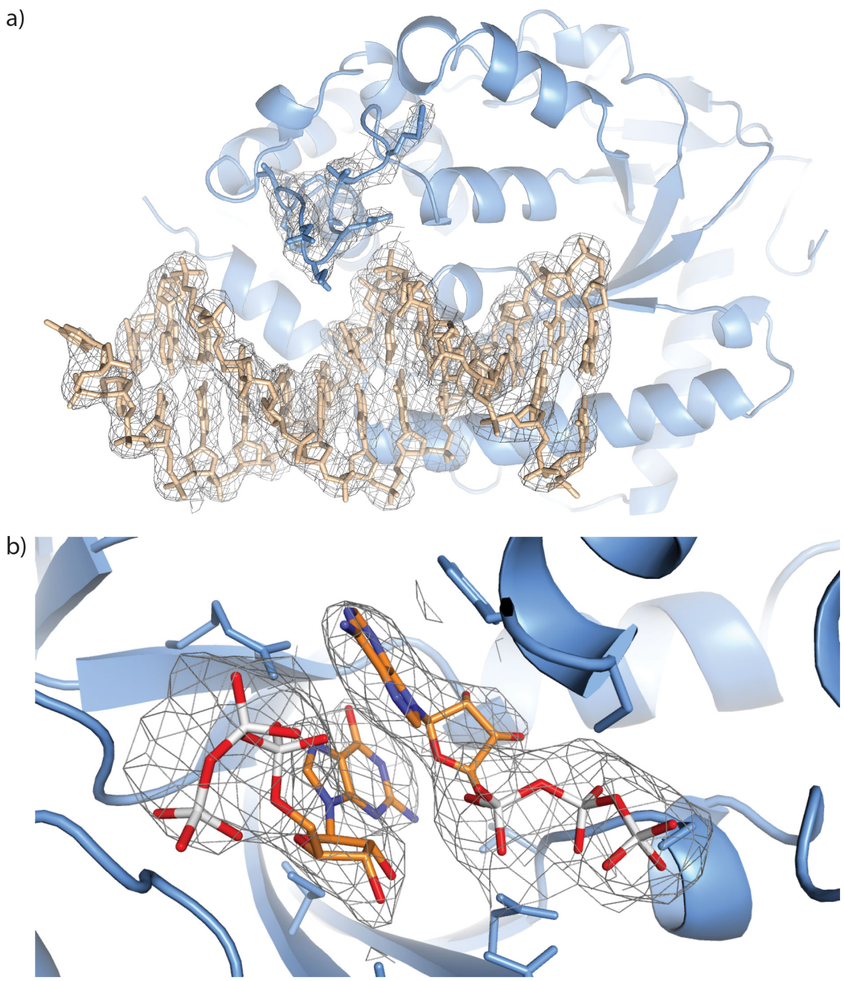
Supplemental Figure S1: Activity assays, electron density and cGAS sequence alignment

a) Activity assays with human and porcine cGAS^{mab21} in the presence of ATP, GTP and ATPα^{32}P. Upper panel left: 2µM human cGAS^{Mab21} + dsDNA (50mer), right: 1µM human cGAS^{Mab21} + 0.5µM ssDNA (40mer). Lower panel left: 2µM human cGAS^{Mab21} + 3µM dsDNA (50mer) using only ATP/ATPα^{32}P and only GTP/GTPα^{32}P, respectively; right: 1µM porcine cGAS ^{Mab21} + 0.5µM dsDNA (40mer). The reactions were stopped at indicated time points by plotting on TLC plates. Both human and porcine cGAS^{Mab21} are activated by dsDNA while ssDNA and single nucleotides fail to induce activity.

b) IFN-β stimulation in THP1 cells by dsDNA^{50mer} or with 5'triphosphate dsRNA (3pRNA). 200ng of indicated ligand was transfected to THP1 cells along with IFN-β promoter reporter plasmid pIFN-β-GLUC. Luciferase activity is plotted: mean ± sd (n=3). The dsDNA^{50mer} used in *in vitro* assays induces interferon production in THP1 cells. 3pRNA, which induces the RIG-I pathway, is used as positive control. The negative control is without ligand.

c) $2F_o$-F_c electron density overlaid with the final model around the thumb (carbons green). The $2F_o$-F_c map is contoured at 2σ.

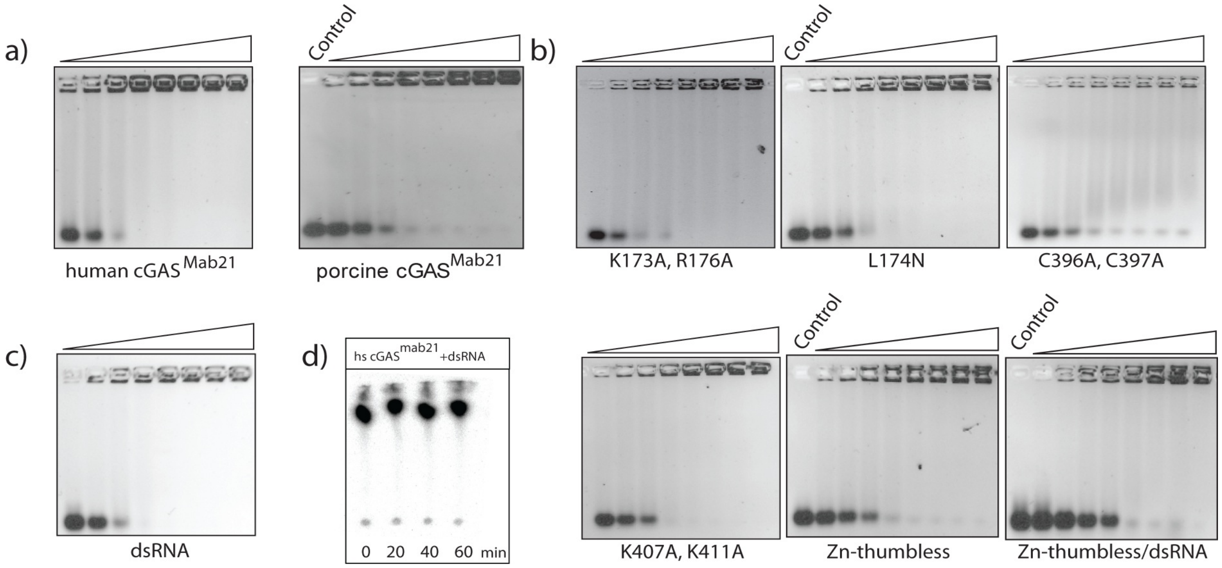
d) Structure based alignment of selected cGAS^{Mab21} sequences (abbreviations: *Sus scrofa*: **ss**, *Homo sapiens*: **hs**, *Mus musculus*: **mm**, *Danio rerio*: **dr**, *Gallus gallus*: **gg**) with highlighted conserved residues and annotated motifs. The secondary structural elements are shown on top of the alignment for the porcine homolog, color coding is analogous to Fig. 1. The squares mark residue contacts: green => DNA, magenta => GTP/ATP, cyan => Mg^{2+}, orange => Zn^{2+}. Stars denote residues that are mutated in this study.



Supplemental Figure S2: Electron densities

a) Ribbon model of $cGAS^{Mab21(td)}$ (blue) with bound dsDNA (beige stick representation). $2F_o$-F_c electron density for the DNA and the Zn-thumb residues are shown at a contour level of 1σ.

b) $2F_o$-F_c electron density around the two nucleotides in the active site (contour level of 1σ).



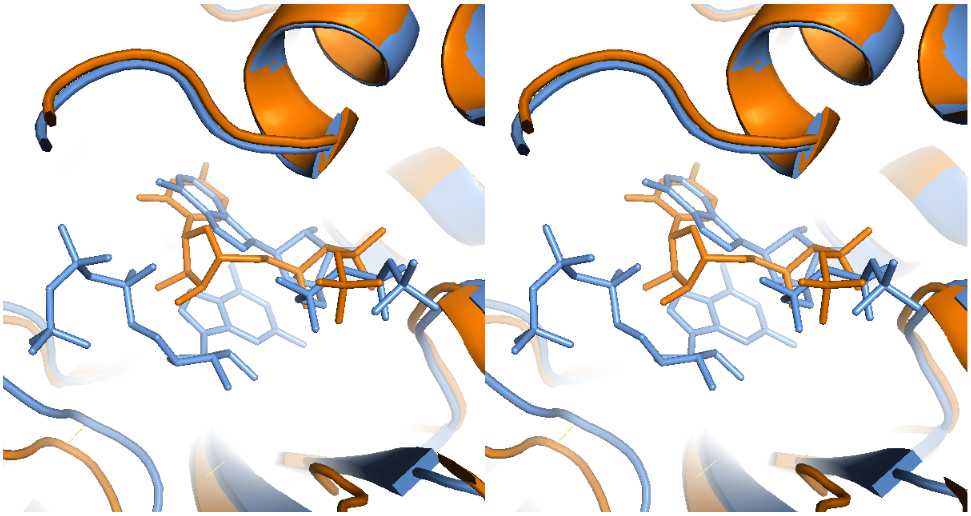
Supplementary Figure S3: Electrophoretic mobility shift and activity assays

a) Electrophoretic mobility shift analysis of dsDNA^{50mer} (0.2 µM) with human or porcine cGAS^{Mab21} (protein concentrations are 0.5, 0.75, 1.00, 1.25, 1.50, 1.75, 2.00, 2.50 µM (triangle)). Control: without protein.

b) Like a), but with human cGAS^{Mab21} mutants.

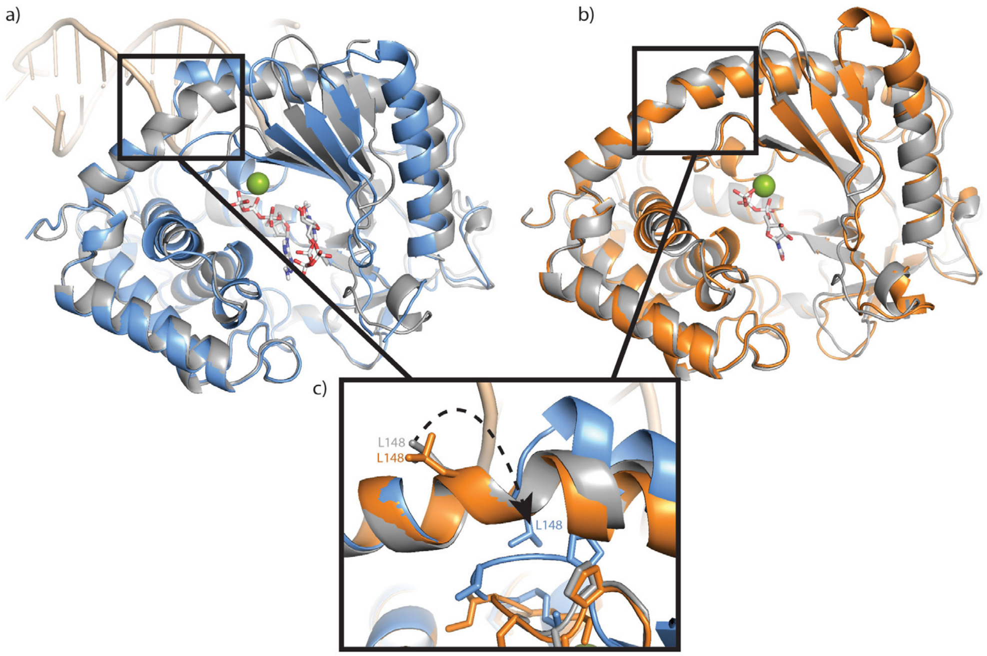
c) Like a) but binding of human cGAS^{Mab21} to dsRNA^{50mer}.

d) Activity assay with 2µM human cGAS^{Mab21} + 3µM dsRNA^{50mer}.



Supplementary Figure S4: Comparison of UTP- and ATP/GTP-bound structure of cGAS^{Mab21}

The stereo figure shows the active site of the superimposed structures of the UTP-bound state of cGAS^{Mab21} (orange) and the cGAS$^{Mab21(td)}$:DNA:GTP:ATP complex (blue), respectively. The ribose moiety is flipped in UTP as compared to ATP.

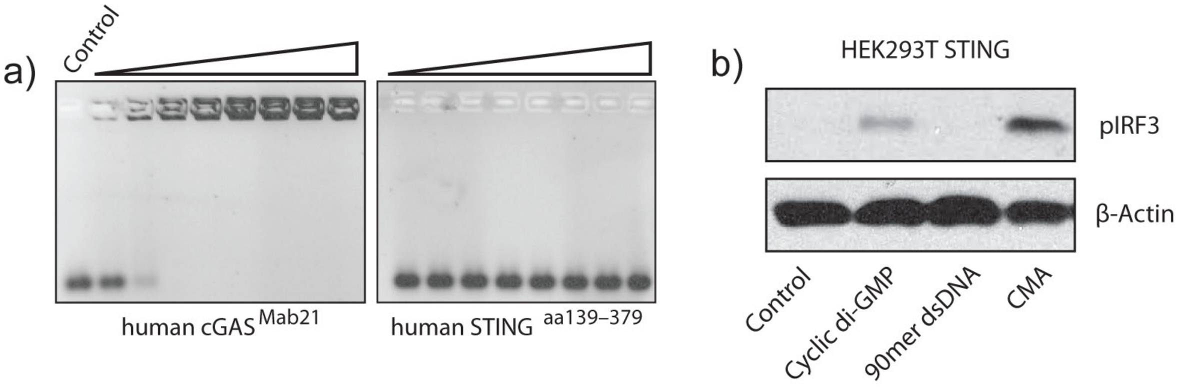


Supplementary Figure S5: Comparison of cGAS$^{\text{Mab21}}$ structures

a) Superposition of the DNA-bound (blue) and apo (grey) cGAS$^{\text{Mab21}}$ structures.

b) Superposition of apo (grey) cGAS$^{\text{Mab21}}$ and UTP-bound (orange) cGAS$^{\text{Mab21}}$.

c) The Leu148 flip upon DNA binding. In the absence of DNA (orange UTP-bound and grey apo structures) Leu148 is solvent exposed. In the DNA binding conformation (blue), Leu148 flips and helps stabilize the nucleotide-binding loop.



Supplementary Figure S6: Comparison of STING and cGAS DNA binding and STING in vivo activity.

a) Electrophoretic mobility shift analysis of dsDNA (0.2 µM) with human cGAS^{Mab21} (left) and human STING$^{aa139\text{-}379}$ (right) (protein concentrations are 0.5, 1.0, 2.0, 3.0, 5.0, 7.0, 10, 15 µM, control: without protein). While human cGAS^{Mab21} readily binds dsDNA at a concentration as low as 0.5 µM, human STING$^{aa139\text{-}379}$ fails binding even at concentration as high as 15 µM.

b) Western blot analysis of IRF3 phosphorylation which indicates interferon stimulation. The HEK293T cells stably expressing STING were induced with cyclic-di-GMP, dsDNA^{90mer} and 10-carboxymethyl-9-acridanone (CMA). The proven STING ligands cyclic-di-GMP and CMA induce IRF3 phosphorylation while DNA^{90mer} fails, suggesting requirement of a sensor that detects DNA upstream of STING. β-actin was used as loading control.

Supplemental Table S1 Data collection and refinement statistics

	SeMet	Apo	+UTP	DNA-GTP-ATP complex
Data collection				
Space group	$C222_1$	$C222_1$	$P2_12_12_1$	$C222_1$
Cell dimensions				
a, *b*, *c* (Å)	47.5, 119.9, 140.9	47.4, 118.0, 142.6	80.6, 97.7, 107.0	86.2, 111.7, 117.6
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	47.1 (2.5)*	47.5 (2.0)*	46.9 (2.28)*	999.0 (3.1)*
R_{merge}	9.0 (44.7)*	4.7 (57.5)*	5.8 (59.1)*	9.0 (90.7)*
$I / \sigma I$	13.8 (3.3)*	20.9 (2.6)*	12.0 (2.0)*	17.0 (1.9)*
Completeness (%)	98.7 (92.0)*	99.6 (98.4)*	95.9 (93.3)*	98.6 (91.7)*
Redundancy	6.8 (6.4)*	6.4 (5.9)*	2.3 (2.3)*	6.9 (6.6)*
Wavelength (Å)	0.97961	1.00665	0.97934	0.97626
Refinement				
Resolution (Å)		45.5 - 2.0	44.6 - 2.27	68.3 - 3.08
No. reflections		51611	72224	10739
R_{work} / R_{free} (%)		18.6 / 21.3	17.0 / 20.4	25.5 / 25.8
No. atoms				
Protein		2878	5969	2899
DNA		-	-	555
Ligand/ion		54	62	73
Water		132	306	2
B-factors				
Protein		52.2	38.0	101.4
DNA		-	-	113.6
Ligand/ion		65.6	41.5	129.5
Water		45.1	38.2	95.59
R.m.s. deviations				
Bond lengths (Å)		0.008	0.009	0.004
Bond angles (°)		1.07	1.19	0.888
Ramachandran (%)				
Favored		97.4	98.6	96.0
Allowed		2.3	1.4	4.0
Outliers		0.3	0	0
PDB Accession Code		4JLX	4JLZ	4KB6

*Values in parentheses are for the highest-resolution shell.

Crystal and solution structure of RIG-I SF2 domain

Tobias Deimling, Sheng Cui, Katja Lammens, Karl-Peter Hopfner & Gregor Witte

Acta Crystallogr F Struct Biol Commun. 2014 Aug 1; **70**(Pt 8): 1027 - 31

Acta Crystallographica Section F

Volume 70 (2014)

Supporting information for article:

Crystal and solution structure of the human RIG-I SF2 domain

Tobias Deimling, Sheng Cui, Katja Lammens, Karl-Peter Hopfner and Gregor Witte

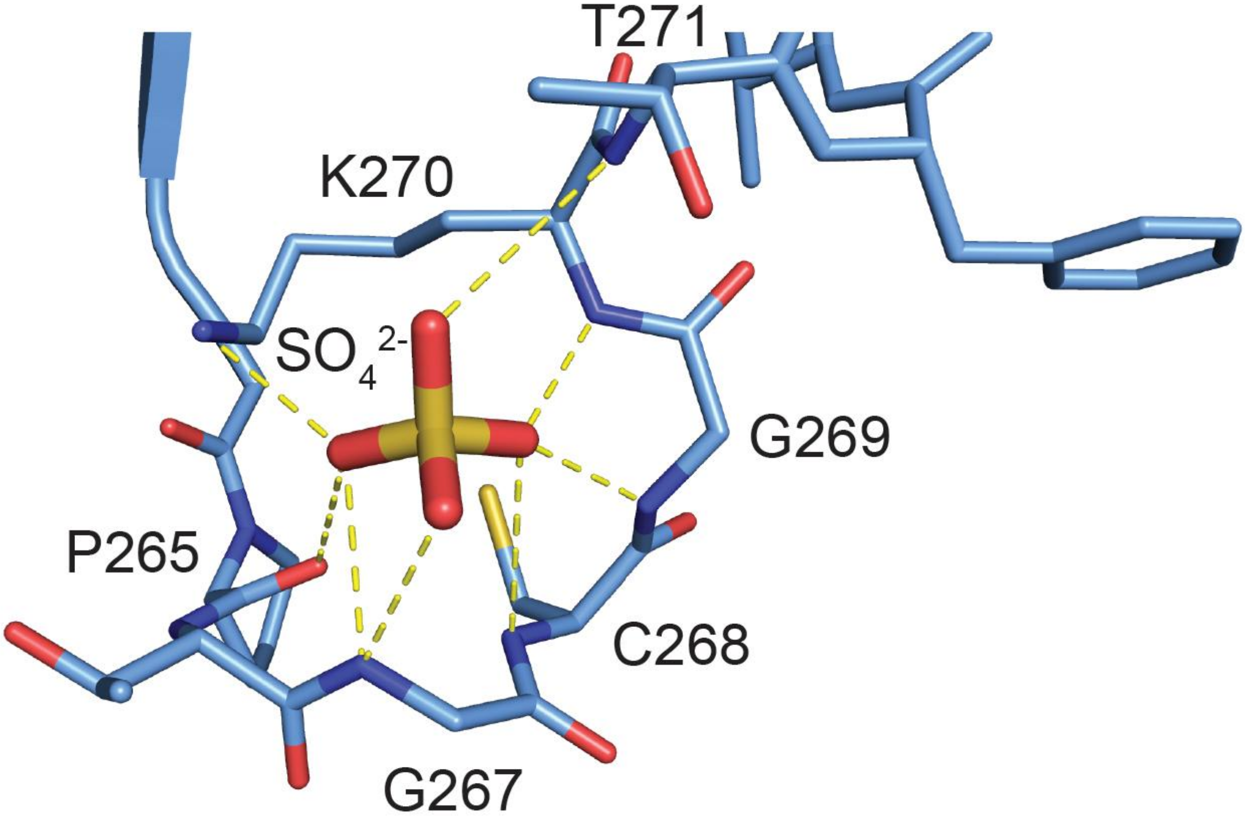


Figure S1 Coordination of the sulfate ion in the ATP-binding loop of RIG-I SF2.

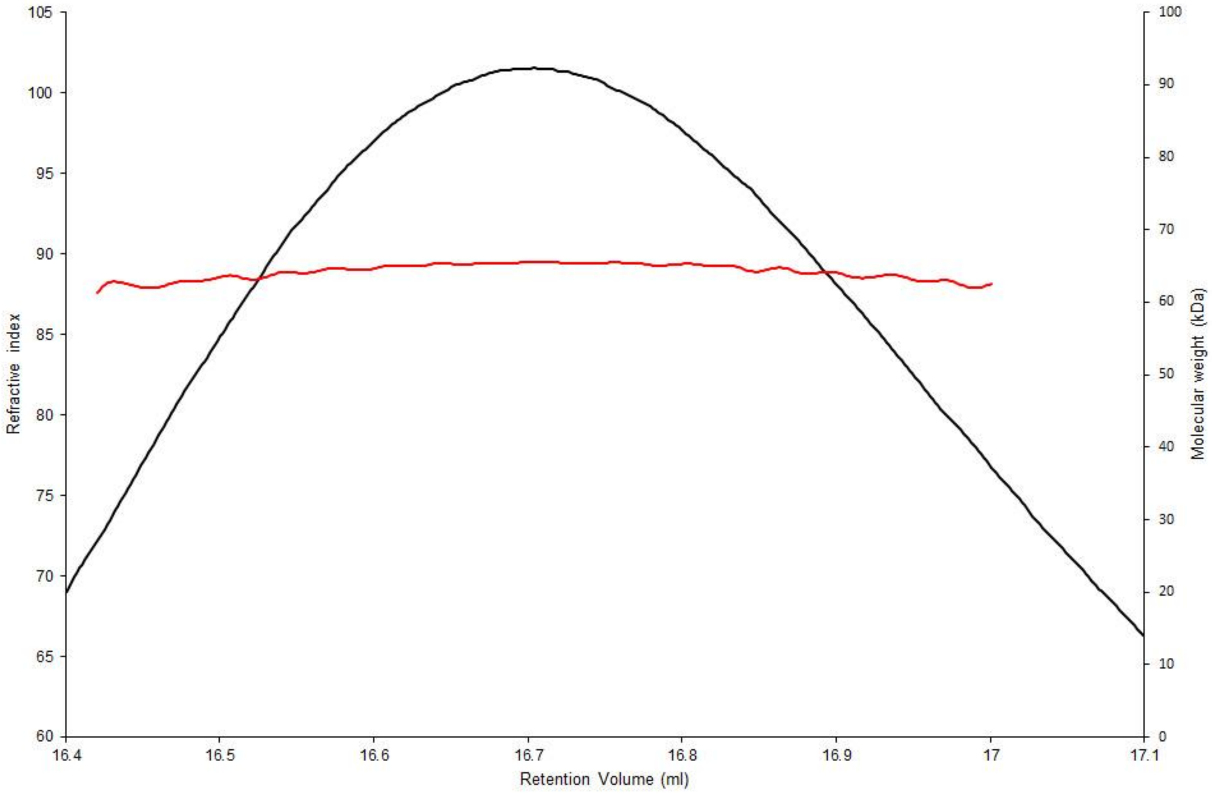


Figure S2 Size-exclusion coupled static light-scattering (SEC-RALS) data of RIG-I SF2. The black curve shows the refractive index (left axis) vs. retention volume of the RIG-I SF2 peak with the corresponding molecular weight (red curve, right axis) of a 3mg/ml sample loaded on the size exclusion column. The molecular weight is constant over the whole range of the peak indicating a very homogenous sample (polydispersity M_w/M_z= 1.0). The average molecular weight determined by SEC-RALS is $M_w^{SEC\text{-}RALS}$ = 64.2 kDa and thus in very good agreement with monomeric RIG-I SF2 ($M_w^{theoretical}$ =66.0 kDa).

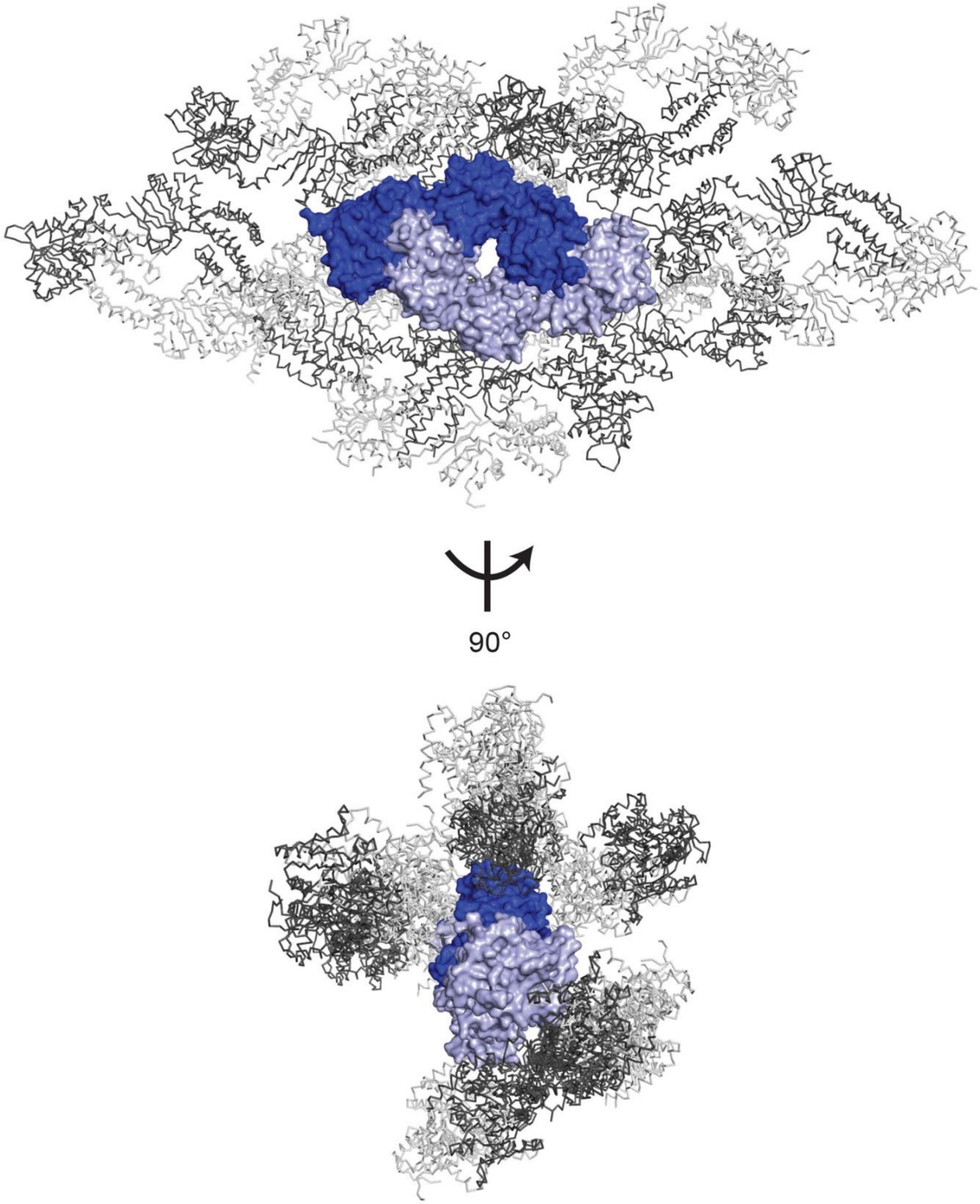


Figure S3 Illustration of the crystal packing of the RIG-I SF2 reported in this study. The crystallographic dimer (shown as surface representation in light and dark blue) and its surrounding symmetry mates shown as grey ribbons. In addition to the stabilizing interactions within the dimer, the symmetry mates influence the overall shape by extensive crystal contacts from top and bottom.