

INNATE IMMUNE SENSING AND SIGNALING OF CYTOSOLIC DNA

APPROVED BY SUPERVISORY COMMITTEE

Dr. Zhijian ‘James’ Chen

Dr. Eric N. Olson

Dr. Melanie H. Cobb

Dr. Joshua T. Mendell (Chair)

DEDICATION

I would like to thank the members of my Graduate Committee, my significant other and so
on and so forth.

INNATE IMMUNE SENSING AND SIGNALING OF CYTOSOLIC DNA

by

JIAXI WU

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

May, 2015

Copyright

by

JIAXI WU, 2015

All Rights Reserved

INNATE IMMUNE SENSING AND SIGNALING OF CYTOSOLIC DNA

Publication No. _____

JIAXI WU, Ph.D

The University of Texas Southwestern Medical Center at Dallas, 2015

Supervising Professor: Dr. Zhijian ‘James’ Chen

In eukaryotic cells, DNA is normally confined within the nucleus and mitochondria. DNA exposed in the cytosol is a danger signal that warns the host of invading microbial pathogens and triggers innate immune responses including the production of type-I interferons (IFNs). Endogenous DNA that is inappropriately cleared can also accumulate in cytosol and drive pathological inflammation and autoimmune diseases such as systemic lupus erythematosus (SLE). It is well known that cytosolic DNA induces IFNs through the STING-TBK1-IRF3 axis. However, how DNA is sensed in the cytosol and how this sensing event leads to the activation of STING remains elusive.

Using a cell-free complementation assay, we identified cyclic GMP-AMP (cGAMP), as a novel eukaryotic second messenger generated by DNA stimulated or DNA virus infected cells. cGAMP contains a unique phosphodiester linkage combination (both 2'-5' and 3'-5'), for which we referred to it as 2'3'-cGAMP. 2'3'-cGAMP bound to STING with nanomolar affinity and induced a dramatic conformational change that led to its activation. Through biochemical purification and quantitative mass spectrometry, we identified the enzyme that synthesizes cGAMP in a DNA dependent manner. This enzyme, which we named cyclic GMP-AMP synthase (cGAS), turned out to be the long sought-after cytosolic DNA sensor. Structural and functional studies revealed that cGAS is activated by DNA-induced dimerization. We further generated and characterized a cGAS knockout mouse strain, which failed to produce IFNs and other cytokines in response to DNA stimulation and was more vulnerable to lethal infection by DNA viruses.

Together, these results not only elucidate the mechanism of cytosolic DNA sensing, but also uncover a novel second messenger-mediated signaling mechanism in innate immunity.

TABLE OF CONTENTS

ABSTRACT.....	v
CHAPTER I: INTRODUCTION	1
CHAPTER II: RESULTS	11
CHAPTER III: DISCUSSION.....	54
CHAPTER IV: METHODOLOGY	62
CHAPTER V: BIBLIOGRAPHY	79

PRIOR PUBLICATIONS

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: 826-30

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: 786-91

Wu J & Chen ZJ. 2014. Innate Immune Sensing and Signaling of Cytosolic Nucleic Acids. *Annual Review of Immunology*, 32: 461-88

Li XD*, **Wu J***, Gao D, Wang H, Sun L, Chen ZJ. 2013. Pivotal Roles of cGAS-cGAMP Signaling in Antiviral Defense and Immune Adjuvant Effects. *Science*, 341: 1390-94

Zhang X*, Shi H*, **Wu J***, Zhang X, Sun L, Chen C, Chen ZJ. 2013. Cyclic GMP-AMP Containing Mixed Phosphodiester Linkages Is An Endogenous High-Affinity Ligand for STING. *Molecular Cell* 51:226-35

Zhang X, **Wu J**, Du F, Xu H, Sun L, Chen Z, Brautigam CA, Zhang X, Chen ZJ. 2014. The Cytosolic DNA Sensor cGAS Forms an Oligomeric Complex with DNA and Undergoes Switch-like Conformational Changes in the Activation Loop. *Cell Reports*, 6: 1-10

Gao D, **Wu J**, Wu YT, Du F, Aroh C, Yan N, Sun L, Chen ZJ. 2013. Cyclic GMP-AMP Synthase Is an Innate Immune Sensor of HIV and Other Retroviruses. *Science*, 341: 903-06

Liu S, Cai X, **Wu J**, Cong Q, Chen X, Li T, Du F, Ren J, Wu Y, Grishin N, Chen ZJ. 2015. Phosphorylation of Innate Immune Adaptor Proteins MAVS, STING, and TRIF Induces IRF3 Activation. *Science*, pii: aaa2630. [Epub ahead of print]

Liu S, Chen J, Cai X, **Wu J**, Chen X, Wu Y, Sun L, Chen ZJ. 2013. MAVS recruits multiple ubiquitin E3 ligases to activate antiviral signaling cascades. *Elife* 2:e00785

Xiong Y, Zhao K, **Wu J**, Xu Z, Jin S, Zhang YQ. 2013. HDAC6 mutations rescue human tau-induced microtubule defects in Drosophila. *PNAS* 110(12):4604-09

LIST OF FIGURES

FIGURE 1	1
FIGURE 2	6
FIGURE 3	8
FIGURE 4	12
FIGURE 5	14
FIGURE 6	16
FIGURE 7	19
FIGURE 8	21
FIGURE 9	24
FIGURE 10	26
FIGURE 11	28
FIGURE 12	30
FIGURE 13	32
FIGURE 14	35
FIGURE 15	37
FIGURE 16	39
FIGURE 17	41
FIGURE 18	43
FIGURE 19	45
FIGURE 20	48
FIGURE 21	51

FIGURE 22	53
FIGURE 23	58
FIGURE 24	60

LIST OF TABLES

TABLE 1.....	49
TABLE 2	67

CHAPTER ONE

Introduction

Innate Immune Detection of Pathogens

Higher organisms are constantly facing the invasion of microorganisms. In vertebrates, two complementary systems have evolved to detect and fight against microbial pathogens: the innate and adaptive immune systems (Figure 1).

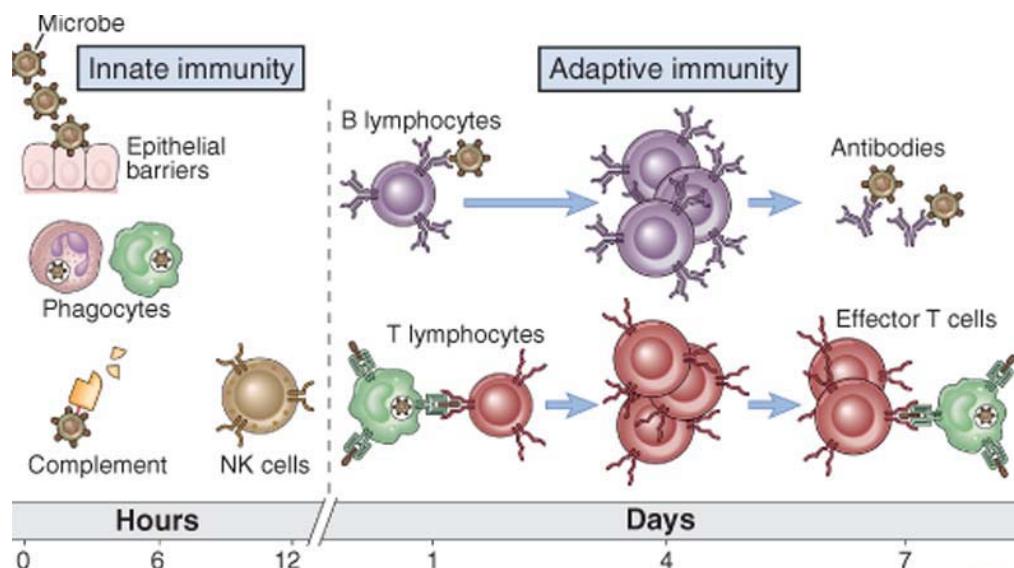


Figure 1. Schematic representation of the innate and adaptive immune systems (From Abbas and Lichtman, 2006)

The innate immune system provides early detection and defense against infections. Components that build innate immunity include epithelial barriers, phagocytes, natural killer cells, and the complement system. Adaptive immunity activates at a later stage and targets antigens in a more specific way. It is mediated by lymphocytes expressing highly specific receptors. Antibodies produced by B cells block infections and eliminate pathogens, and T cells kill intracellular microbes.

Adaptive immunity, which is activated at a later phase of infection, eliminates pathogens majorly through antibody responses and cell mediated immune responses that are carried out by two different types of lymphocytes (B cells and T cells). The production of these

lymphocyte clones involves gene rearrangement events that lead to the generation of a vast number of different antigen receptors that are uniquely expressed on different lymphocyte clones. In contrast, as the first line of host defense, the innate immune system deploys a limited number of germline-encoded receptors called pattern-recognition receptors (PRRs) to detect and respond to the presence of pathogens. PRRs recognize conserved molecular structures known as pathogen-associated molecular patterns (PAMPs) that are essential for the life cycle of the pathogen (Akira et al., 2006).

Detection of Nucleic Acids as Fundamental PAMPs

Many PAMPs, such as lipopolysaccharides, peptidoglycans, and flagellin, are found in microbes but not in the host, allowing the host to distinguish non-self from self through PRRs. One apparent exception is the detection of pathogen-derived nucleic acids. The fact that nucleic acids could serve as PAMPs greatly expands the repertoire of microorganisms detectable by the host immune system. In principle, all microbes use DNA and/or RNA as genetic information carriers in their life cycle and could therefore potentially activate host nucleic acid sensors. However, this also introduces the risk of self nucleic acid recognition by innate immune sensors, an event that causes several autoimmune and autoinflammatory diseases (Wu and Chen, 2014).

Innate immune sensors for nucleic acids can be generally divided into two groups on the basis of their subcellular localization and expression pattern. The first group includes several members of the Toll-like receptor (TLR) family that function mostly in immune cells, such as dendritic cells (DCs), macrophages, and B cells. These TLRs localize on the endosomal membrane and monitor the lumen of endosomes and lysosomes to detect various forms of

nucleic acids from bacteria and viruses. The second group of receptors, which have not been well characterized until recently, detects nucleic acids in the cytoplasm of almost all cell types. Cytosolic nucleic acid sensors include proteins that detect cytoplasmic DNA as well as the RIG-I-like receptor (RLR) family members that detect pathogen-derived RNA in the cytosol (Wu and Chen, 2014).

Upon recognition of nucleic acids, the endosomal TLRs and cytosolic sensors activate a signaling cascade that culminates in production of type I interferons (IFNs), which primarily include numerous subtypes of IFN- α and a single IFN- β , as well as proinflammatory cytokines such as tumor necrosis factor- α (TNF- α) and interleukin-1 β (IL-1 β). Type I IFNs then induce a large array of antiviral genes through the activation of the Janus kinase (JAK)-signal transducer and activator of transcription (STAT) pathway (Platanias, 2005). The induction of type I IFNs is regulated by the transcription factors nuclear factor κ -light-chain enhancer of activated B cells (NF- κ B), and the IFN regulatory factors IRF3 and IRF7. The activation of NF- κ B requires phosphorylation of its inhibitor I κ B by the I κ B kinase (IKK) complex. The phosphorylation of I κ B targets this inhibitor for polyubiquitination and subsequent proteasome-dependent degradation, releasing NF- κ B to the nucleus to regulate downstream genes (Silverman and Maniatis, 2001). Activation of IRF3 and IRF7 requires phosphorylation by two IKK-related kinases, TBK1 and IKK ϵ (Fitzgerald et al., 2003; Sharma et al., 2003). After phosphorylation, these two IRFs undergo homodimerization and translocate into the nucleus, where they form an enhanceosome complex together with NF- κ B and other transcription factors to turn on the transcription of type I IFN genes (Honda and Taniguchi, 2006).

Sensing Nucleic Acids in the Endosome by Toll-like Receptors

Toll-like receptors (TLRs) are among the best-studied groups of PRRs. TLRs are single transmembrane proteins with ectodomains containing leucine-rich repeats for PAMP recognition, a transmembrane domain, and a cytosolic Toll/IL-1 receptor (TIR) domain responsible for transducing signals to downstream adaptors including TRIF and MyD88 (Kawai and Akira, 2010). Among the 10 and 13 identified TLRs in human and mouse, respectively, 5 are involved in nucleic acid sensing: TLR3, TLR7, TLR8, TLR9, and TLR13. These receptors monitor the endolysosomal lumen for pathogen-derived nucleic acids and function via two signaling pathways: TLR3 activates TRIF, whereas TLR7, TLR8, TLR9, and TLR13 activate MyD88. Both adaptor proteins lead to the activation of NF- κ B, whereas IRF3 is activated by the TRIF pathway and IRF7 by the MyD88 pathway (Kawai and Akira, 2010).

Using gene-targeted mice, studies with synthetic ligands and microbes have shed light on the ligand specificity of these nucleic acid-sensing TLRs. TLR3 was originally identified as a sensor for double-stranded RNA (dsRNA) using the synthetic analog polyinosinic-polycytidylic acid (poly[I:C]) as the stimulus (Alexopoulou et al., 2001). Later studies showed that TLR3 was also involved in the host response to respiratory syncytial virus (RSV), encephalomyocarditis virus (EMCV), West Nile virus (WNV), and coxsackievirus (Akira et al., 2006; Kawai and Akira, 2010; Negishi et al., 2008). The two phylogenetically close TLRs, TLR7 and TLR8, share general ligand specificity for single-stranded RNA (ssRNA) derived from RNA viruses (Diebold et al., 2004; Heil et al., 2004) and imidazoquinoline derivatives such as resiquimod (R848) (Hemmi et al., 2002). However,

ssRNA activates only TLR8 and not TLR7 in humans, and mice lack TLR8 but respond normally to these agonists. Thus, TLR7 and TLR8 appear to have distinct functions in different species (Akira et al., 2006; Kawai and Akira, 2010). TLR9 recognizes unmethylated cytosine-guanosine (CpG) DNA motifs that are common for bacterial but rare in mammalian genomes (Hemmi et al., 2000; Kawai and Akira, 2010). Recently, our group and Kirschning's group independently identified murine TLR13 as the sensor for bacterial 23S ribosomal RNA (Li and Chen, 2012; Oldenburg et al., 2012). A sequence containing 13 nucleotides near the active site of 23S rRNA ribozyme is both necessary and sufficient to activate TLR13.

Cytosolic Sensing of RNA and DNA

The ligand-binding domain of the nucleic acid–sensing TLRs faces the lumen of the endosome, rendering these TLRs “blind” to microbes that have successfully invaded the host cytoplasm and replicated inside the cells (Goubau et al., 2013). In addition, most nonimmune cells, such as epithelial cells and fibroblasts that normally line the mucosal surface and thus are most susceptible to infections, do not express the nucleic acid–sensing TLRs but can nevertheless mount effective innate immune responses against microbial infections. Thus, a cell-intrinsic, cytoplasmic surveillance system must exist to defend against microbes that invade both immune and non-immune cells.

RNAs from pathogens invading the cytosol are sensed by RIG-I-like receptors (RLRs) including RIG-I and MDA5. Both RIG-I and MDA5 consist of N-terminal tandem caspase activation and recruitment domains (CARDs), which are important for signal activation, and a central DExH/C motif helicase domain and C-terminal domain (CTD), which together function as an RNA sensing unit (Yoneyama and Fujita, 2009; Yoneyama et al., 2004)

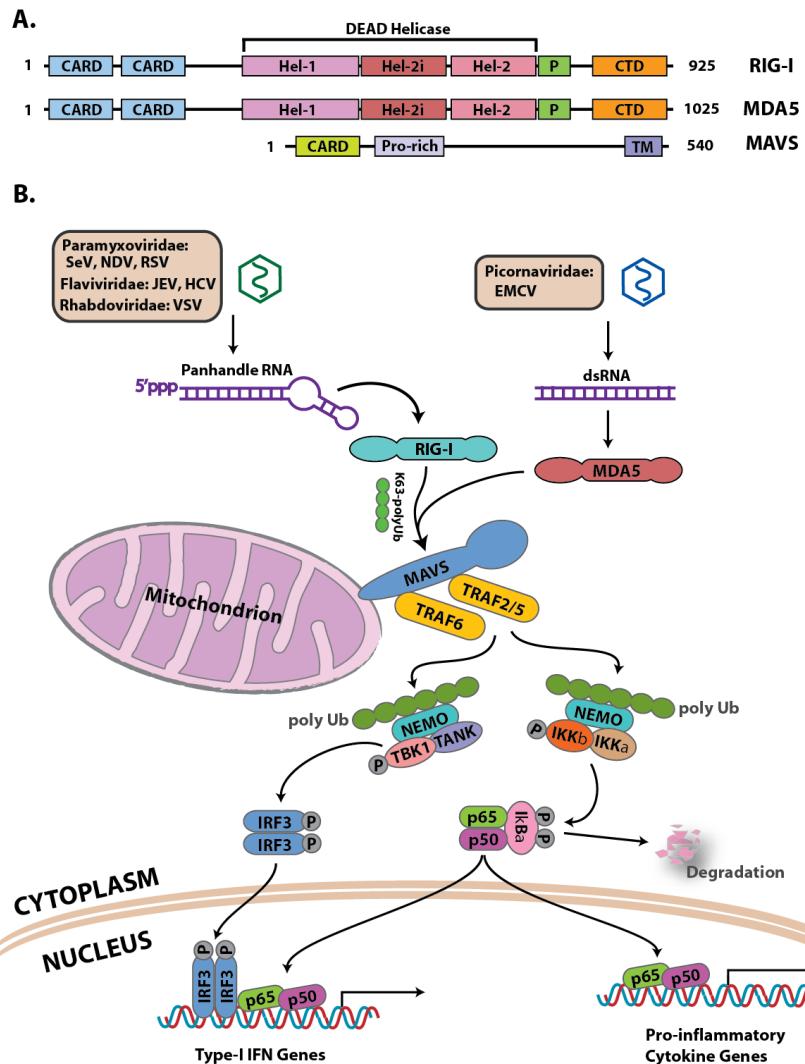


Figure 2. Innate immune sensing and signaling of cytosolic RNA

(A) Domain organization of RLRs and their adaptor MAVS.

(B) RIG-I and MDA5 play nonredundant roles in cytosolic RNA sensing by recognizing different groups of viral RNAs. Whereas RIG-I detects viral RNA containing 5'ppp and panhandle-like secondary structures, MDA5 recognizes long dsRNA in the viral genome. RIG-I and MDA5 activation induce the prion-like polymerization of MAVS, which in turn recruits and activates E3 ligases TRAF2, TRAF3, TRAF5, and TRAF6. These E3 ligases then synthesize polyubiquitin chains that are sensed by NEMO through its ubiquitin-binding domains. NEMO then recruits IKK and TBK1 complexes to the MAVS polymer, where the kinases phosphorylate $\text{I}\kappa\text{B}\alpha$ and IRF3, respectively, leading to the induction of type I interferons and other cytokines.

(Figure 2A). RNA binding induces a conformational change in RLRs, which exposes CARDs

and allows them to interact with the N-terminal CARD of the downstream adaptor MAVS

(Seth et al., 2005). Recently studies suggested that, in addition to RNA ligands, K63-linked ubiquitin chains are also required for the activation of RLRs (Jiang et al., 2012; Peisley et al., 2014; Zeng et al., 2010). When activated by RNA and K63 ubiquitin chains, RIG-I forms oligomers to convert monomeric MAVS on mitochondria into functional prion-like aggregates, which then recruit ubiquitin E3 ligases TRAF2, TRAF5, and TRAF6 to synthesize polyubiquitin chains (Hou et al., 2011; Liu et al., 2013); in turn, these ubiquitin chains recruit and activate kinases IKK and TBK1, which subsequently activate the transcription factors IRF3 and NF- κ B to trigger IFN production (Liu et al., 2013) (Figure 2B).

The presence of cytosolic DNA triggers potent immune responses including inflammasome activation and type I IFN induction. The AIM2 protein recognizes cytosolic DNA in the inflammasome pathway, which leads to pyroptosis and the processing of inflammatory cytokines including interleukin-1 β (IL-1 β) and IL-18 (Burckstummer et al., 2009; Fernandes-Alnemri et al., 2009; Hornung et al., 2009) (Figure 3). Cytosolic DNA induces type I IFNs through an adaptor protein STING, which is localized on the endoplasmic reticulum membrane. STING further activates the kinase TBK1 and transcription factor IRF3 to induce IFNs (Ishikawa and Barber, 2011; Ishikawa et al., 2009) (Figure 3). In addition, it is also known that STING can be activated by bacterial cyclic-dinucleotides (CDNs) (Burdette et al., 2011). However, it remained elusive how cytosolic DNA is sensed and how this sensing event leads to the activation of STING. Many proteins, such as DAI, IFI16 and DDX41 had been proposed as putative DNA sensors (Takaoka et al.,

2007; Unterholzner et al., 2010; Zhang et al., 2011). However, none of them has been met with universal acceptance.

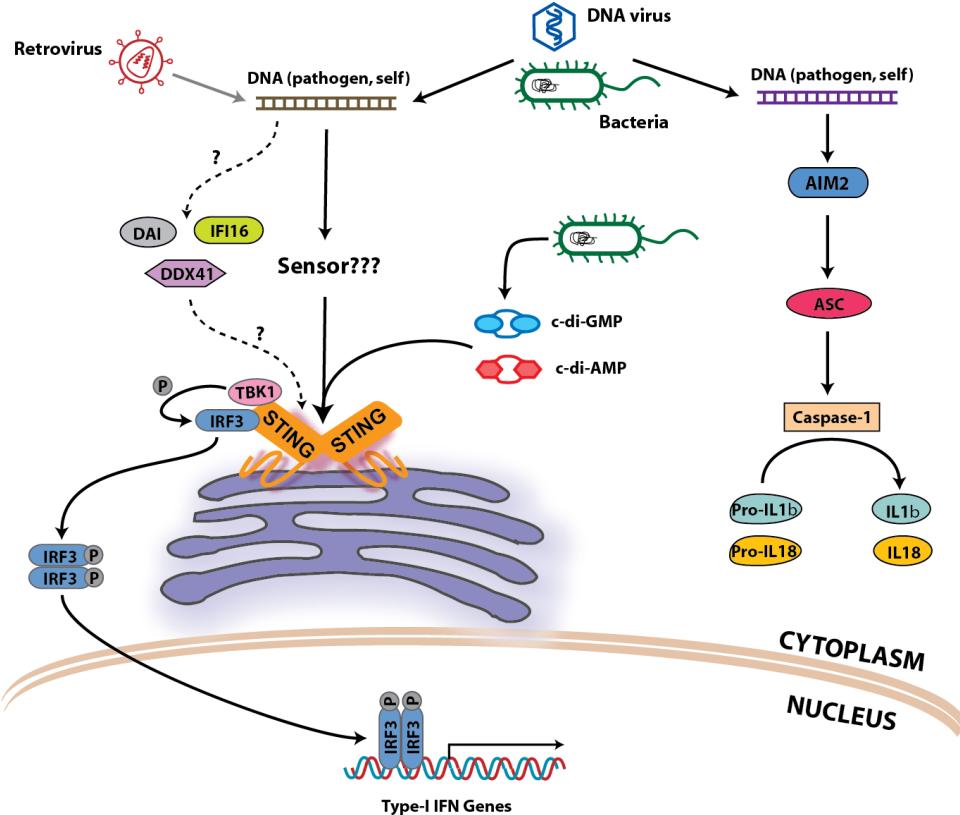


Figure 3. Innate immune sensing and signaling of cytosolic DNA

Cyttoplasmic DNA of self or microbial origin will activate a putative DNA sensor, which then transduce signals to the ER-localized adaptor protein STING. STING recruits and activates the kinase TBK1, which then activates the transcription factor IRF3 to induce type I IFNs. The bacterial second messengers c-di-GMP and c-di-AMP are directly sensed by STING. In addition to inducing type I IFNs, cytosolic DNA can also activate the AIM2 inflammasome, leading to caspase-1 activation and IL-1 β maturation.

Cytosolic DNA Sensing in Human Physiology and Pathology

Even though the identity of the cytosolic DNA sensor that activates IFNs remained elusive, it had been widely accepted that the STING-dependent DNA sensing pathway plays crucial roles in both human physiology and pathology.

Accumulating evidence demonstrated that the cytosolic DNA sensing pathway plays critical roles in host defense against a broad spectrum of infectious pathogens, including DNA viruses (Ishikawa et al., 2009), retroviruses (Silvin and Manel, 2015; Yan et al., 2010), RNA viruses (Ishikawa et al., 2009), bacteria (Dey et al., 2015; Hansen et al., 2014b; Manzanillo et al., 2012), and parasites (Sharma et al., 2011). Animals with compromised function of this pathway were more susceptible to lethal infections by many of the pathogens listed above. Conversely, as this pathway is activated by the presence of cytosolic DNA independent of its source and sequence, aberrant detection of self-DNA that has leaked into the cytosol can also lead to sterile inflammation and autoimmunity. For example, deficiency of the endogenous DNA exonuclease Trex-1 will cause Aicardi-Goutières Syndrome (Crow et al., 2006), an encephalopathy exhibiting phenotypic overlaps with the autoimmune disease SLE. Mechanistically, deficiency in Trex1 results in cytosolic accumulation of DNA derived from endogenous retroelements or replication debris, which further activate the cytosolic DNA sensing pathway (Gall et al., 2012; Stetson et al., 2008). In addition, gain-of-function mutations in STING have been shown to cause a systemic vascular and pulmonary inflammation called STING-Associated Vasculopathy with onset in Infancy (SAVI) (Jeremiah et al., 2014; Liu et al., 2014).

DNA vaccines are a new type of vaccine that involve direct injection of plasmids encoding an antigen into a living host to induce humoral and cellular immune responses against the antigen. The activation of the cytosolic DNA sensing pathway in different cell types had been implicated in mediating the immunogenicity of DNA vaccines (Coban et al., 2014; Ishii et al., 2008; Ishikawa et al., 2009). More recently, the cytosolic DNA sensing

pathway was shown to play important roles in anticancer immunity. Spontaneous CD8⁺ T cell priming against immunogenic tumors requires the DNA sensing adaptor STING (Woo et al., 2014). In a radiotherapy model, STING dependent induction of IFN- β in dendritic cells by irradiated-tumor cells was required for radiation-induced antitumor immunity (Deng et al., 2014).

CHAPTER TWO

Results

Activity Assay for Cellular Factors That Activate the STING Pathway

We hypothesized that DNA binds to and activates a putative cytosolic DNA sensor, which then directly or indirectly activates STING, leading to the activation of TBK1 and IRF3 (Figure 4A). If this is the case, DNA stimulated cells should contain an active factor; this active factor could be the DNA sensor itself in its active form or other protein(s) recruited and activated by the DNA sensor. When this putative active factor is delivered to reporter cells, which contain all the downstream components of the pathway, it should activate IRF3 (Figure 4B). To test this model, an in vitro complementation assay was established by Dr. Lijun Sun in the laboratory, using the murine fibrosarcoma cell line L929. L929 cells are known to induce IFN- β in a STING-dependent manner (Tanaka and Chen, 2012). An L929 cell line stably expressing an shRNA against STING was used such that DNA transfection would only activate factor(s) upstream of STING. The L929-shSTING cells were transfected with different types of DNA, and then cytoplasmic extracts from these cells were mixed with reporter cells (human monocytic cell line THP1 or murine macrophage cell line Raw264.7), which were permeabilized with perfringolysin O (PFO; Figure 4B). PFO treatment creates holes in the plasma membrane (Rossjohn et al., 2007), allowing the cytoplasm to diffuse in and out of cells, while retaining organelles including the endoplasmic reticulum (which contains STING) and the Golgi apparatus inside the cells. As shown in Figure 4C, cytoplasmic extracts from L929-shSTING cells transfected with a DNA sequence known as interferon-stimulatory DNA (ISD), activated IRF3 in permeabilized THP1 cells.

This activity was also detected in the cytosolic extracts from cells transfected with poly(dA:dT), a GC-rich 50–base pair double-stranded DNA (G:C50), poly(deoxyinosine-deoxycytidine) [poly(dI:dC)], or herring testis DNA (HT-DNA) (Figure 4D), indicating that L929 cells generated this DNA-induced activator independent of the sequence of DNA.

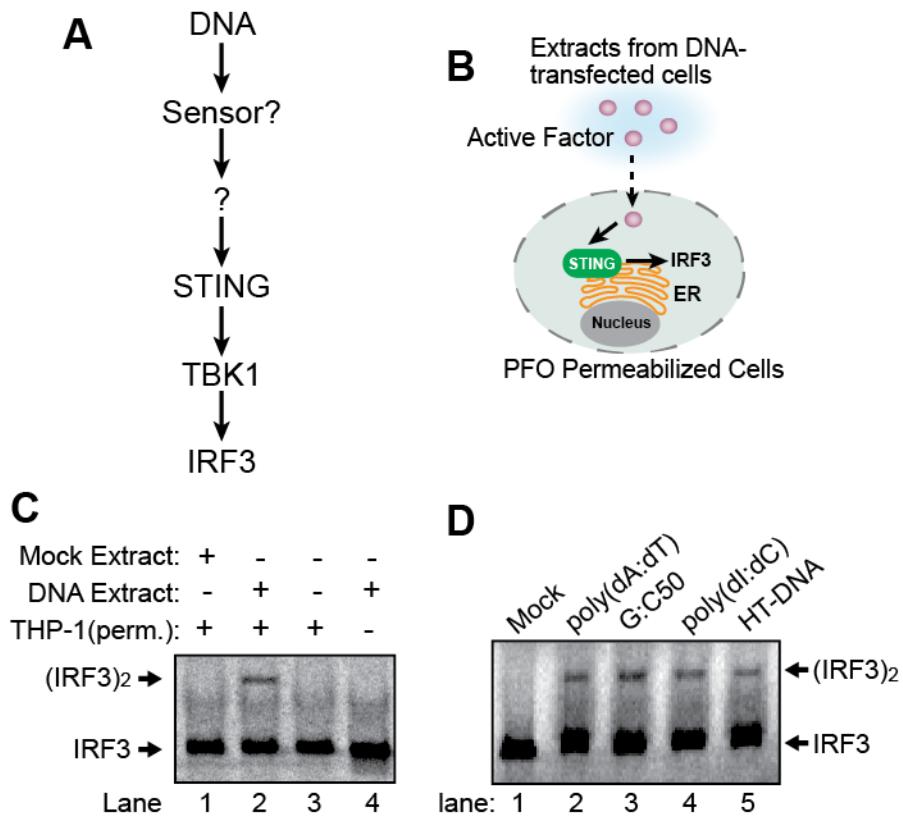


Figure 4. An in vitro complementation assay for cellular factors that activate the STING pathway

- (A) A proposed model of the cytosolic DNA sensing and signaling pathway.
- (B) Schematic view of an activity assay for monitoring the STING activator.
- (C) Cytosolic extracts from mock or ISD-transfected L929-shSTING cells were incubated with PFO-permeabilized THP1 cells together with ³⁵S-labeled IRF3. Dimerization of IRF3 was analyzed by native gel electrophoresis followed by autoradiography.
- (D) Similar to (C), except that the indicated DNA were used for transfection.

Characterization of the STING Activator

To determine whether the STING activator is a protein, we incubated the cytoplasmic extracts at 95°C to denature most proteins and then incubated the “heat supernatant” with

permeabilized THP1 cells. Surprisingly, the heat supernatant from the ISD-transfected or HT-DNA–transfected cells caused IRF3 dimerization (Figure 5A). This activity was resistant to treatment with Benzonase (Novagen, Figure 5B), which degrades both DNA and RNA, and proteinase K (Figure 5B). Thus, the STING activator is probably not a protein, DNA, or RNA.

To test whether DNA could stimulate the generation of the heat-resistant activator in vitro, we incubated HT-DNA with L929- shSTING cytoplasmic extracts (S100) in the presence of ATP (Figure 5C). The reaction mixture was heated at 95°C to denature proteins. Remarkably, incubation of the supernatant with permeabilized Raw264.7 cells led to IRF3 dimerization (Figure 5D, lane 2). This activity depended on the addition of DNA to the cytoplasmic extracts. Other DNAs, including poly(dA:dT), poly(deoxyguanosine-deoxycytidine), and ISD, also stimulated the generation of the STING activator in L929-shSTING cytoplasmic extracts, whereas poly(inosine-cytidine) [poly(I:C)] and single-stranded RNA had no activity (Figure 5D).

This activator induces IRF3 dimerization through STING, as knockdown of STING in the permeabilized THP1 cells abolished IRF3 activation by the heat-resistant factor generated by DNA transfected into L929 cells or DNA added to L929 cytosolic extracts (Figure 5E). I also tested cytoplasmic extracts from several cell lines for their ability to produce the heat-resistant STING activator (Figure 5F). Incubation of HT-DNA with extracts from primary mouse embryo fibroblasts (MEFs), mouse bone marrow derived macrophages (BMDMs), and L929 cells led to generation of the heat- resistant factor that activated IRF3. Human cell extracts from THP1, but not human embryonic kidney (HEK) 293T cells, were also able to

produce this STING activator. These results are consistent with previous findings in the lab that primary MEFs, BMDMs, and L929 and THP1 cells, but not HEK293T cells, possessed the STING-dependent cytosolic DNA sensing pathway to induce type I interferons (Chiu et al., 2009).

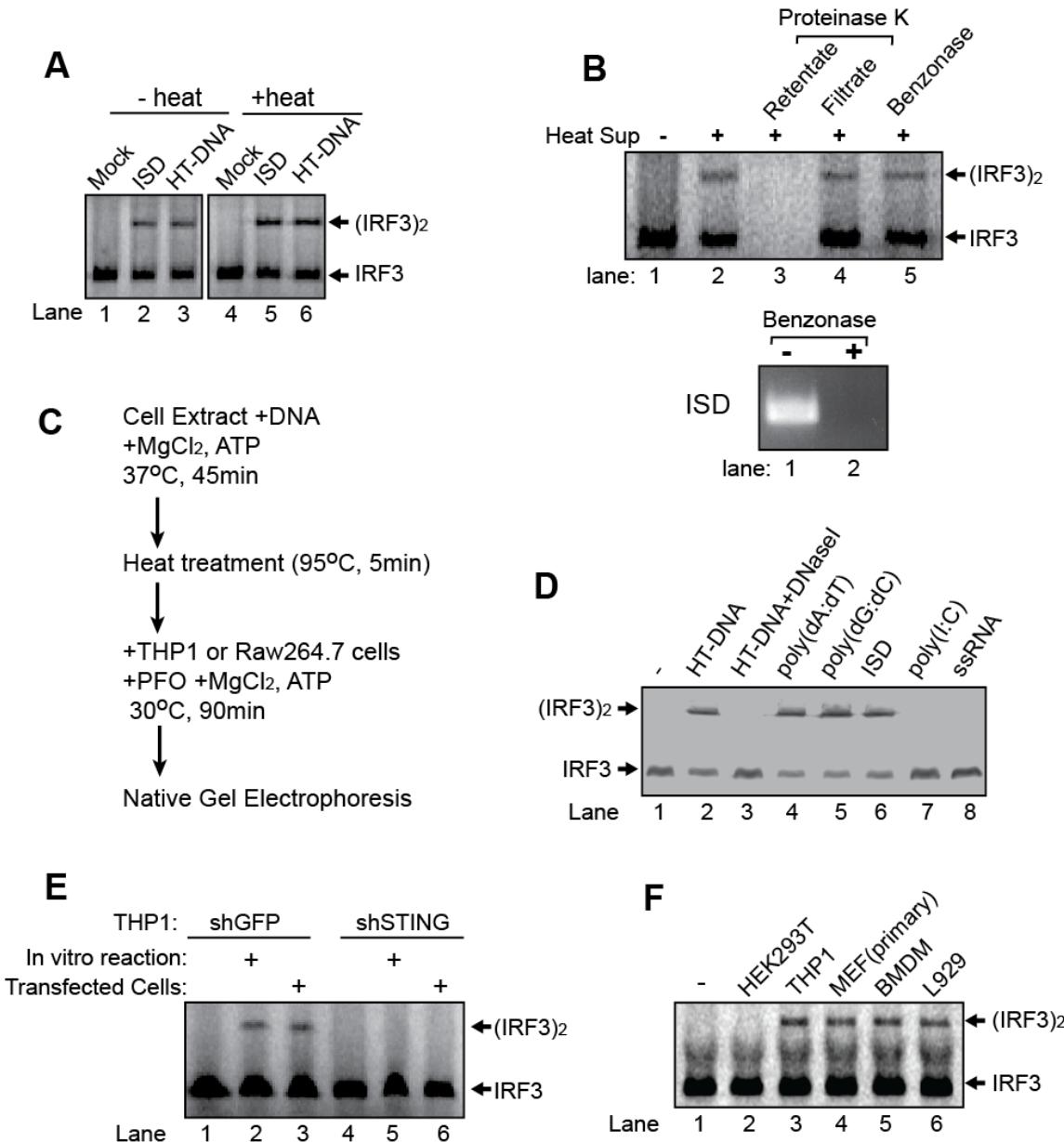


Figure 5. DNA-dependent generation of a heat-resistant small molecule activates the STING pathway

- (A) Cytosolic extracts from mock, ISD or HT-DNA transfected L929-shSTING cells were heated at 95°C for 5 min to denature proteins and then the heat-resistant supernatant was incubated with PFO-permeabilized THP1 cells together with ³⁵S-labeled IRF3. Dimerization of IRF3 was analyzed by native gel electrophoresis followed by autoradiography.
- (B) Similar to (A), except that instead of being heated, cytoplasmic extracts from HT-DNA transfected L929 cells were incubated with proteinase K (lanes 3 and 4), Benzonase (lane 5) or untreated (lane 2). The proteinase K treated samples were separated by a filter with a molecular weight cut-off of 3 kDa. The retentate (large molecules; lane 3) and filtrate (small molecules, lane 4) as well as other samples were incubated with PFO permeabilized THP1 cells and then IRF3 dimerization was examined. The effectiveness of Benzonase treatment was verified by its ability to digest DNA (bottom).
- (C) Illustration of a cell-free system that synthesizes and detects the heat-resistant STING activator using DNA-supplemented L929 cytosolic extracts.
- (D) L929-shSTING cytosolic extracts were incubated with the indicated nucleic acids in the presence of ATP, and then the heat-resistant supernatant was assayed for its ability to stimulate IRF3 dimerization in permeabilized Raw264.7 cells.
- (E) THP1 cells stably expressing shRNA against GFP (control) or STING were permeabilized with PFO and then incubated with the heat-resistant supernatant from the reaction mixture containing DNA-supplemented L929 cytosolic extracts (lanes 2 and 5) or from DNA-transfected L929 cells (lanes 3 and 6). IRF3 activation was analyzed by native gel electrophoresis.
- (F) Cytosolic extracts from the indicated cell lines were incubated with HT-DNA, and then heat-resistant supernatants were assayed for their ability to stimulate IRF3 dimerization in permeabilized Raw264.7 cells.

Purification and Identification of cGAMP as the STING Activator

We next used several chromatographic steps to purify the heat-resistant STING activator from L929 cell extracts (Figure 6A). The purification procedure includes a STING-Flag affinity pull-down, an anion-exchange column mono Q, a size exclusion column Superdex Peptide and a reverse-phase column C18. We were able to get a decent recovery of the activity from the last step C18 column that allowed further identification of the active factor (Figure 6B).

I then analyzed the fractions from the C18 column using nano–liquid chromatography–mass spectrometry (nano-LC-MS). Examination of the MS spectra revealed two ions with mass-to-charge ratios (*m/z*) of 675.1 ($z = 1^+$) and 338.1 ($z = 2^+$), which were present in the

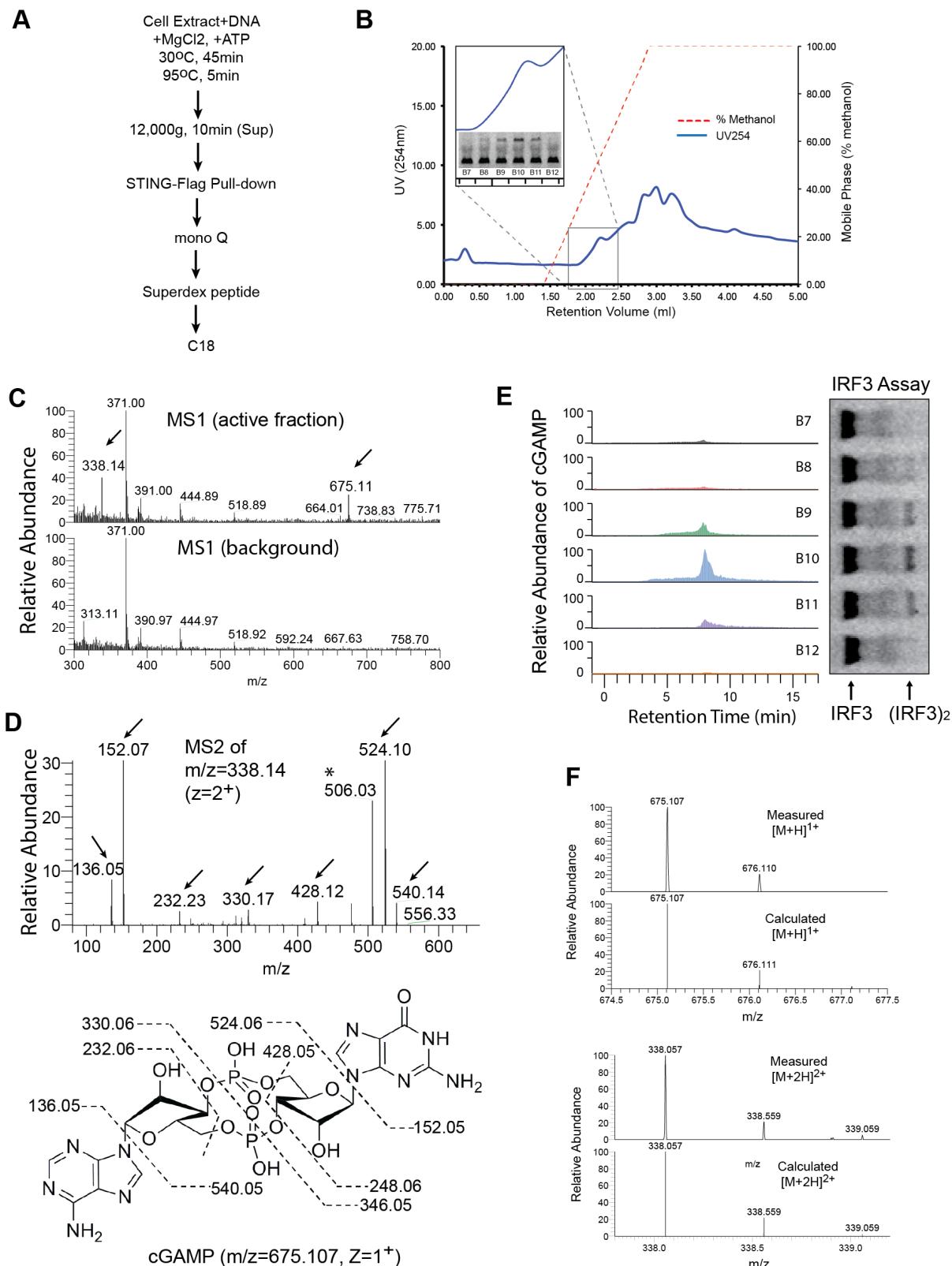


Figure 6. Purification and identification of cGAMP as the STING activator

- (A) Illustration of the synthesis and purification of the small molecule STING activator from L929 cytosolic extracts.
- (B) Purification of the small molecule STING activator by HPLC using a C18 column. The starting materials for this column had been partially purified through three steps as described in A. The inset indicates IRF3 dimerization assays for selected fractions from the C18 column.
- (C) Full scan nano-LC-MS spectra of active and inactive fractions from the C18 column. Arrows indicate an ion at +1 (675.11) and +2 (338.14) charge states present only in the active fraction.
- (D) Tandem mass (MS2) spectra after CID fragmentation of the ion with $m/z = 338.14$ ($z = 2$) from the MS1 scan shown in (C). Arrows indicate the m/z values of the expected fragmentation patterns of cyclic GMP-AMP (cGAMP, bottom). Asterisk indicates an ion ($m/z = 506$) that resulted from a neutral loss of a water molecule (18) from the ion with $m/z = 524$.
- (E) Fractions (B7 to B12) from the C18 column were analyzed for the presence of cGAMP by selective reaction monitoring of the expected ions and for their ability to stimulate IRF3 dimerization.
- (F) High resolution high accuracy measurement of m/z values of the ions present in the purified fractions containing the STING activator. The measurement was performed using Orbitrap in Q Exactive. The calculated m/z values of cGAMP are shown for comparison. The lower intensity peaks represent isotopes of the major ions.

active fractions but absent in the background spectra (Figure 6C). These m/z values, despite the low mass accuracy of the mass spectrometer (LTQ, Thermo), were equivalent to the average calculated m/z values of previously identified bacterial second messengers c-di-GMP and c-di-AMP [$675 = (691 + 659)/2$]. This observation suggested that the detected ion might be a hybrid of c-di-GMP and c-di-AMP—that is, cyclic GMP-AMP, or cGAMP ($m/z = 675.107$, $z = 1^+$; $m/z = 338.057$, $z = 2^+$). This possibility was further bolstered by the Collision-induced dissociation (CID) fragmentation spectra of the ion $m/z = 338.1(z = 2^+)$, which revealed several prominent ions with m/z values expected of the product ions of cGAMP (Figure 6D).

Quantitative mass spectrometry using selective reaction monitoring (SRM) showed that the abundance of the ions representing cGAMP in the fractions from the C18 column correlated very well with their IRF3-stimulatory activities (Figure 6E). To further confirm

the identity of the heat-resistant STING activator, I used a high-resolution high-accuracy mass spectrometer (Q Exactive, Thermo) to perform nano-LC-MS analysis. The cell-derived STING activator had m/z values of 675.107 ($z = 1^+$) and 338.057 ($z = 2^+$), which exactly matched the theoretical values of cGAMP (Figure 6F). Together, these results identified the STING activator as cGAMP, a novel second messenger in metazoa.

DNA Transfection and DNA Virus Infection Induce IFN- β Through cGAMP

To determine whether DNA or DNA virus stimulation leads to the production of cGAMP in cells, I infected L929 cells with Herpes Simplex Virus 1 (HSV-1) lacking ICP34.5, a viral protein known to antagonize interferon production in infected cells (Mossman and Smiley, 2002). Like DNA transfection, HSV-1 Δ ICP34.5 infection led to IRF3 activation in L929 cells (Figure 7A, upper). Cell extracts from the DNA-transfected or virus-infected cells contained a heat-resistant factor that could activate IRF3 in permeabilized Raw264.7 cells (Figure 7A, lower panel). As a control, I infected L929 cells with vesicular stomatitis virus (VSV), an RNA virus known to trigger strong interferon production through the RIG-I pathway (Sun et al., 2006). In contrast to HSV-1, VSV-infected cells did not contain the heat-resistant IRF3 activator in the same in vitro assay, although VSV infection did induce IRF3 activation in L929 cells (Figure 7A). The heat-resistant factor in these cells was further enriched by HPLC using a C18 column and then quantified by nano-LC-MS using SRM. The result suggests that cGAMP was induced in DNA-transfected or HSV-1-infected cells, but not mock-treated or VSV-infected L929 cells (Figure 7B). To test whether DNA viruses could induce cGAMP production in human cells, we infected THP1 cells with HSV1 or vaccinia virus (VACV) (Figure 7C). Both viruses induced IRF3 dimerization in the

cells, and both viruses also triggered the production of cGAMP that activated IRF3 (Figure 7C, lower panel).

I also conducted a kinetic experiment, which showed that after DNA was transfected into L929 cells, cGAMP was generated before IRF3 dimerization and IFN- β induction could be detected (Figure 7D). When chemically synthesized cGAMP was delivered to cells, it strongly induced IFN- β RNA and protein, as shown by qRT-PCR and ELISA assays (Figure 7E). Titration experiments showed that cGAMP induced IFN- β RNA robustly even at

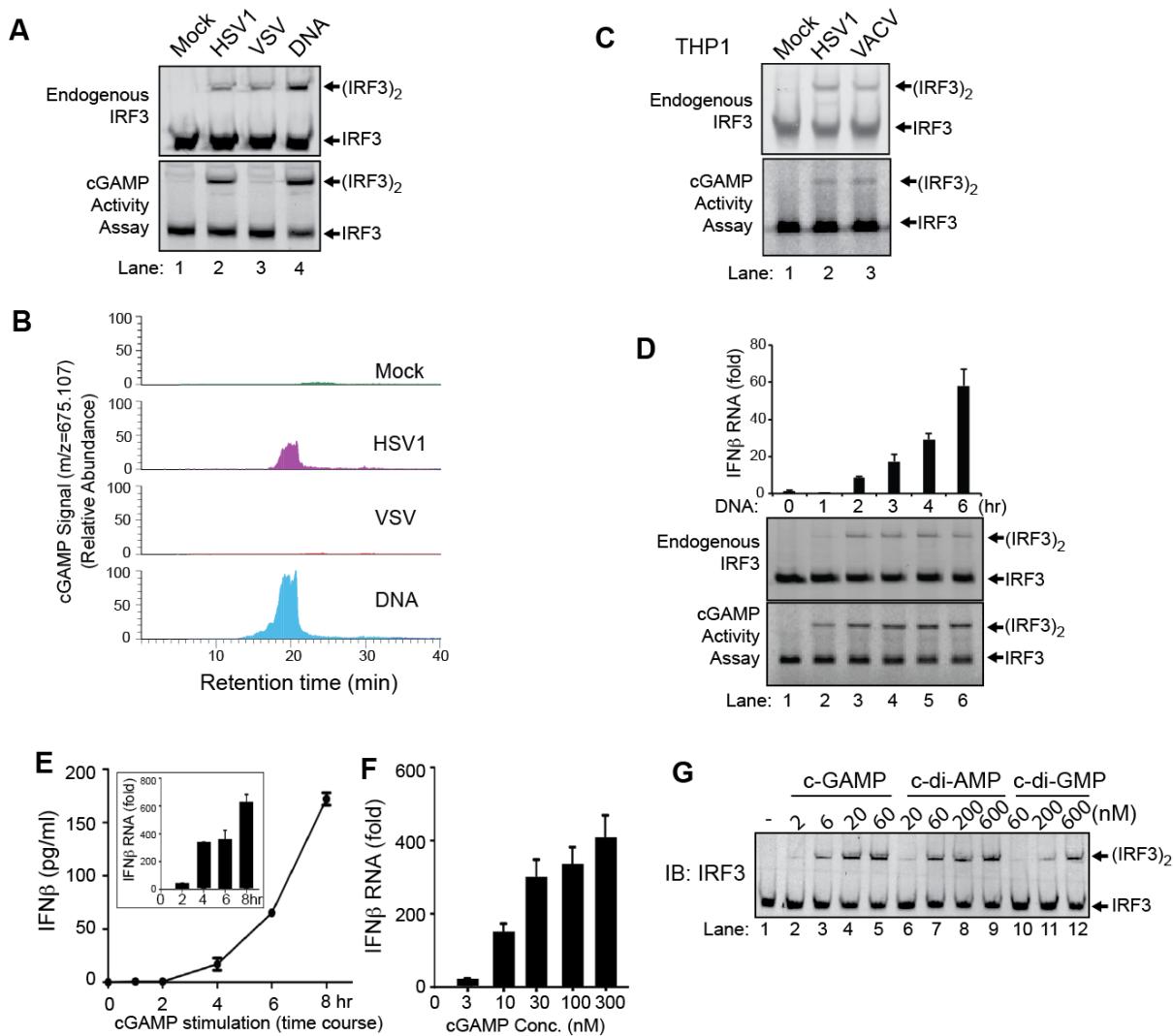


Figure 7. DNA transfection and DNA virus infection induce IFN- β through cGAMP

- (A) L929 cells were infected with HSV-1Δ34.5 or VSV-GFP, transfected with HT-DNA, or mock-treated. An aliquot of the cell extracts was directly analyzed for IRF3 dimerization (top), whereas another aliquot was heated to denature proteins and the heat-resistant supernatant was assayed for its ability to stimulate IRF3 dimerization in permeabilized Raw264.7 cells (bottom). (B) The heat-resistant supernatant from (A) was fractionated by HPLC using a C18 column, and the presence of cGAMP in the fractions was measured by mass spectrometry using SRM. (C) THP1 cells were infected with HSV-1Δ34.5 and VACV for 6 hrs, and then the activation of endogenous IRF3 and generation of cGAMP activity were measured. (D) L929 cells were transfected with HT-DNA (4 μ g/ml) for the indicated time, then IFN- β RNA was measured by qRT-PCR and IRF3 dimerization was analyzed by native PAGE. Aliquots of the cell extracts were tested for the presence of cGAMP on the basis of its ability to induce IRF3 dimerization after delivery into Raw264.7 cells. (E) Chemically synthesized cGAMP (100 nM) was delivered to digitonin-permeabilized L929 cells for the indicated times, and then IFN- β RNA and secreted protein were measured by qRT-PCR (inset) and ELISA, respectively. (F) Similar to (E), except that different concentrations of cGAMP were delivered into L929 cells for 8 hrs, followed by qRT-PCR analyses of IFN- β RNA. (G) Different concentrations of cGAMP, c-di-GMP and c-di-AMP were delivered into L929 cells followed by IRF3 dimerization assays.

concentrations as low as 10 nM (Figure 7F). cGAMP was also more potent than bacterial second messengers c-di-GMP and c-di-AMP in activating IRF3 (Figure 7G). Collectively, these results indicate that DNA transfection and DNA virus infections in cells produced cGAMP, which led to IRF3 activation.

cGAMP Activates IRF3 in a STING-Dependent Manner

I carried out loss-of-function and gain-of-function experiments to test whether cGAMP activates the pathway through STING. I first tested the response of L929 and L929-shSTING cells to cGAMP (Figure 8A). Similar to ISD and c-di-GMP, cGAMP-induced IRF3 dimerization was dependent on STING. As a control, poly(I:C) still induced IRF3 dimerization in the absence of STING. These results suggest that STING is necessary for cGAMP to activate IRF3.

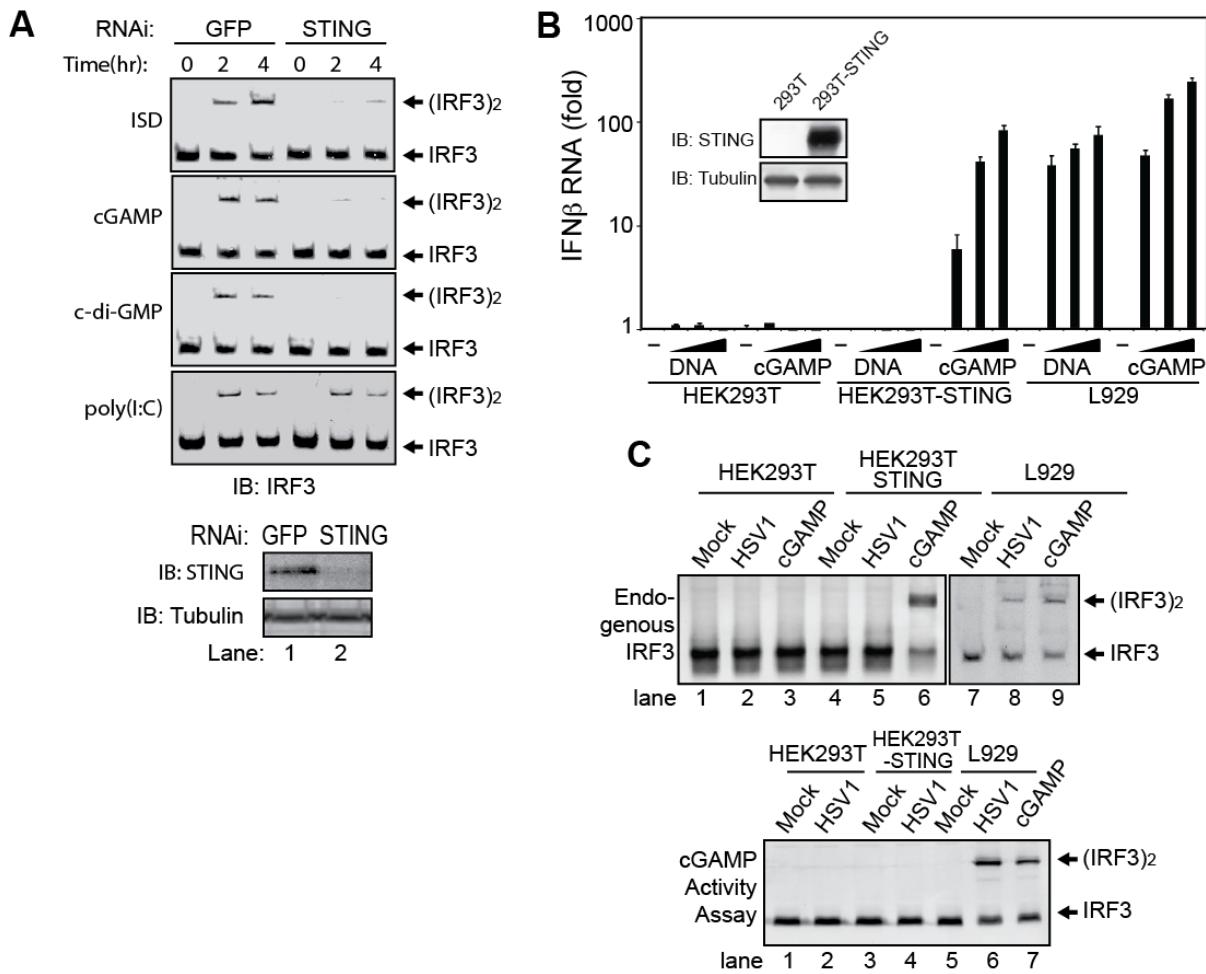


Figure 8. cGAMP activates IRF3 and IFN- β through STING

(A) cGAMP, c-di-GMP, ISD, or poly(I:C) were delivered into L929 cells stably expressing an shRNA against GFP or STING for the indicated time, followed by analysis of IRF3 dimerization.

(B) Increasing concentrations of HT-DNA or cGAMP were delivered to indicated cells, and the induction of IFN- β was measured by qRT-PCR. Inset shows immunoblots of STING and β -tubulin in the cell lines.

(C) Indicated cell lines were infected with HSV1 Δ 34.5 or permeabilized with digitonin and then incubated with cGAMP. Activation of endogenous IRF3 was analyzed by native gel electrophoresis (top). Aliquots of the cytosolic extracts were heated to denature proteins, and the supernatant was assayed for its ability to stimulate IRF3 in permeabilized Raw264.7 cells (bottom).

It is well known that HEK293T cells do not contain the endogenous DNA sensing pathway due to a lack of the STING protein. I stably expressed STING in this cell line and

stimulated it with cGAMP, and then measured IFN- β induction by quantitative RT-PCR (Figure 8B). Even though parental HEK293T cells did not respond to cGAMP, stable expression of STING rendered them responsive to cGAMP, as shown by a high level of IFN- β induction after cGAMP delivery. However, DNA did not stimulate HEK293T-STING cells to induce IFN- β , suggestive of a defect of HEK293T cells in producing cGAMP in response to DNA stimulation. In contrast, L929 cells produced IFN- β in response to both cGAMP and DNA. HSV-1 infection-induced IRF3 dimerization in L929 cells but not in HEK293T or HEK293T-STING cells (Figure 8C, upper panel), which suggests that the generation of cGAMP is important for HSV-1 to activate IRF3 in cells. Indeed, extracts from HSV1-infected L929 cells, but not from HEK293T or HEK293T-STING cells, contained the cGAMP activity that led to IRF3 dimerization in permeabilized Raw264.7 cells (Figure 8C, lower panel). These results indicate that the expression of STING in HEK293T cells restored the ability of the cells to activate IRF3 and induce IFN- β in response to cGAMP, but was insufficient to install the response to DNA or DNA viruses due to an inability of HEK293T cells to synthesize cGAMP.

Purification and Identification of cGAMP Synthase (cGAS)

To identify the enzyme responsible for cGAMP production, I fractionated cytosolic extracts (S100) from the L929 cells, which contains the cGAMP-synthase (cGAS) activity. This activity was assayed by incubating the column fractions with ATP and GTP in the presence of HT-DNA. After digesting the DNA with Benzonase and heating at 95°C to denature proteins, the heat-resistant supernatants that contained cGAMP were supplemented to PFO-permeabilized Raw264.7 cells. cGAMP-induced IRF3 dimerization in these cells was

analyzed by native gel electrophoresis. Using this assay, Dr. Lijun Sun and I carried out three independent routes of purification, each containing four steps of chromatography but differing in the columns or the order of the columns used (Figure. 9A). Specifically, all three routes started with a column packed with Heparin Sepharose, which can bind to proteins via both affinity and ion exchange mechanisms. It is also worth mentioning that the third route included an affinity pull-down using biotinylated-ISD. It is estimated that we achieved a range of 8000- to 15,000-fold purification and 2 to 5% recovery of the activity from these routes of fractionation. I performed silver staining to visualize the fractions from the last step of each fraction route. However, it did not reveal distinct protein bands that copurified with the cGAS activity, which suggests that the abundance of the putative cGAS protein might be very low in L929 cytosolic extracts.

As I did not see any activity-correlated band that could be identified by mass spectrometry directly, I developed a quantitative mass spectrometry strategy to identify a group of proteins that copurified with the cGAS activity at the last step of each purification route. It was very likely that an overlap of the lists from all the three routes would give us a short list of pursueable candidates, as the putative cGAS protein must copurify with its activity in all three purification routes, whereas most “contaminating” proteins would not. Thus, from the last step of each purification route, I chose fractions that contained most of the cGAS activity (peak fractions) and adjacent fractions that contained very weak or no activity (Figure 9B). The proteins in each fraction were separated by SDS-PAGE and identified by tandem mass-spectrometry. The data were analyzed by label-free quantification using the MaxQuant software (Cox and Mann, 2008); the number of proteins that co-purified

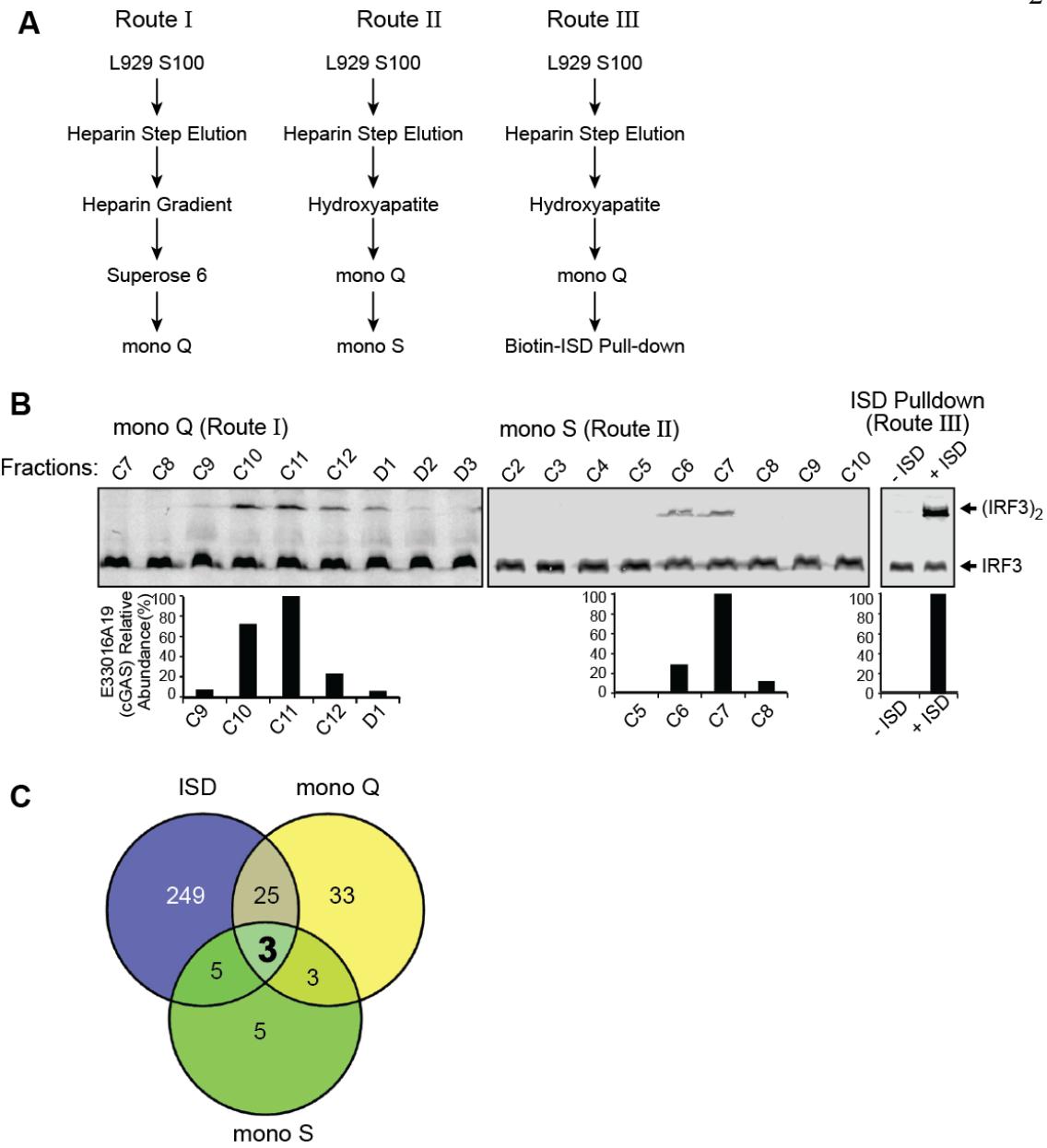


Figure 9. Purification of a cGAMP synthase (cGAS) and its identification by quantitative mass spectrometry

(A) Diagram of three different routes to purify endogenous cGAS activity from L929 cell extracts. (B) Fractions from the last step of each purification route were tested for the ability to generate cGAMP, which was delivered to permeabilized Raw264.7 cells followed by measurement of IRF3 dimerization. The relative abundance of E33016A19 (m-cGAS) was estimated by label free quantification of mass spectrometry data generated from 80 samples (gel slices). (C) Venn diagram depicting the number of proteins that overlapped among different purification routes based on quantitative mass spectrometry analyses.

with the cGAS activity are shown in a Venn diagram (Figure 9C). Strikingly, although many proteins co-eluted with the cGAS activity in one or two purification routes, only three proteins co-purified in all three routes. All three were putative uncharacterized proteins: E330016A19 (NCBI accession number NP_775562), Arf-GAP with dual PH domain containing protein 2 (NP_742145), and signal recognition particle 9 kD protein (NP_036188). Among these, more than 24 unique peptides were identified in E330016A19, covering 41% its 507 amino acids.

Bioinformatic analysis immediately drew my attention to E330016A19, which exhibited structural and sequence homology to the catalytic domain of human oligoadenylate synthase (OAS1) (Figure 10A). In particular, E330016A19 contains the conserved active site residues hG[GS], [DE]h[DE]h and h[DE]h (h indicates a hydrophobic residue) that were found in the nucleotidyltransferase (NTase) family (Kuchta et al., 2009). Besides OAS1, this family includes adenylate cyclase, polyadenylate polymerase, and DNA polymerases. The C-terminus of E330016A19 contained a Mab21 (male abnormal 21) domain, which was first identified in the *Caenorhabditis elegans* protein Mab21. Sequence alignment revealed that the C-terminal NTase and Mab21 domains are highly conserved from zebrafish to human (Figure 10, B and C), whereas the N-terminal sequences are much less conserved. The human homolog of E330016A19, C6orf150 (also known as MB21D1), was recently identified as one of several viral replication inhibitors in an overexpression screen for more than 380 human interferon-stimulated genes (ISGs) (Schoggins et al., 2011). Based on the experimental evidence shown below, we named the mouse protein E330016A19 m-cGAS and the human homolog C6orf150 h-cGAS.

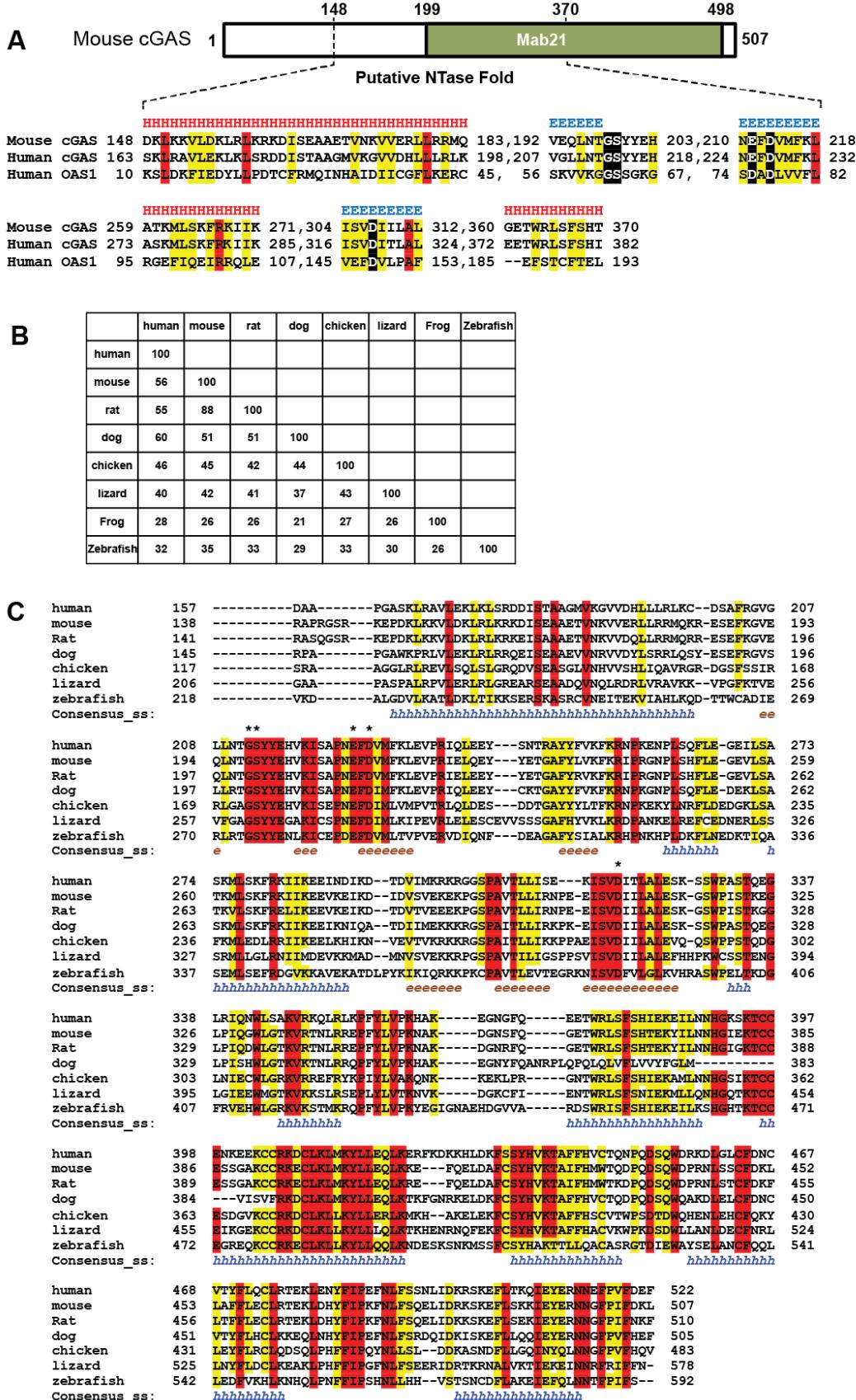


Figure 10. Bioinformatic analyses of cGAS

(A) Multiple sequence and structure alignment of the putative NTase domain of mouse cGAS, human cGAS, and human OAS1 using the PROMALS3D program. Conserved active site residues of the NTase superfamily are highlighted in black, identical amino acids in red, and conserved amino acids in yellow. Predicted secondary structure is indicated above the alignment as α helices (H) and β strands (E).

(B) Identity scores between amino acid sequences of cGAS from pairs of vertebrate species.

(C) Multiple sequence alignment of the C-terminal conserved regions of cGAS proteins from various species using the *PROMALS3D* program. Identical amino acids are indicated in red and conserved amino acids in yellow. Predicted secondary structure elements are indicated below the alignment as alpha helices (*h*) and beta strands (*e*). Conserved active site residues of NTase superfamily are marked with an asterisk (*).

cGAS Activates IRF3 and Induces IFN- β in a STING-Dependent Manner

As mentioned before, HEK293T cells lack STING expression as well as cGAS activity.

Overexpression of m-cGAS in HEK293T-STING cells but not HEK293T cells strongly induces IFN- β (Figure 11A), suggesting that cGAS activates the pathway through STING. Point mutations of the putative catalytic residues G198 and S199 to alanine abolished the ability of m-cGAS to induce IFN- β . This double mutant, as well as another mutant harboring alanine mutations on the other putative catalytic residues E211 and D213, also abrogated the ability of m-cGAS to induce IRF3 dimerization in HEK293T-STING cells (Figure 11B).

These results indicated that cGAS activated the pathway through its catalytic activity, or by producing cGAMP.

The level of IFN- β triggered by cGAS was similar to that induced by the RNA sensing pathway adaptor MAVS and was several orders of magnitude higher than that induced by other previously reported DNA sensors, including DAI, IFI16, and DDX41 (Figure 11C). To determine whether overexpression of cGAS and other putative DNA sensors led to the production of cGAMP in cells, I incubated supernatants from heat-treated cell extracts with

PFO-permeabilized Raw264.7 cells, followed by measurement of IRF3 dimerization. Among all the proteins expressed in HEK293T-STING cells, only cGAS was capable of producing the cGAMP activity in the cells (Figure 11D, bottom).

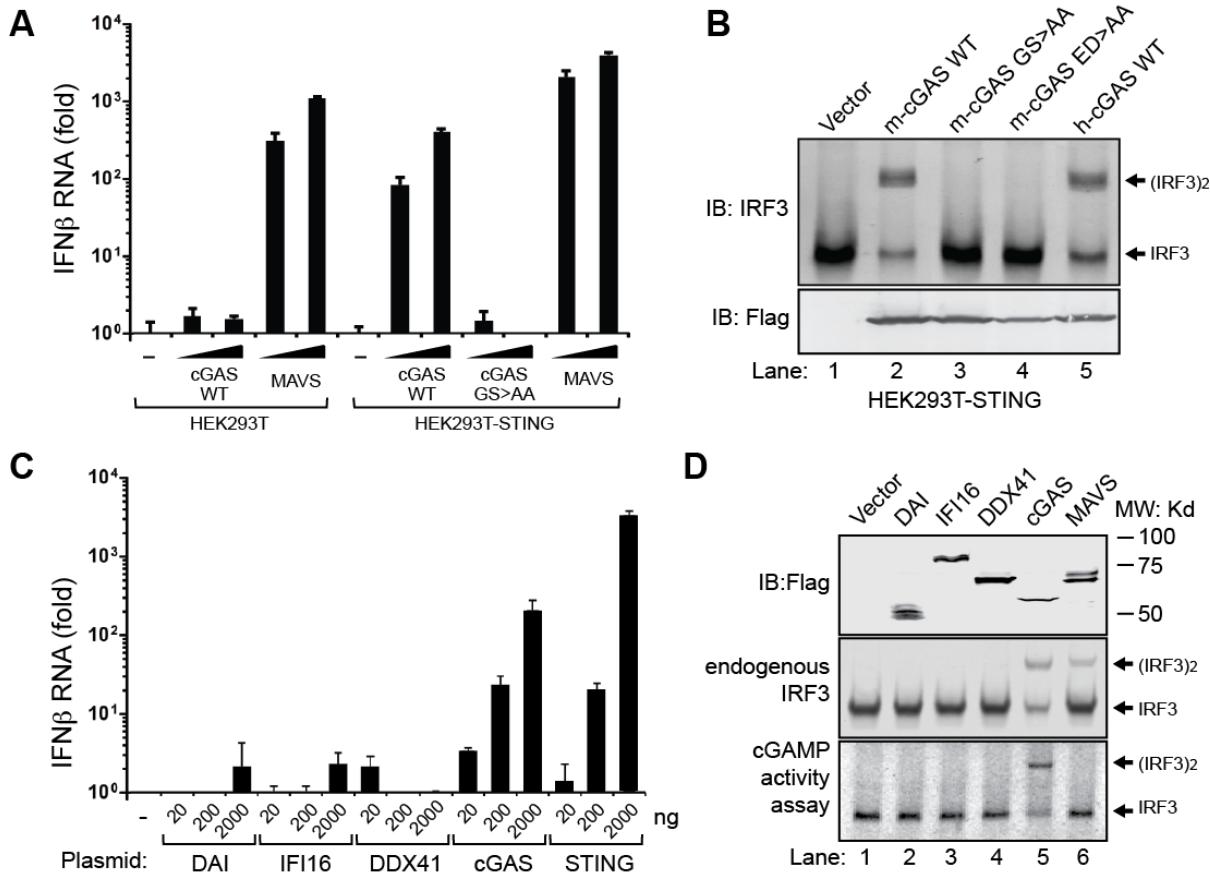


Figure 11. cGAS activates IRF3 and induces IFN-β in a STING-dependent manner

(A) Expression plasmid (100 and 500ng) encoding Flag-tagged mouse cGAS (m-cGAS), its active site mutants G198A and S199A (designated as GS>AA), and MAVS were transfected into HEK293T cells or the same cell line stably expressing STING (HEK293T-STING). IFN-β RNA was measured by qRT-PCR 24 hrs after transfection.

(B) Similar to (A), except that cell lysates were analyzed for IRF3 dimerization by native gel electrophoresis (top). Expression levels of the transfected genes were monitored by immunoblotting with antibody to Flag (bottom). h-cGAS, human cGAS; ED>AA, E211A and D213A in mouse cGAS.

(C) Expression vectors for the indicated proteins were transfected into HEK293T-STING cells, followed by measurement of IFN-β by qRT-PCR.

(D) Cell lysates shown in (C) were immunoblotted with antibodies to Flag and IRF3 after SDS-PAGE and native PAGE, respectively (top two panels). Aliquots of the cell extracts were assayed for the presence of cGAMP activity, which was measured by detecting IRF3 dimerization after delivery into permeabilized Raw264.7 cells (bottom).

Knockdown of cGAS Inhibits DNA and DNA Virus-Induced IRF3 Activation and IFN- β Production

To test if cGAS is required for IRF3 activation and IFN induction by cytosolic DNA, I established two L929 cell lines stably expressing shRNAs that targeted distinct regions of m-cGAS (Figure 12A). The ability of these cells to induce IFN- β in response to HT-DNA was severely compromised relative to a control cell line expressing an shRNA against GFP protein (Figure 12B). To rule out possible off-target effects of these cGAS shRNAs, I expressed cGAS in the L929-shcGAS cells, which restored IFN- β induction (Figure 12C). In contrast, expression of cGAS or delivery of cGAMP failed to induce IFN- β in L929-shSTING cells, whereas expression of STING or MAVS restored IFN- β induction in both shcGAS and shSTING cells (Figure 12, C and D). These results indicate that cGAS functions upstream of STING and is required for IFN- β induction by DNA transfection.

To determine if cGAS is required for DNA virus infection-induced IRF3 activation, I infected L929-shcGAS and control L929-shGFP cells with HSV1 and monitored IRF3 dimer formation after different lengths of time. Cells expressing shRNA against m-cGAS, but not against GFP, were severely compromised in IRF3 dimerization in response to HSV-1 infection (Figure 12E). In contrast, knockdown of cGAS did not affect IRF3 activation by Sendai virus, an RNA virus (Figure 12F).

I also tested the importance of cGAS in the DNA sensing pathway in human cells by establishing a THP1 cell line stably expressing an shRNA targeting h-cGAS (Figure 12G, left). The knockdown of h-cGAS strongly inhibited IFN- β induction by HT-DNA transfection or infection by vaccinia virus, another DNA virus, but not by Sendai virus (Figure 12G, right). The knockdown of h-cGAS also inhibited IRF3 dimerization induced by

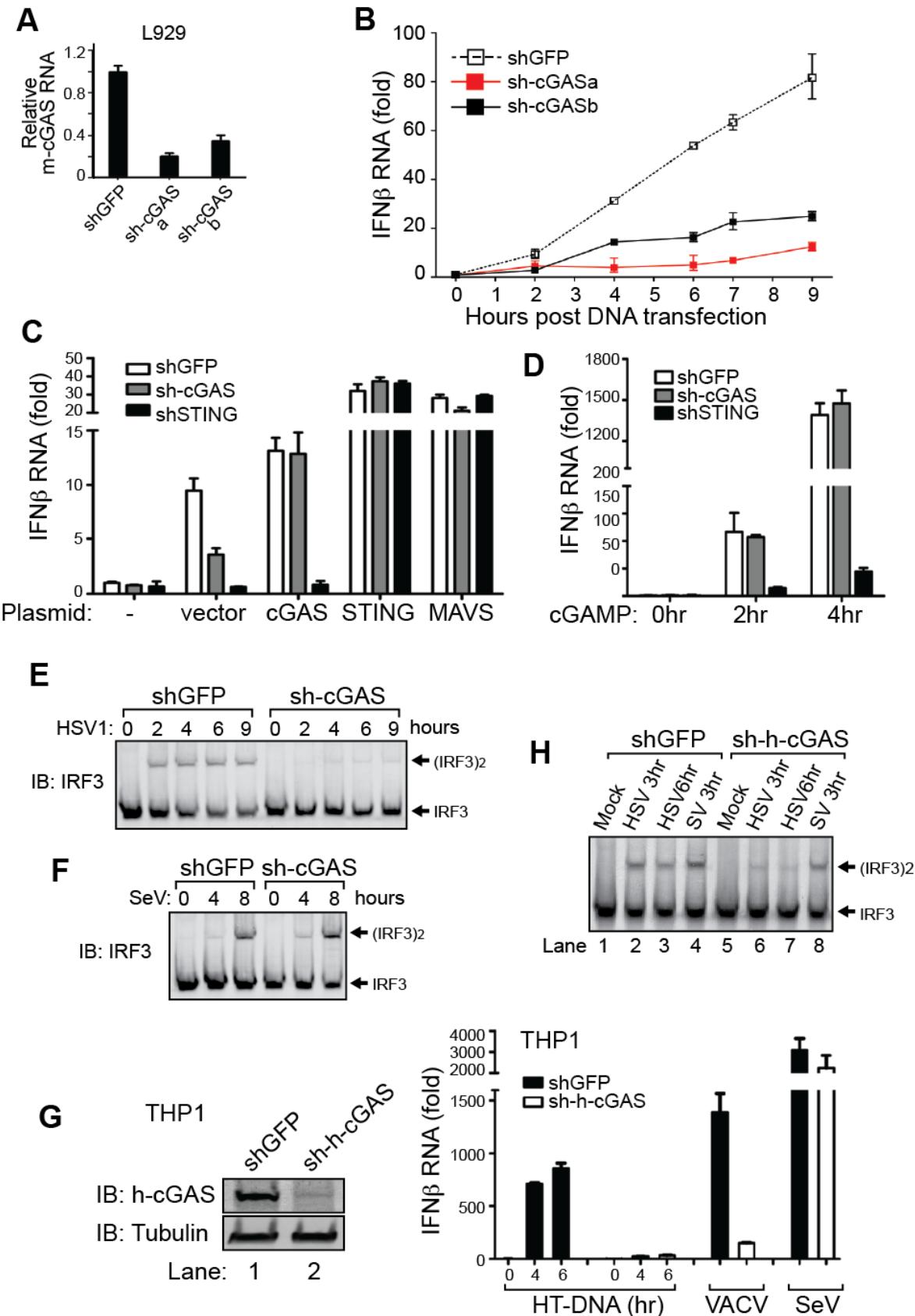


Figure 12. Knockdown of cGAS inhibits DNA and DNA virus-induced IRF3 activation and IFN- β production

- (A) qRT-PCR analyses of m-cGAS RNA in L929 cell lines stably expressing shRNA targeting GFP (control) or two different regions of m-cGAS.
- (B) Cells in (A) were transfected with HT-DNA for the indicated times, followed by measurement of IFN- β RNA by qRT-PCR.
- (C) L929 cells stably expressing shRNA against GFP, cGAS, or STING were transfected with pcDNA3 (vector) or the same vector driving the expression of the indicated proteins. IFN- β RNA was measured by qRT-PCR 24 hrs after transfection.
- (D) cGAMP (100 nM) was delivered to digitonin-permeabilized L929 cells stably expressing shRNA against GFP, cGAS, or STING. IFN- β RNA was measured by qRT-PCR at the indicated times after cGAMP delivery.
- (E and F) L929 cells stably expressing shRNA against GFP or cGAS were infected with HSV1 (Δ ICP34.5) (D) or Sendai virus (SeV) (E) for the indicated times, followed by measurement of IRF3 dimerization.
- (G) THP1 cells stably expressing an shRNA against GFP or human cGAS were transfected with HT-DNA or infected with VACV or SeV. 6 hrs later, IFN- β RNA was measured by qRT-PCR. The knockdown of h-cGAS in cells was assessed by immunoblotting.
- (H) Cells in (G) were infected with HSV1 or SeV for 3hrs or 6 hrs, and then IRF3 dimerization was analyzed by native gel electrophoresis.

HSV-1 infection in THP1 cells (Figure 12H). Collectively, the RNAi data from both mouse and human cell lines demonstrate a key role of cGAS in the STING-dependent DNA sensing pathway.

cGAS is Required for the Generation of cGAMP in Cells

I first measured the relative expression level of cGAS RNA in various cell lines by qRT-PCR, which showed that the expression of m-cGAS was low in immortalized mouse embryo fibroblasts (MEFs) but high in L929 cells, Raw264.7 cells, and BMDMs (Figure 13A). Similarly, the expression of h-cGAS was undetectable in human embryonic kidney (HEK) 293T cells but high in the human monocytic cell line THP1 (Figure 13B). Thus, the expression levels of m-cGAS and h-cGAS in different cell lines are consistent with the ability of these cells to produce cGAMP in response to cytosolic DNA (Figure 5F).

To determine whether cGAS is required for the generation of cGAMP in cells, I transfected HT-DNA into L929-shGFP and L929-shcGAS cells or infected these cells with HSV-1, then prepared heat-resistant fractions that contained cGAMP, which were subsequently delivered to permeabilized Raw264.7 cells to measure IRF3 activation. Knockdown of cGAS largely abolished the cGAMP activity generated by DNA transfection or HSV-1 infection (Figure 13C, bottom). Quantitative mass spectrometry using selective reaction monitoring (SRM) showed that the abundance of cGAMP induced by DNA transfection or HSV-1 infection was markedly reduced in L929 cells depleted of cGAS (Figure 13D). These results demonstrate that cGAS is essential for producing cGAMP and activating IRF3 in response to DNA transfection or HSV-1 infection.

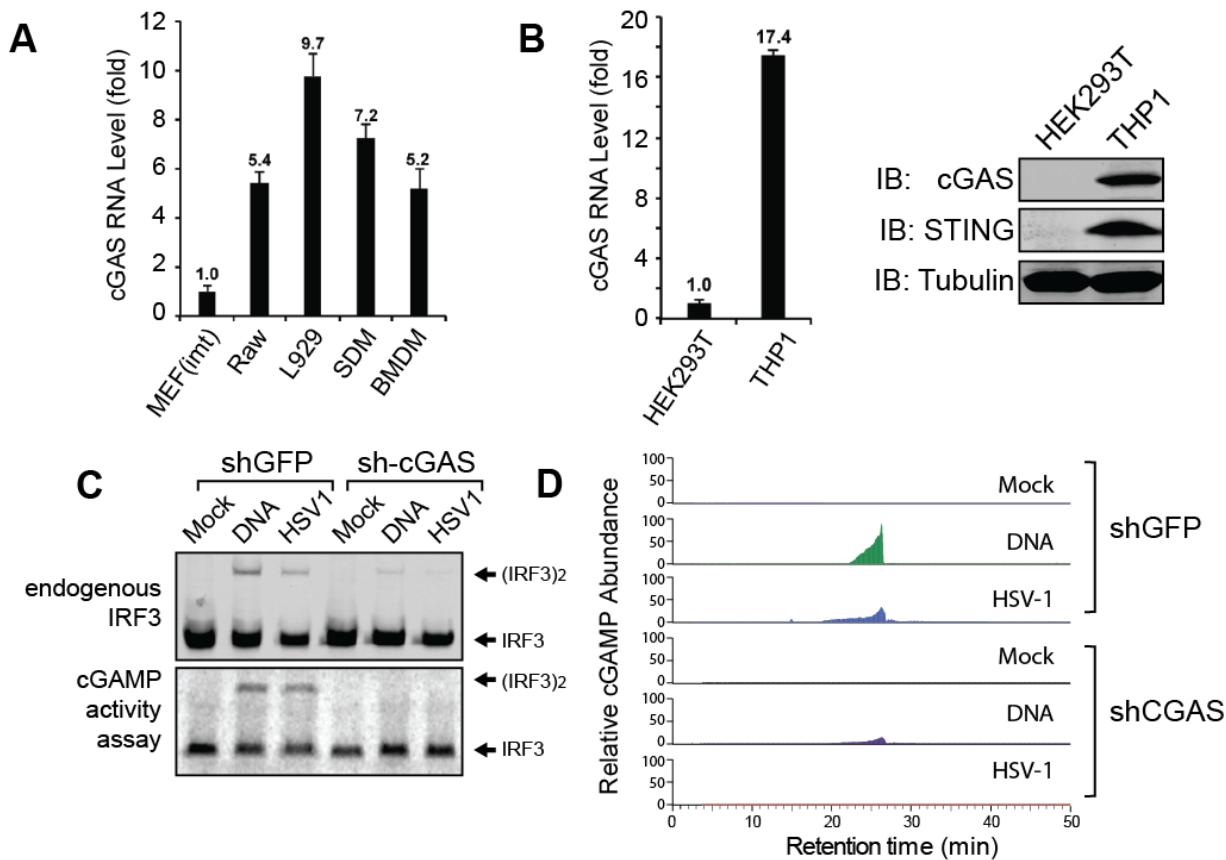


Figure 13. cGAS is required for the generation of cGAMP in cells

- (A) Quantitative RT-PCR analyses of cGAS RNA levels in different murine cell lines. MEF(imt), immortalized MEF; Raw, Raw264.7; SDM, spleen-derived macrophage; BMDM, bone marrow-derived macrophage.
- (B) Quantitative RT-PCR and immunoblotting analyses of endogenous human cGAS RNA and protein levels in HEK293T and THP1 cells with the indicated antibodies.
- (C) L929 cells stably expressing shRNA against GFP or cGAS were transfected with HT-DNA or infected with HSV1 for 6 hrs, followed by measurement of IRF3 dimerization (top). Extracts from these cells were used to prepare heat-resistant supernatants, which were delivered to permeabilized Raw264.7 cells to stimulate IRF3 dimerization (bottom).
- (D) The heat- resistant supernatants in (C) were fractionated by high-performance liquid chromatography using a C18 column; the abundance of cGAMP was quantified by mass spectrometry using SRM.

cGAS is Essential for IRF3 Activation and IFN- β Induction by DNA Transfection and DNA Virus Infection in Fibroblasts and Macrophages

To provide definitive evidence for the essential role of cGAS in cytosolic DNA sensing and study the function of cGAS in vivo, Dr. Xiao-Dong Li, a postdoctoral researcher in the lab, generated a cGAS knockout mouse strain. The cGas^{-/-} mice were born at the Mendelian ratio, and did not show any noticeable developmental defect (Li et al., 2013b). I conducted qPCR analyses of RNA from lung fibroblasts and bone marrow derived macrophages (BMDMs), which confirmed that the cGas^{-/-} cells were defective in producing cGAS RNA, whereas cGas^{+/-} cells produced intermediate levels of cGAS RNA (Figure 14A).

I isolated lung fibroblasts from wild-type (WT), cGas^{+/-}, and cGas^{-/-} mice as well as the Sting goldenticket (gt/gt) mouse, which has a point mutation that results in the loss of expression of STING (Sauer et al., 2011). Transfection of different types of DNA, including HT-DNA, *Escherichia coli* DNA, and ISD, into the lung fibroblasts from WT and cGas^{+/-} mice led to robust production IFN- β protein, as measured by ELISA (Figure 14B). In contrast, the cGas^{-/-} and Sting^{gt/gt} cells failed to produce any detectable level of IFN- β . As a control, dsRNA analog Poly[I:C], which activates RIG-I-like receptor (RLR) pathway,

induced IFN- β in the absence of cGAS or STING. Interestingly, poly[dA:dT], which was previously shown to induce type I interferons through the RNA polymerase III–RIG-I–MAVS pathway (Chiu et al., 2009), induced IFN- β normally in the cGas^{-/-} and Sting^{gt/gt} cells. qPCR analyses further confirmed that cGAS is essential for IFN- β RNA induction by different types of synthetic or bacterial DNA, except poly[dA:dT] (Figure 14C). Time course experiments showed that IFN- β induction by ISD was completely abolished in cGas^{-/-} lung fibroblasts at all the time points tested (Figure 14D), indicating that cGAS is essential for IFN- β induction by cytosolic DNA.

I also infected the lung fibroblasts with DNA viruses HSV1, VACV, and a mutant strain of HSV1 known as d109, which has a deletion of viral proteins such as ICP0 that are known to counteract host immune responses (Samaniego et al., 1998). IFN- β induction by each of these viruses was largely abolished in cGas^{-/-} and Sting^{gt/gt} cells, and drastically decreased in cGas^{+/-} cells (Figure 14E). In contrast, IFN- β induction by RNA virus Sendai virus was not affected by the loss of cGAS or STING. Consistent with the RNAi experiments described above, delivery of cGAMP into the cytoplasm rescued IFN- β induction in cGas^{-/-} cells but not Sting^{gt/gt} cells (Figure 14E). Similarly, induction of the chemokine CXCL10 by the DNA viruses was dependent on cGAS and STING (Figure 14F). Measurement of IRF3 dimerization showed that cGas^{-/-} cells failed to activate IRF3 in response to HT-DNA transfection or infection by WT HSV1 or the HSV1 strain 7134, which also lacks the interferon antagonist ICP0 (Figure 14G) (Melroe et al., 2004). The cGAS deficiency did not impair IRF3 activation by Sendai virus. Thus, cGAS plays an indispensable and specific role for IRF3 activation and IFN induction by DNA viruses in mouse lung fibroblasts.

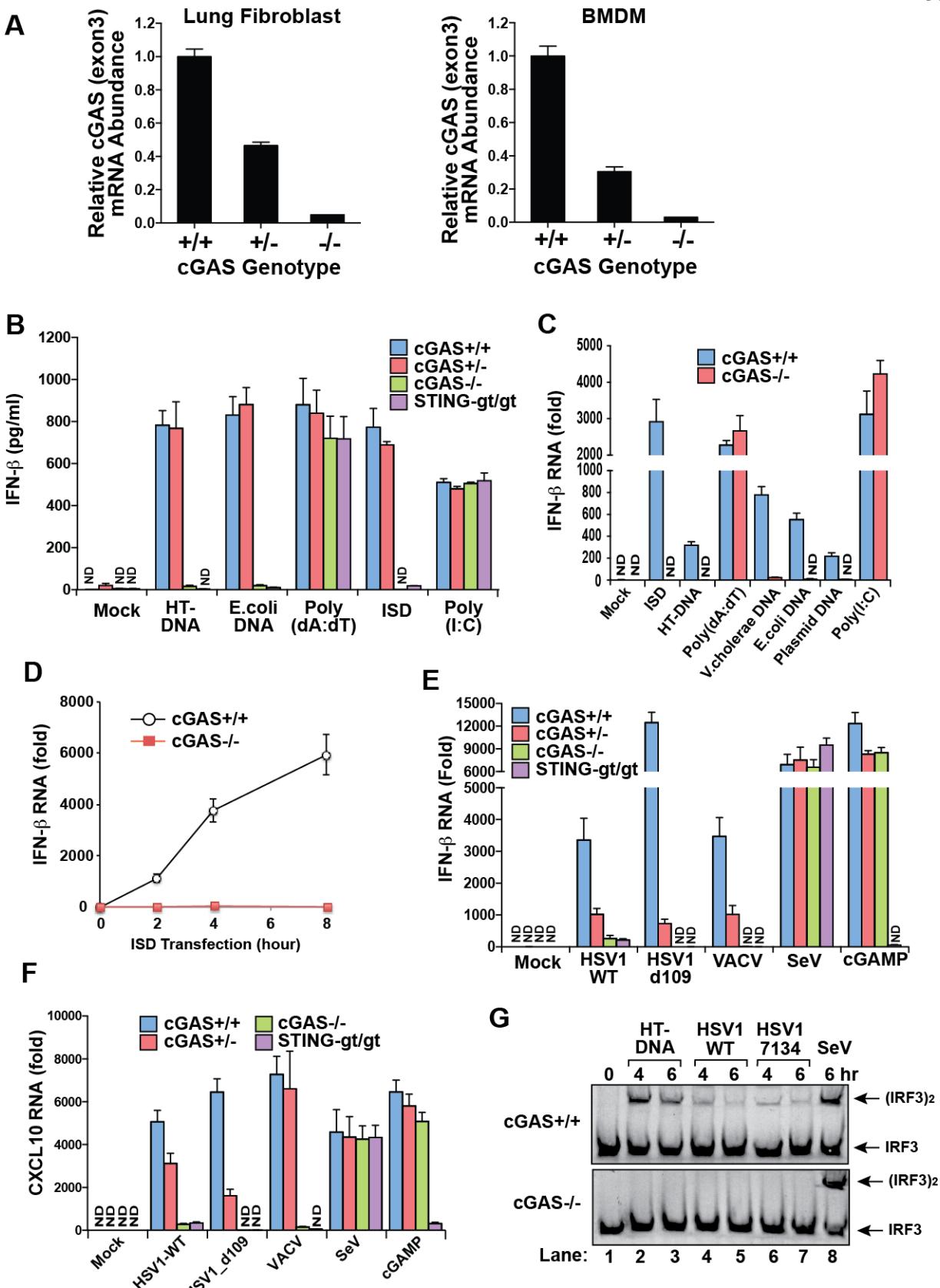


Figure 14. cGAS is essential for IRF3 activation and IFN- β induction by DNA transfection and DNA virus infection in fibroblasts

- (A) qRT-PCR analyses of cGAS RNA amplified from the total RNA of lung fibroblasts or BMDM.
- (B) Mouse lung fibroblasts were transfected with different forms of DNA or poly[I:C] (2 μ g/ml) for 24 hrs, followed by measurement of IFN- β protein by ELISA. Unless indicated otherwise, lipofectamine 2000 was used in all transfection experiments.
- (C) Mouse lung fibroblasts were transfected with different forms of nucleic acids (2 μ g/ml) for 10 hrs followed by measurement of IFN- β RNA by qRT-PCR.
- (D) Lung fibroblasts were transfected with ISD (3 μ g/ml) for indicated lengths of time followed by measurements of IFN- β RNA by qRT-PCR.
- (E and F) Lung fibroblasts were infected with the indicated viruses or stimulated with 2'3'-cGAMP (200 nM) for 9 hrs followed by measurement of IFN- β (E) or CXCL10 (F) RNA by qRT-PCR.
- (G) Lung fibroblasts were transfected with HT-DNA or infected with HSV1 or Sendai virus for the indicated times. Cell extracts were analyzed for IRF3 dimerization by native gel electrophoresis.

BMDMs from cGas^{-/-} and Sting^{gt/gt} mice were also incapable of producing IFN- β in response to HT-DNA or ISD transfection (Figure 15A). IFN- β induced by VACV and the HSV1 strains d109 and 7134 was largely diminished in cGas^{-/-} and Sting^{gt/gt} BMDMs. Interestingly, IFN- β induction by WT HSV1 was severely but not completely abrogated in either cGas^{-/-} or Sting^{gt/gt} BMDMs, suggesting the existence of another pathway that could partially compensate for the loss of the cGAS-STING pathway to detect WT HSV1 infection. The loss of cGAS or STING in BMDMs did not affect IFN- β induction by Sendai virus. Time course experiments showed that IFN- β induction by ISD and HSV1-d109 was abolished in cGas^{-/-} BMDMs at all time points tested (Figure 15, B and C). Similarly to IFN- β , the induction of TNF α by HT-DNA or ISD was abolished in cGas^{-/-} or Sting^{gt/gt} BMDMs (Figure 15D). qPCR analyses showed that the induction of IFN- β , interleukin-6 (IL-6), and CXCL10 RNA by transfection of HT-DNA, ISD, or infection with HSV1-d109 completely relied on cGAS and STING (Figure 15, E to G). In contrast, the RNA levels of these cytokines induced by poly[I:C] or Sendai virus were not affected by the loss of cGAS or

STING. Thus, cGAS plays an essential and specific role in DNA or DNA virus-induced production of IFN and other cytokines in BMDMs.

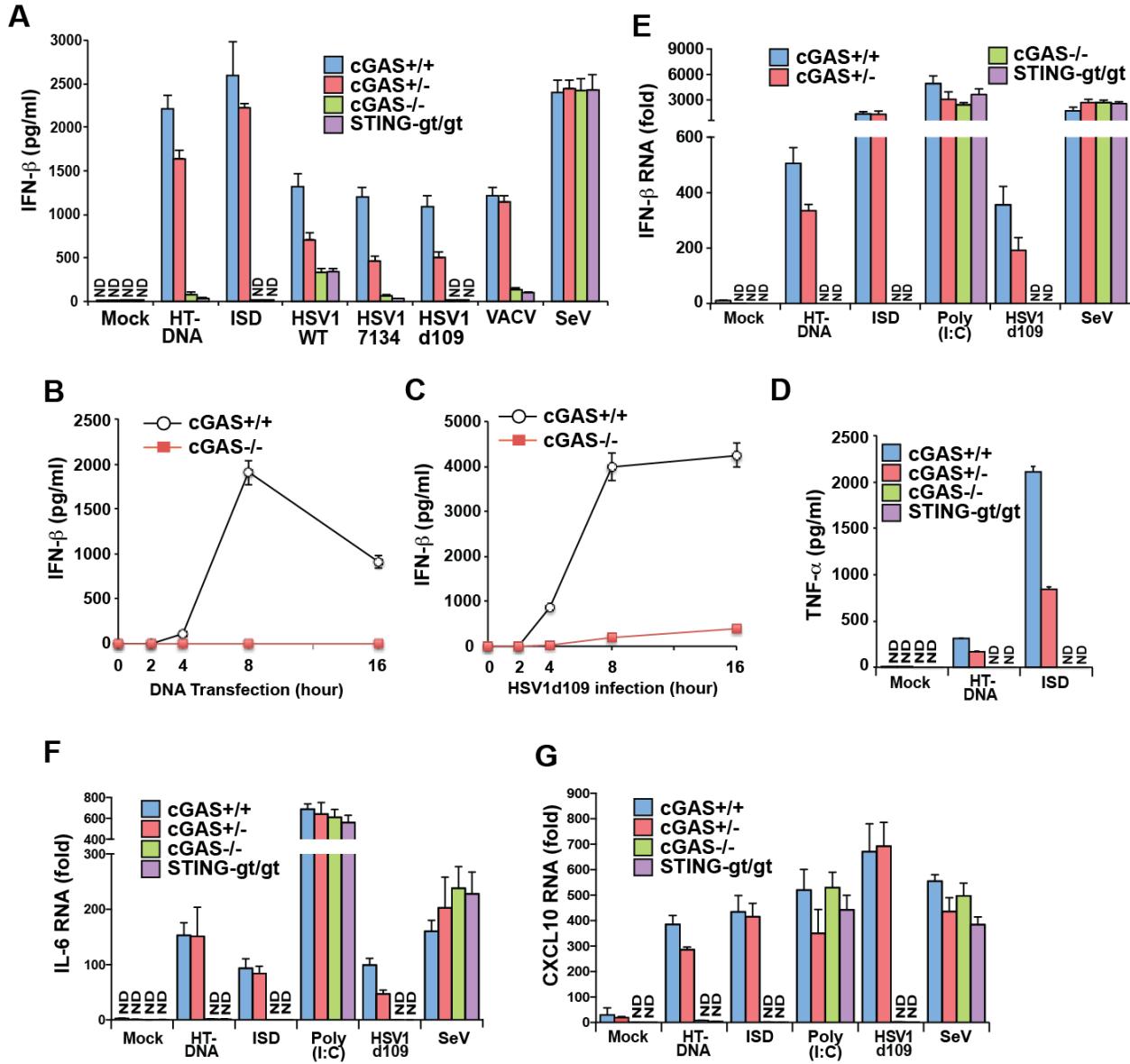


Figure 15. cGAS is essential for IFN and cytokine induction by DNA transfection and DNA virus infection in macrophages

(A) BMDMs were transfected with HT-DNA or ISD or infected with the indicated virus for 20 hrs, and then IFN- β levels were measured by ELISA.

(B and C) BMDMs were transfected with ISD (B) or infected with HSV1-d109 (C) for the indicated time points followed by measurements of IFN- β secretion by ELISA.

(D) BMDMs were transfected with HT-DNA or ISD for 20 hrs followed by measurement of TNF α by ELISA.

(E-G) BMDMs were transfected with HT-DNA, ISD, poly(I:C) or infected with HSV1-d109 or Sendai virus for 9 h, and then the levels of IFN- β (E), IL-6 (F) and CXCL10 (G) RNA were measured by qRT-PCR.

cGAS is Essential for Type I Interferon Induction by DNA or DNA Virus Stimulation in Dendritic Cells

Conventional dendritic cells (cDC) and plasmacytoid DCs (pDC) were obtained by culturing bone marrow in conditioned media containing granulocyte-macrophage colony-stimulating factor (GM-CSF) and Flt3 ligand (Flt3L), respectively. The GM-CSF DCs, which majorly contain cDCs, from the cGas^{-/-} and Sting^{gt/gt} mice are incapable of producing IFN- α or IFN- β in response to transfection of HT-DNA or ISD (Figure 16, A and B). Similar to BMDMs, deficiency in cGAS or STING in GM-CSF DCs led to diminished IFN- β induction by HSV1-d109 and VACV and partially inhibited IFN- β induction by WT HSV1. In contrast, the lack of cGAS or STING did not reduce IFN- α or IFN- β induction by Sendai virus. qPCR experiments further confirmed that cGAS and STING were essential for the induction of IFN- β , IL-6, and CXCL10 RNA by transfection with HT-DNA or ISD, or infection with HSV1-d109, whereas the induction of these cytokines by poly[I:C] or Sendai virus remained normal in cGAS or STING deficient cDCs.

pDCs are known to express TLR9 that is responsible for the induction of type I interferons by synthetic CpG DNA containing phosphorothioate backbone (Akira et al., 2006). As expected, when the CpG DNA was used to stimulate Flt3L DCs, which contain largely pDCs, in the presence or absence of liposome (lipofectamine 2000), it induced robust production of IFN- α and IFN- β even in the cGas^{-/-} and Sting^{gt/gt} cells (Figure 16. C and D). In contrast, other forms of DNA, including ISD, and genomic DNA from *E. coli* and *Vibrio cholerae*, induced IFN- α in pDCs only in the presence of liposome, and this induction was totally dependent on cGAS and STING. It is worth noting that, unlike in fibroblasts, poly[dA:dT]-induced IFN- α induction also completely relies on cGAS and STING in pDCs,

suggesting that the cGAS-STING pathway, but not the Pol-III–RIG-I pathway, plays a dominant role in sensing DNA in these cells. The Flt3L DCs from the *cGas*^{-/-} and *Sting*^{gt/gt} mice induced IFN- α and IFN- β in response to infection by Sendai virus but not HSV1 (Figure 16, D and E). Together, these results demonstrate that cGAS is responsible for detecting various natural DNA species (e.g., bacterial DNA) and DNA virus infection in dendritic cells.

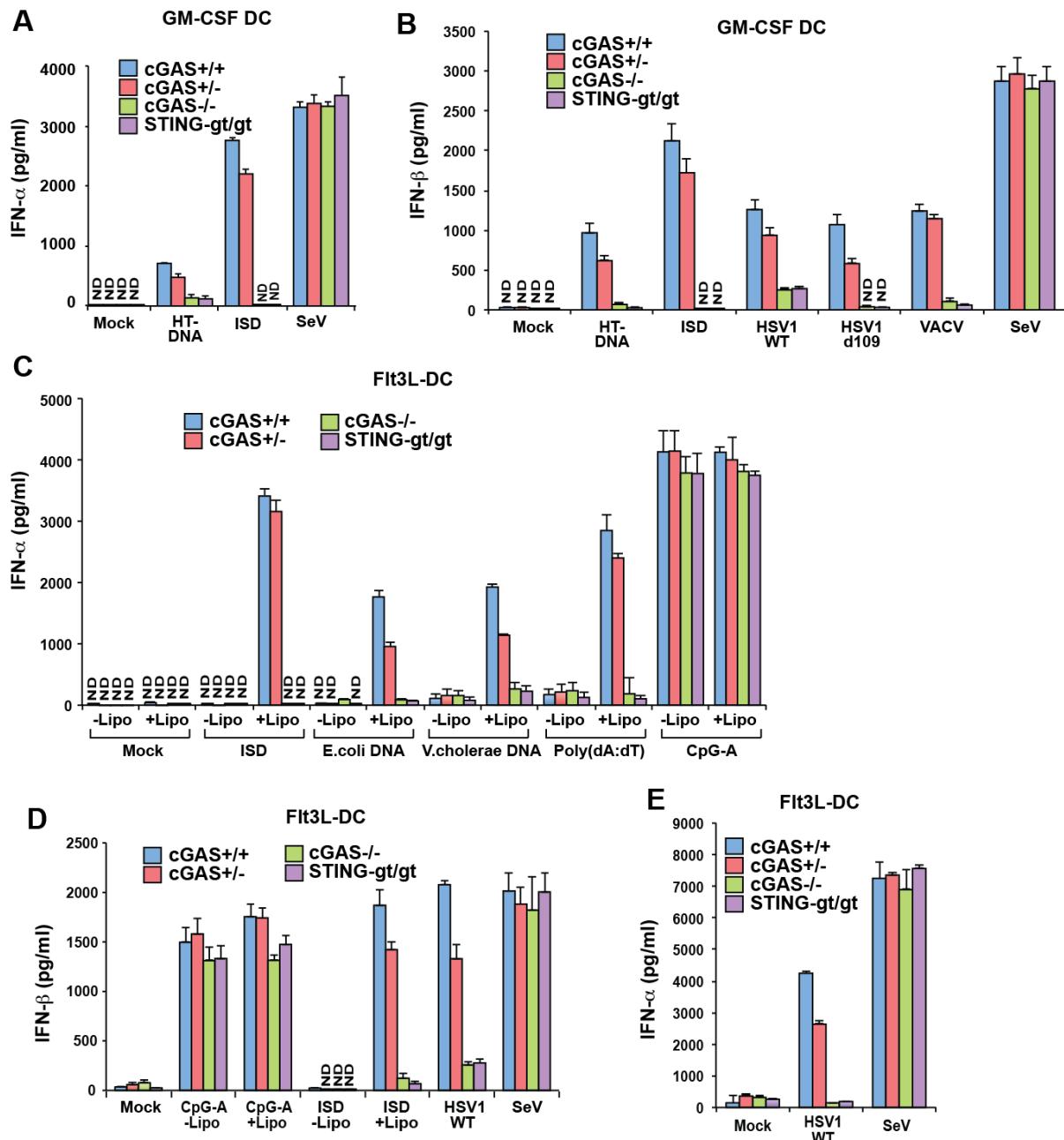


Figure 16. cGAS is essential for IFN induction by DNA transfection and DNA virus infection in dendritic cells

- (A) GM-CSF-induced DCs were transfected with HT-DNA or ISD or infected with Sendai virus for 20 hrs, followed by measurement of IFN- α by ELISA.
- (B) GM-CSF DCs were transfected with the indicated DNA or viruses for 20 hrs, and then IFN- β was measured by ELISA.
- (C) Flt3L-induced dendritic cells (Flt3L-DCs) were incubated with the indicated DNA for 16 hrs in the absence or presence of lipofectamine, followed by measurement of IFN- α by ELISA.
Lipo, Lipofectamine 2000.
- (D) Flt3L-DCs were incubated with CpG DNA and ISD as indicated, or infected with HSV1 or Sendai virus for 16 hrs, followed by measurement of IFN- β by ELISA.
- (E) Flt3L-DCs were infected with HSV1 or Sendai virus for 16 hrs, followed by measurement of IFN- α by ELISA.

Recombinant cGAS Protein Catalyzes cGAMP Synthesis from ATP and GTP in a DNA-Dependent Manner

To test whether cGAS alone is sufficient to catalyze cGAMP synthesis, I expressed Flag-tagged h-cGAS in HEK293T cells and purified it to a single band on a silver stained gel (Figure 17A). The purification contained two steps including a flag pulldown, followed by heparin chromatography. In the presence of HT-DNA, purified c-GAS protein catalyzed the production of cGAMP activity, which stimulated IRF3 dimerization in permeabilized Raw264.7 cells (Figure 17B). The activity of cGAS was dependent on DNA, as including DNase I in the reaction abolished cGAMP production. The cGAS activity was also stimulated by other DNAs, including poly[dA:dT], poly[dG:dC], and ISD, but not by the RNA poly[I:C] (Figure 17B). To determine the substrate requirement of cGAS, I set up cGAS reactions in presence of different combinations of nucleotide triphosphates. The synthesis of cGAMP by cGAS required both ATP and GTP but not cytidine triphosphate (CTP) or uridine triphosphate (UTP) (Figure 17C). These results indicated that the cyclase activity of purified cGAS protein was stimulated by DNA but not by RNA, and ATP and GTP were the substrates utilized by cGAS. To provide formal evidence that cGAS is

sufficient to catalyze the synthesis of cGAMP, I analyzed the reaction products by nano-LC-MS using SRM. cGAMP was detected in a 60-min reaction containing purified cGAS, ATP, and GTP (Figure 17D). The identity of cGAMP was further confirmed by ion fragmentation using collision-induced dissociation (CID).

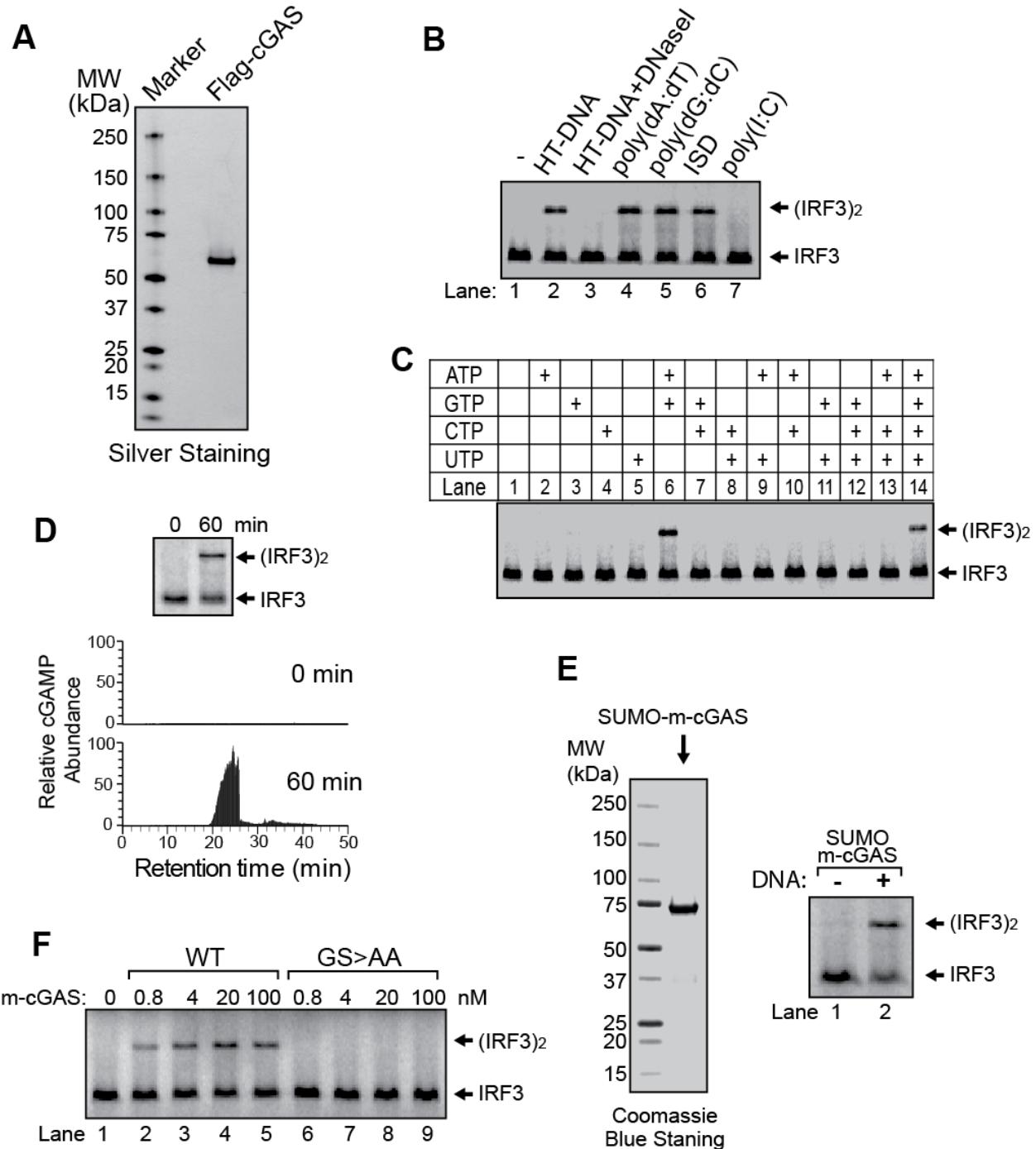


Figure 17. Recombinant cGAS protein catalyzes cGAMP synthesis from ATP and GTP in a DNA-dependent manner

- (A) Silver staining of Flag-h-cGAS expressed and purified from HEK293T cells.
- (B) Purified Flag-h-cGAS as shown in (A) was incubated with ATP and GTP in the presence of different forms of nucleic acids as indicated. Generation of cGAMP was assessed by its ability to induce IRF3 dimerization in Raw264.7 cells.
- (C) Similar to (B), except that reactions contained HT-DNA and different combinations of nucleotide triphosphates as indicated.
- (D) Purified m-cGAS from *E. coli* was incubated with ATP, GTP, and DNA for 0 or 60 min, and the production of cGAMP was analyzed by IRF3 dimerization assay (top) and mass spectrometry using SRM (bottom).
- (E) Coomassie blue staining of Sumo-m-cGAS expressed and purified from *E. coli* (left). The protein was incubated with ATP and GTP in the presence or absence of HT-DNA, and then the reaction mixtures were heated to prepare heat-resistant supernatants, which were delivered to permeabilized Raw264.7 cells to measure IRF3 activation (right).
- (F) Similar to (B), except that wild-type and mutant cGAS proteins were expressed and purified from *E. coli* and assayed for their activities at the indicated concentrations.

I also expressed sumo tagged m-cGAS in *E. coli*. After purification, m-cGAS generated the cGAMP activity in a DNA-dependent manner (Figure 17E). Titration experiments showed that recombinant cGAS protein can generate cGAMP activity at concentrations less than 1 nM, whereas the catalytically inactive mutant of cGAS failed to produce cGAMP activity at all concentrations tested (Figure 17F).

cGAS Binds to DNA in the Cytoplasm

The fact that the activity of cGAS is strictly dependent on DNA suggests that cGAS is a DNA sensor (Figure 17B). To test whether cGAS bound to DNA directly, I conducted a pulldown experiment using biotinylated ISD-conjugated streptavidin beads. Both GST (glutathione S-transferase) tagged m-cGAS (GST-m-cGAS) and GST-h-cGAS, but not GST– RIG-I N terminus [RIG-I(N)], were precipitated by biotinylated ISD (Figure 18A). In a complementary experiment, I also showed that only biotinylated ISD, but not biotinylated RNA could bind cGAS (Figure 18B).

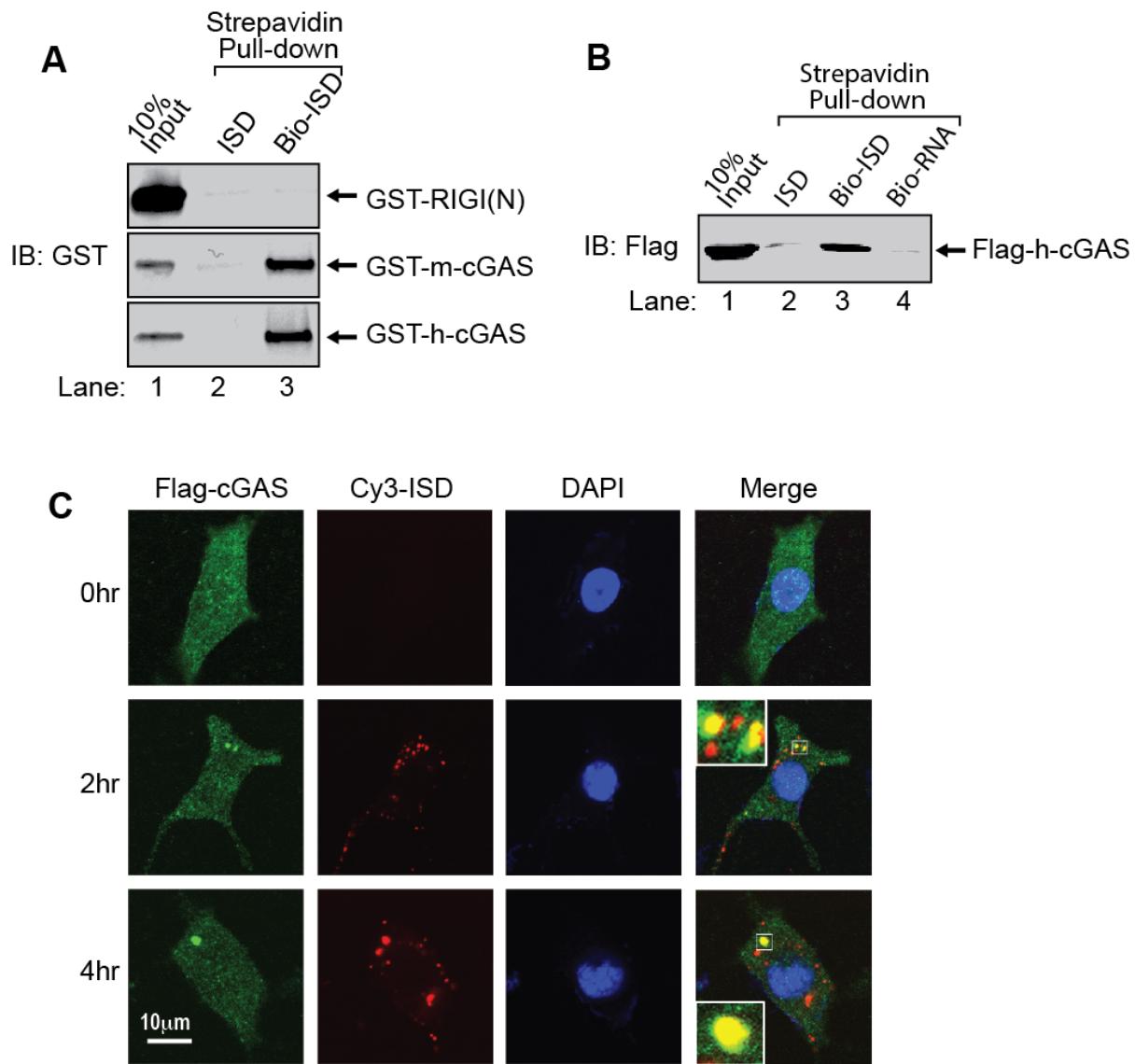


Figure 18. cGAS binds to DNA in the cytoplasm

(A) The indicated GST fusion proteins were expressed and purified from *E. coli* and then incubated with streptavidin beads in the presence of ISD or biotin-ISD. Bound proteins were eluted with SDS sample buffer and detected by immunoblotting with a GST antibody.

(B) Flag-h-cGAS was expressed and purified from HEK293T cells and then incubated with streptavidin beads in the presence of ISD, biotin-ISD or biotin-RNA. Bound proteins were eluted with SDS sample buffer and detected by immunoblotting with a Flag antibody.

(C) L929 cells stably expressing Flag-cGAS (green) were transfected with Cy3-ISD (red). At different time points after transfection, cells were fixed, stained with antibody to Flag or with DAPI, and imaged by confocal fluorescence microscopy. The insets in the merged images are magnifications of the small-boxed areas.

I also examined the localization of cGAS by confocal immunofluorescence microscopy of L929 cells stably expressing Flag-m-cGAS (Figure 18C). The cGAS protein was distributed throughout the cytoplasm but could also be observed in the nuclear or perinuclear region. After the cells were transfected with Cyanine 3 (Cy3)-labeled ISD for 2 or 4 hrs, punctate forms of cGAS were observed, and they overlapped with the DNA fluorescence. Such colocalization and apparent aggregation of cGAS and Cy3-ISD was observed in more than 50% of the cells under observation. These results, together with the biochemical evidence of direct binding of cGAS to DNA, suggest that cGAS binds DNA in the cytoplasm.

Determination of the Phosphodiester Linkages in the cGAS Product

In nature, a phosphodiester bond connecting nucleotides has either a 2'-5' or 3'-5' linkage. While the experiments described above clearly demonstrated that cGAMP is an endogenous second messenger produced by cGAS in mammalian cells, the exact nature of the two internal phosphodiester linkages between GMP and AMP in cGAMP was not determined, in part because mass spectrometry alone could not unambiguously distinguish these linkages without comparing to the spectra of standards. To provide standard reference for the determination of cGAMP structure, I decided to chemically synthesize cGAMP molecules containing all four possible phosphodiester linkage combinations (Figure 19A). The chemical synthesis of cGAMP isomers was performed by Dr. Heping Shi in Professor Chuo Chen's Lab in the department of Biochemistry. For simplicity, we named these cGAMP molecules according to the OH position of GMP followed by the OH position of AMP that form the phosphodiester bonds; for example, 2'3'-cGAMP contains a

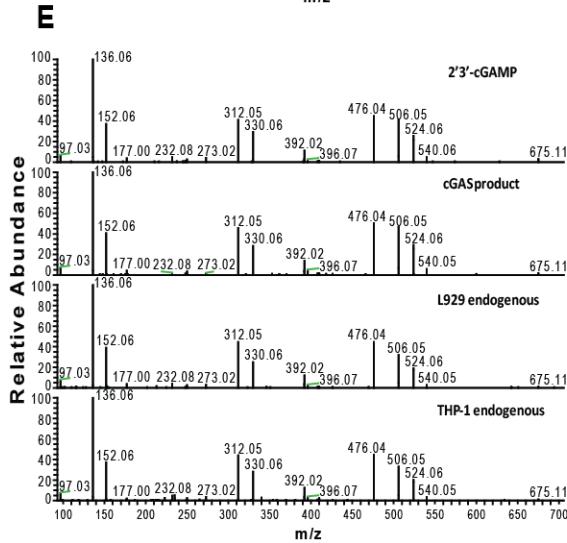
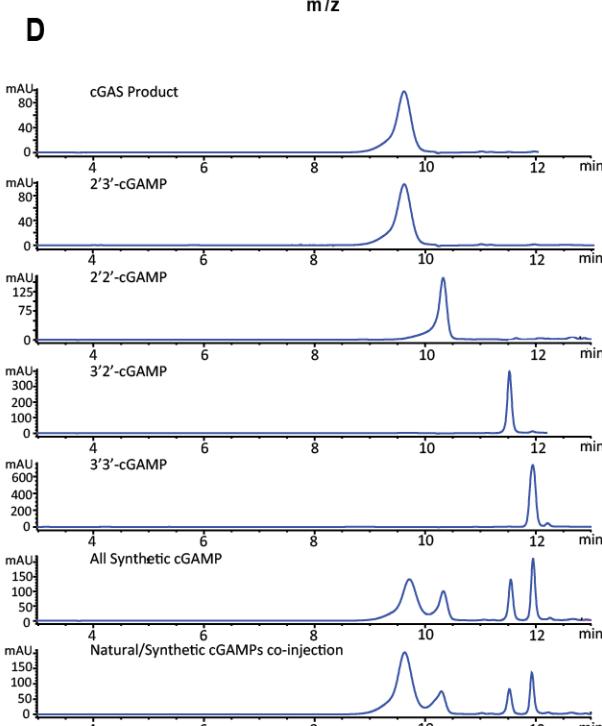
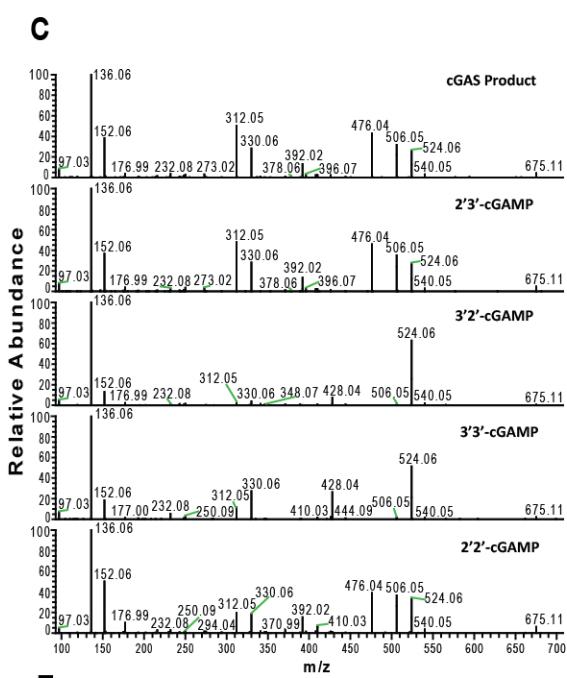
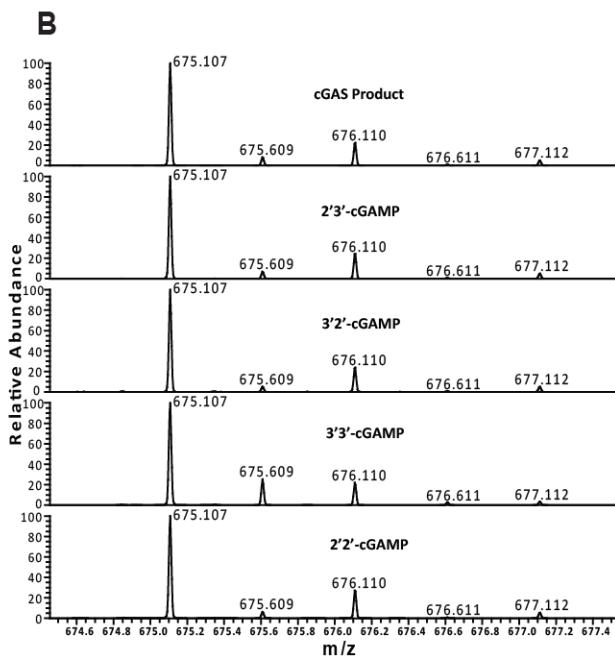
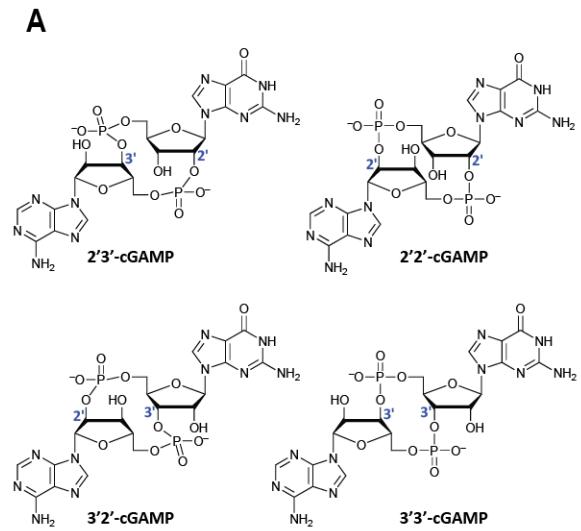


Figure 19. Determination of the phosphodiester linkages in cGAMP

- (A) Depiction of cGAMPs containing different combinations of phosphodiester bonds.
(B) Isotopically resolved high-resolution mass spectra of singly charged ($[M+H]^+$) ions were acquired using a Q Exactive mass spectrometer.
(C) Tandem mass spectra of natural and synthetic cGAMPs resulting from higher-energy collision dissociation (HCD) of the precursor ion ($[M+H]^+ = 675.107$).
(D) Reverse phase HPLC elution profiles of natural and synthetic cGAMPs on a C18 column.
(E) L929 and THP1 cells were transfected with HT-DNA, and the cell extracts containing endogenous cGAMP were analyzed by tandem mass spectrometry along with the cGAS product and synthetic 2'3'-cGAMP.

phosphodiester linkage between 2'-OH of GMP and 5'-phosphate of AMP, and another between 3'-OH of AMP and 5'-phosphate of GMP. We also used purified cGAS protein to enzymatically synthesize the natural cGAMP in vitro from ATP and GTP in the presence of DNA. The purified cGAS product and synthetic cGAMP isomers were analyzed by mass spectrometry.

The total mass of each of these singly charged molecules ($[M+H]^+$) was 675.107, exactly matching the theoretical mass of cGAMP (Figure 19B). The tandem mass (MS/MS) spectra of the cGAS product, which was fragmented using higher-energy collision dissociation (HCD), were identical to those of synthetic 2'3'-cGAMP, and similar but not identical to those of 2'2'-cGAMP and 3'3'-cGAMP (Figure 19C). The MS/MS spectra of 3'2'-cGAMP appeared to be most distinct from those of 2'3'-cGAMP and the cGAS product. Reverse-phase HPLC analysis showed that the cGAS product co-eluted with 2'3'-cGAMP, but not other cGAMP isomers (Figure 19D). Collectively, these results demonstrated that cGAS synthesizes 2'3'-cGAMP in vitro. To test whether mammalian cells could produce endogenous cGAMP that contains the mixed phosphodiester linkages, I transfected mouse L929 cells and human THP1 cells with HT-DNA, heated cell lysates at 95°C to denature proteins, and prepared supernatants for analysis of endogenous cGAMP by mass

spectrometry. The MS/MS spectra of the endogenous molecule from both cell lines were identical to those of cGAS product and 2'3'-cGAMP, indicating that the endogenous second messenger is 2'3'-cGAMP (Figure 19E).

2'3'-cGAMP is a High-Affinity Ligand for STING

To measure the binding affinity (K_d) between STING and different cGAMP isomers, Dr. Xu Zhang and I performed isothermal titration calorimetry (ITC) experiments. The C-terminal domain (CTD) containing residues 139–379 of human STING, which was previously shown to mediate binding to the bacterial second messenger cyclic di-GMP (Burdette et al., 2011; Yin et al., 2012), was expressed in *E. coli* and purified to apparent homogeneity for the ITC experiments. In agreement with previous reports (Ouyang et al., 2012; Shu et al., 2012; Yin et al., 2012), we found that c-di-GMP bound to STING with a K_d of 1.21 μM (Figure 20A and Table 1). Remarkably, the measured K_d values for STING binding to cGAS product and synthetic 2'3'-cGAMP were 4.59 nM and 3.79 nM, respectively (Table 1), representing affinities that are several orders of magnitude higher than that of c-di-GMP. It is worth noting that such a high affinity made the curve fitting difficult in the direct titration assay, and we had to use a competition assay to calculate the accurate K_d values. This assay was conducted by titrating different amounts of cGAS product and synthetic 2'3'-cGAMP as competitors into the STING-c-di-GMP complex (Figures 20B and 20C). We also obtained the K_d values of 3'2'-, 2'2'- and 3'3'-cGAMP binding to STING, which are 1.61 μM , 287 nM and 1.04 μM , respectively (Table 1). Thus, the K_d of 2'3'-cGAMP was ~300-fold lower than that of c-di-GMP, 3'2'-cGAMP, and 3'3'-cGAMP, and ~75-fold lower than that of 2'2'-cGAMP. In addition, unlike the binding of c-di-GMP, which

is an exothermic process, the binding of 2'3'-cGAMP to STING was endothermic, suggesting that the energy may be used for STING conformational changes. Indeed, the crystal structure of 2'3'-cGAMP bound to STING, which was solved by Dr. Xu Zhang, revealed a dramatic conformational rearrangement compared to the STING-c-di-GAMP complex structure (Zhang et al., 2013).

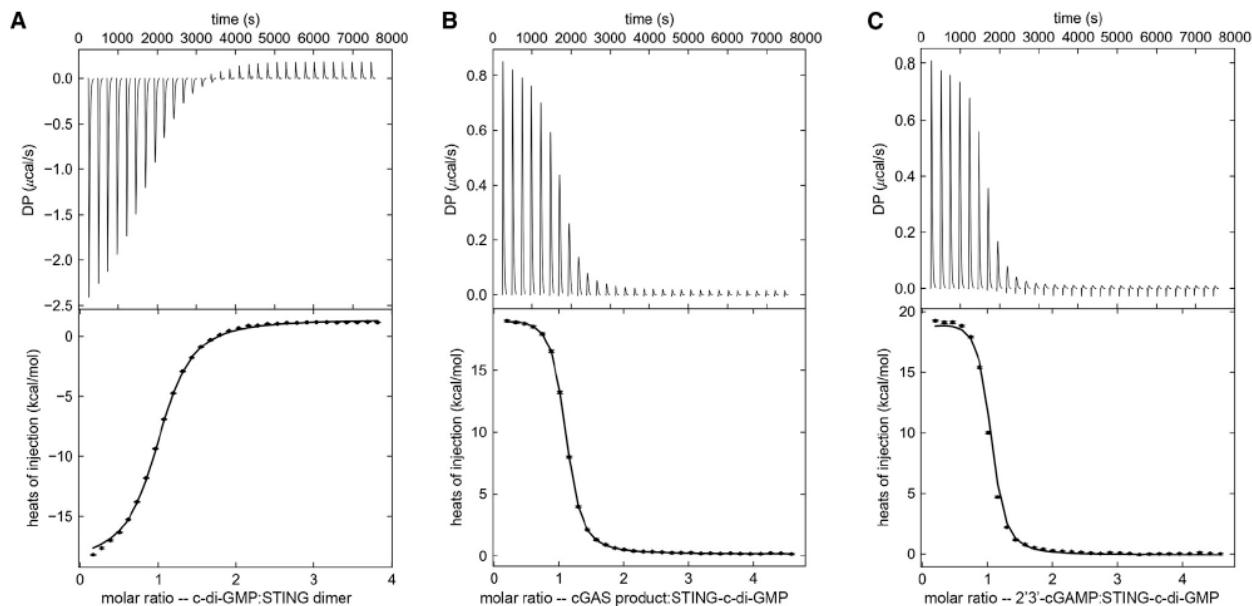


Figure 20. Isothermal titration calorimetry measurements of cGAMP binding to STING

(A) The original titration traces (top) and integrated data (bottom) of ITC experiments, in which c-di-GMP was titrated into a solution of STING dimer.

(B) cGAS product was titrated into a solution containing preformed c-di-GMP-STING complex as a competitive binding assay. (The tight binding of 2'3'-cGAMP to STING made it difficult to do curve fitting if it were directly titrated into apo-STING dimer.)

(C) Same as in (B), except that synthetic 2'3'-cGAMP was used for titration.

ligand	c-di-GMP	cGAS product	2'3'-cGAMP	3'2'-cGAMP	2'2'-cGAMP	3'3'-cGAMP
STING dimer concentration (μM)	25	7.5	7.5	25	25	25
ligand concentration (μM)	500	180	180	500	500	500
molar ratio (c-di-GMP:STING dimer, pre-bound)	-	3.55:1	3.55:1	3.55:1	-	-
K_d	1.21 μM	4.59 nM	3.79 nM	1.61 μM	287 nM	1.04 μM
68.3% interval	[1.04 μM, 1.40 μM]	[3.42 nM, 8.49 nM]	[2.55 nM, 6.71 nM]	[0.59 μM, 84.2 μM]	[262 nM, 316 nM]	[0.98 μM, 1.14 μM]
ΔH	-20.19	1.34	0.71	0.03	14.02	-4.55

Table 1. Binding affinities of STING to different cyclic dinucleotides, as measured by ITC. For c-di-GMP, synthetic 2'2'- and 3'3'-cGAMP, the ligands were titrated into free STING dimer. For others, the ligands were titrated into c-di-GMP-bound STING (3.55:1 in molar ratio).

cGAS is Activated by DNA-Induced Dimerization

Several months after I discovered cGAS, several groups reported the structures of cGAS alone or in complex with DNA (Civril et al., 2013; Gao et al., 2013b; Kranzusch et al., 2013), which provided important information for studying the mechanism of cGAS activation by DNA. To better understand how DNA binding activates cGAS, Dr. Xu Zhang and I solved the crystal structure of mouse cGAS in complex with DNA. The same mouse cGAS protein and DNA were previously used to obtain the crystal structure of a cGAS-DNA complex, which was interpreted to contain one cGAS and one DNA (Gao et al., 2013b). Surprisingly,

through careful structural analysis, we found that cGAS forms a 2:2 complex with dsDNA instead of the 1:1 complex described in previous studies (Figure 21A) (Civril et al., 2013; Gao et al., 2013b). In addition to the primary DNA binding surface (surface 1), which is identical to the ones reported previously (Civril et al., 2013; Gao et al., 2013b), cGAS has another surface area, composed of two positive patches, that interacts with DNA (Figure 21B). This higher order complex composed of dimeric cGAS bound to two molecules of DNA was also reported in a concurrent study (Li et al., 2013a).

To rule out the possibility that this 2:2 complex is caused by crystal packing artifacts, I also set up AUC experiments in collaboration with Dr. Chad Brautigam in the Department of Biophysics. In the absence of DNA, m-cGAS existed as a monomer in solution (Figure 21C). However, in the presence of a 16-mer dsDNA, higher-order complexes are present at 4.0 S, 5.3 S, and 6.5 S (Figure 21D). The latter peak is close to that expected for a 2:2 complex. Dr. Brautigam also carried out a multisignal SV (Balbo et al., 2005; Padrick et al., 2010) analysis, which demonstrated that the molar ratio of protein:DNA in this peak is 0.9:1. Thus, cGAS also formed a 2:2 complex with DNA in solution. In support of this, a single mutation of residue Lys335, which is involved in both cGAS-cGAS and cGAS-DNA interactions, to Glu was sufficient to disrupt the formation of the fast-sedimenting 2:2 complex (Figure 21D).

To further prove the functional importance of this 2:2 complex, I mutated several DNA binding residues within the newly observed secondary binding interface as well as the dimeric interface in human cGAS. Point mutations of Lys347 (Lys335 in m-cGAS), Arg353, or Lys394 to Glu in human cGAS completely abolished its ability to activate IRF3 and induce IFN- β (Figure 21E). Although single mutations of Arg236, Lys254, and Lys327 to

Glu only partially impaired the activity of cGAS, double (R236E/K254E and K254E/K327E) and triple (R236E/K354E/K327E) mutations abolished the activity. These results indicate that both the second DNA binding site and the dimeric interface in cGAS play critical roles in DNA-induced activation of the type I IFN pathway.

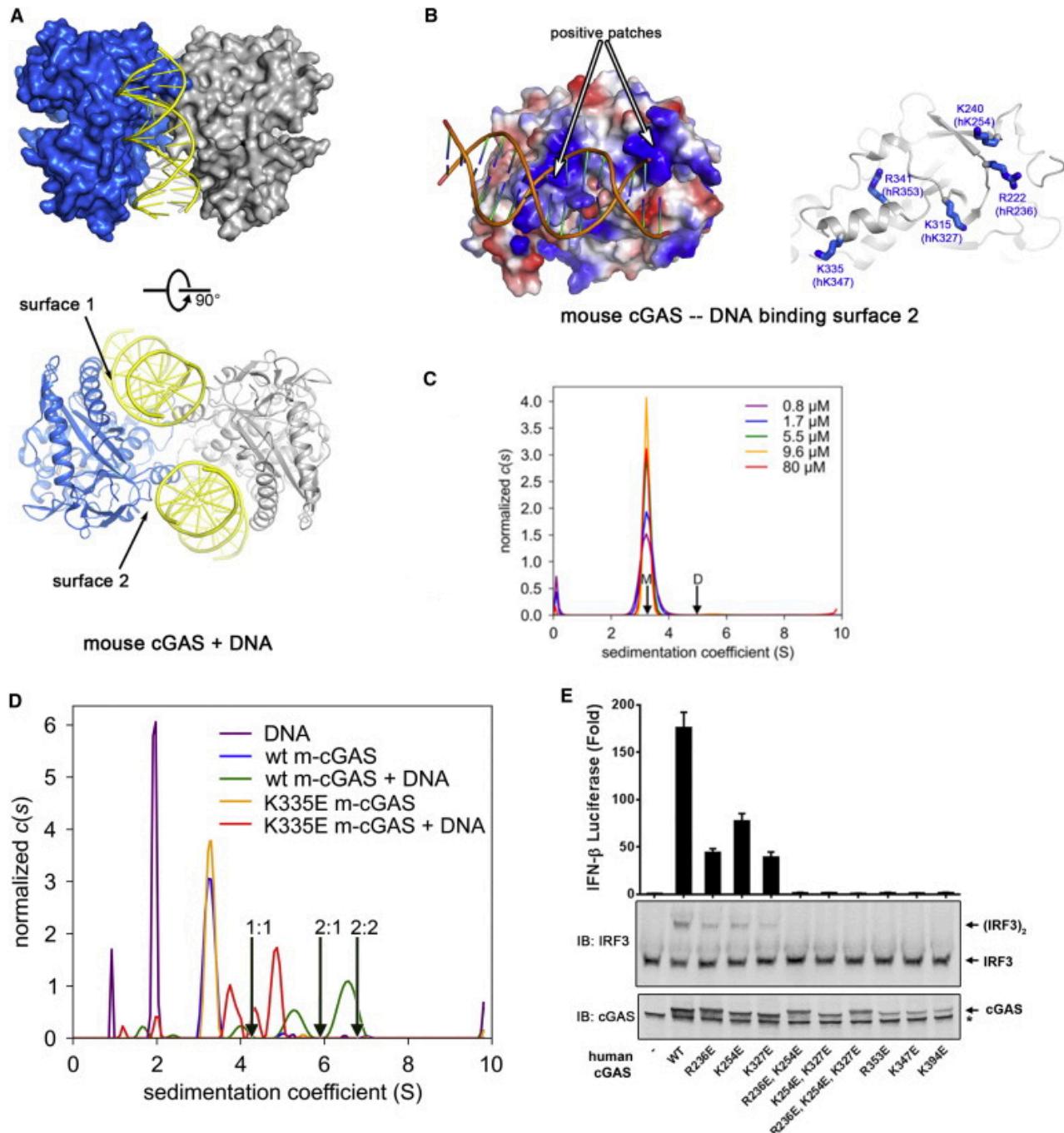


Figure 21. cGAS is activated by DNA-induced dimerization

- (A) Crystal structure of mouse cGAS in complex with a 16 bp dsDNA. Each asymmetric unit contains one 2:2 complex, composed of two protein molecules and two DNA molecules. Two perpendicular views are shown.
- (B) Electrostatic view of DNA binding surface 2 on mouse cGAS (left). Conserved positively charged residues are shown on the right; their corresponding residues in human cGAS are shown in parentheses.
- (C) AUC of m-cGAS alone at indicated concentrations. Arrows indicate the theoretical sedimentation coefficients for m-cGAS monomer ('M') and dimer ('D').
- (D) AUC of m-cGAS (WT or K335E) and DNA in different combinations. Concentrations for cGAS and DNA in each experiment were: DNA alone, 120 μ M; WT m-cGAS, 80 μ M; WT m-cGAS/DNA, 80 and 120 μ M, respectively; K335E m-cGAS, 94 μ M; and K335E m-cGAS/DNA, 94 μ M and 120 μ M, respectively. The arrows indicate the theoretical sedimentation coefficients of 1:1, 2:1, and 2:2 m-cGAS/DNA complexes.
- (E) The second DNA binding surface and the dimeric interface are important for cGAS activity. Expression vectors encoding WT or mutant human cGAS proteins as indicated were transfected into HEK293T-STING-IFN β luciferase reporter cells, followed by measurement of luciferase activities (top). Aliquots of the cell extracts were immunoblotted with an IRF3 antibody following native gel electrophoresis (middle) or with a cGAS antibody following SDS-PAGE (bottom). The error bars represent variation ranges of duplicate experiments.

A Working Model of the Cytosolic DNA Sensing Pathway

Based on the experimental evidence presented above, a model can be proposed for the cytosolic DNA sensing pathway (Figure 22). Briefly, during microbial infection or cellular damage, DNA from pathogens or self is sensed in the cytosol by the cytosolic DNA sensor cGAS. DNA binding induces dimerization and a conformational change of the enzyme cGAS, allowing it to synthesize the second messenger 2'3'-cGAMP from ATP and GTP. cGAMP then binds to and activates the ER-localized adaptor protein STING, which in turn recruits and activates the kinase TBK1 and transcription factor IRF3 to induce type I interferons.

Conclusion of research project

In summary, my thesis study not only elucidated the mechanism of cytosolic DNA

sensing and signaling, but also revealed a new class of second messengers in eukaryotes.

Furthermore, it also establishes the enzyme cGAS as a novel therapeutic target for the

treatment of autoimmune diseases.

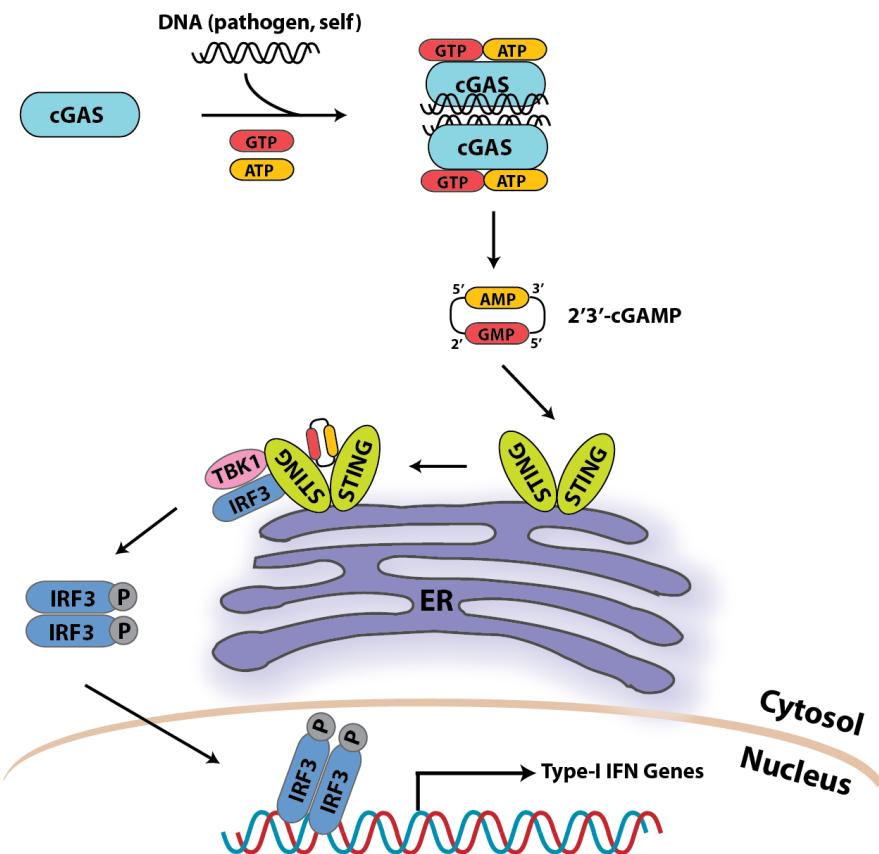


Figure 22. Working model for the cGAS mediated cytosolic DNA sensing pathway

CHAPTER THREE

Discussion

Strategy for the Identification of cGAS

The failure in our initial attempts to purify the cGAMP producing enzyme to homogeneity using traditional protein purification prompted us to develop a new strategy that combined quantitative mass spectrometry with conventional protein purification to identify biologically active proteins partially purified from crude cell extracts. As a proof of principle, we successfully used this strategy to identify the mouse protein E330016A19 as the enzyme that synthesizes cGAMP, which turned out to have very low abundance in L929 cell cytosolic extracts. With the recent advancement in quantitative mass spectrometry techniques, this strategy may be generally applicable to proteins that are difficult to purify to homogeneity using traditional biochemical tools due to very low abundance, labile activity, or scarce starting material.

cGAS is the Long Sought-after DNA Sensor

Using a combination biochemical, genetic, chemical and biophysical techniques, I have provided several key lines of evidence supporting cGAS as the long sought-after cytosolic DNA sensor. First, cGAS is predominantly a cytoplasmic protein, and it colocalizes with fluorophore-labeled DNA in the cytoplasm. Second, purified cGAS protein binds to DNA of varying sequences directly in a biochemical assay. Structural studies suggest that the binding involves two distinct binding surfaces on cGAS and the sugar-phosphate backbone of DNA.

Mutations of the DNA-binding residues abrogate cGAS-induced IFN production. Third, the enzymatic activity of cGAS is strictly dependent on DNA binding. It synthesizes cGAMP only in the presence of DNA. Fourth, overexpression of cGAS activates IRF3 and induces IFN- β in a STING dependent manner. IFN- β induced by cGAS is several orders of magnitude greater than that of other putative DNA sensors. Fifth, various cell types, including macrophages, fibroblasts, and DCs, from cGAS-deficient mice show a complete loss of type I IFN production in response to DNA transfection or DNA virus infection. Furthermore, cGAS knockout mice are more sensitive to lethal infection by DNA viruses than their wild-type littermates (Li et al., 2013b; Schoggins et al., 2014). Thus, cGAS is a bona fide sensor that mediates DNA induced immune responses irrespective of cell type or DNA sequence, and plays essential roles in anti-microbial host defense.

Functional Breadth of cGAS

As a sequence-independent DNA sensor, cGAS is in principle capable of detecting all microorganisms that harbor or generate DNA in their life cycles. However, because many microorganisms likely generate multiple PAMPs (e.g., LPS, CDNs, RNA), it remains to be determined to what extent cGAS contributes to the overall immune responses to each pathogen. Nevertheless, accumulating evidence suggests that cGAS plays important roles in sensing and triggering host defense against a wide class of pathogens. DNA viruses, such as HSV1, VACV and adenovirus trigger immune responses through cGAS (Lam et al., 2014). cGAS also senses the DNA stage of retroviruses, including HIV, SIV and MLV, to trigger the production IFN and other cytokines (Gao et al., 2013a; Lahaye et al., 2013). In addition, cytokine responses to invading bacteria, such as *Mycobacterium tuberculosis*, *Chlamydia*

trachomatis, *Francisella novicida* and *Listeria monocytogenes*, are also mediated by cGAS (Dey et al., 2015; Hansen et al., 2014a; Storek et al., 2015; Zhang et al., 2014b).

In addition to DNA of microbial origin, cGAS also detects self-DNA from different sources under different physiological conditions. Trex-1 deficiency-induced spontaneous cytokine production is mediated by cGAS (Ablasser et al., 2014). During UV or ROS damage, oxidized self-DNA, which is resistant to TREX1-mediated degradation, can potently activate cGAS (Gehrke et al., 2013). Mitochondrial DNA release during mitochondrial stress can also activate cGAS to prime the antiviral state of cells (West et al., 2015). Thus, cGAS monitors both extrinsic and intrinsic dangers and integrates the signals to modulate the antiviral state and homeostasis of a cell.

cGAMP as Novel Second Messenger in Eukaryotes

We identified cGAMP as the first cyclic dinucleotide in eukaryotes and showed that it is a potent immunostimulatory molecule. As a novel second messenger, the function of cGAMP is similar to that of the well-known second messenger cAMP (Beavo and Brunton, 2002). Like cAMP, which is synthesized by adenylate cyclase upon its activation by upstream ligands, cGAMP is synthesized by the cyclase cGAS in response to stimulation by a DNA ligand. cAMP binds to and activates protein kinase A and other effector molecules. Similarly, cGAMP binds to and activates STING to trigger downstream signaling cascades.

The utilization of a second messenger represents a new signaling mechanism in innate immunity. It affords several advantages to the cytosolic DNA sensing pathway. First, the enzymatic synthesis of cGAMP by cGAS provides a mechanism of signal amplification for a

robust and sensitive immune response. Second, cGAMP can rapidly and efficiently transduce signal from sites of DNA detection to ER restricted STING adaptors, due to its small size and faster diffusion rate in the highly-condensed cytosol. Last but not least, cGAMP can also function nonautonomously by spreading from producing cells to neighboring cells through gap junctions (Ablasser et al., 2013). This intercellular communication through a second messenger may provide a rapid response mechanism that protects neighboring cells from virus infection, especially when some viruses produce antagonists that prevent the production of type I IFNs in the virus-infected cells.

Phosphodiester linkages in cGAMP

Endogenous cGAMP contains a unique combination of 2'-5' and 3'-5' phosphodiester bonds. Even though this uniqueness was highlighted by the fact that 2'3'-cGAMP binds to STING with a much higher affinity than cGAMP isomers containing other phosphodiester linkages, all four cGAMP isomers induced IFN- β with comparable EC₅₀ values (Figure 23), which were much lower than that of c-di-GMP. Thus, all cGAMP isoforms are potent inducers of IFN- β , and it is not clear why vertebrates have evolved to produce 2'3'-cGAMP as the endogenous second messenger to trigger innate immune responses. It is also possible that cGAMP containing other phosphodiester linkages might exist in nature, perhaps in some lower organisms. Indeed, the bacteria *Vibrio cholera* contains a cyclase that can produce 3'3'-cGAMP to regulate its chemotaxis and colonization (Davies et al., 2012; Kranzusch et al., 2014). In addition, the mixed phosphodiester linkage may offer a tighter control of cGAMP activity. On one hand, it makes cGAMP resistant to bacterial 3'-5' specific CDN

phosphodiesterases, allowing more robust and sustained host defense responses. On the other hand, the host may encode a 2'-5' specific phosphodiesterase that is capable of degrading 2'3'-cGAMP to avoid autoimmunity. Indeed, ENPP1 was recently identified as an enzyme that contains specific hydrolyzing activity towards 2'3'-cGAMP, but not 3'3'-cGAMP (Li et al., 2014). It will be interesting to see whether mice lacking this enzyme develop autoimmune related phenotypes.

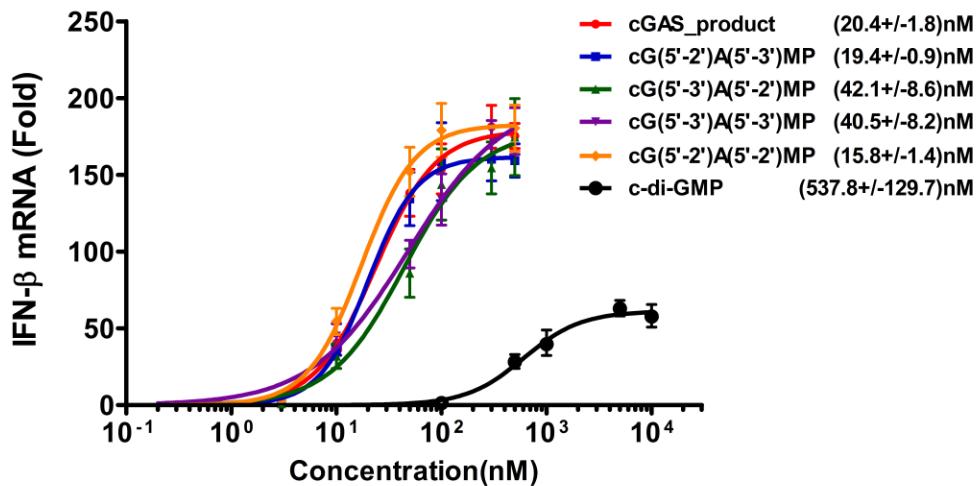


Figure 23. cGAMP isomers are potent inducers of type I Interferon

Different concentrations of the cGAS product, synthetic cGAMPs, and c-di-GMP were delivered into digitonin-permeabilized L929 cells. Four hours later, IFN- β RNA was measured by qRT-PCR. Dose-response curves and the half-maximal effective concentration (EC_{50}) for each compound were generated using GraphPad Prism 5.0 software. The error bars indicate standard deviations of triplicate experiments.

Evolution of cGAS and STING

Detection of foreign DNA by cGAS and the resulting production of endogenous cGAMP vastly expand the repertoire of pathogens detectable by the innate immune system, from CDN-producing bacteria to potentially any microorganism that carries or generates DNA in

its life cycle. It remains to be determined whether STING evolved first to detect bacterial CDNs or endogenous 2'3'-cGAMP produced as a result of cGAS activation by DNA from different microbes. The fact that endogenous 2'3'-cGAMP binds to STING with a much higher affinity and serves as a more potent activator of IFN seems to suggest that the cGAS-cGAMP-STING pathway has a selective advantage through evolution.

To gain more insights into the evolution of cGAS and STING, I search individual genomes of major animal phyla for homologs of human cGAS and STING proteins using BlastP. Whereas in vertebrates cGAS homologs are confidently identified and all contain the conserved residues involved in catalysis and DNA binding, the invertebrate cGASs are more orthologous to human protein MAB21L2 (Figure 24A). Interestingly, cGAS from the cephalochordate *Branchiostoma floridae* also contains key catalytic residues and is in close proximity to the highly conserved vertebrate cGAS cluster in the phylogenetic tree (Figure 24A). This suggests that cGAS originated during the transition from invertebrates to vertebrates. In contrast, STING proteins are found in almost all phyla, indicating a disparate evolution pattern to cGAS. However, further phylogenetic analysis of STING proteins from representative species reveals that cephalochordate and vertebrate STINGs form a group that is distant from those in invertebrates (Figure 24B). Thus, it remains to be determined if these invertebrate STINGs are functional sensors for CDNs. It is still possible that both full-fledged cGAS and STING emerge from the vertebrate lineage, functioning cooperatively to detect pathogen infections.

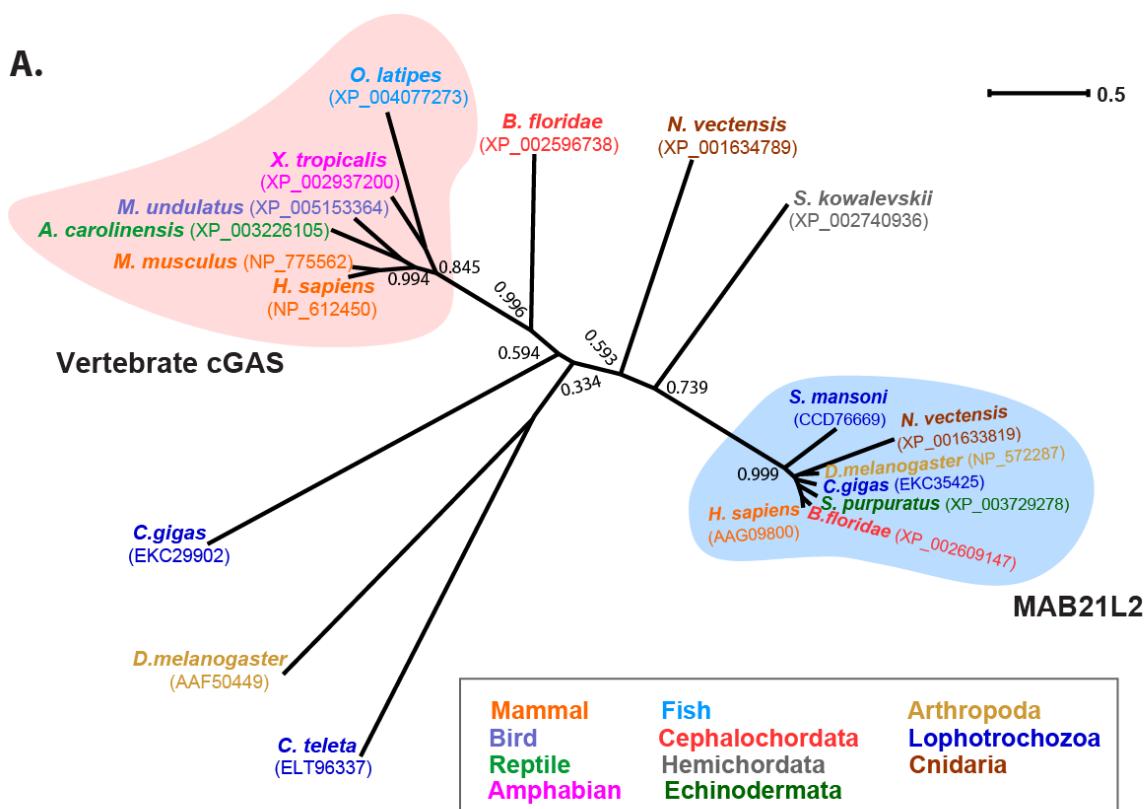
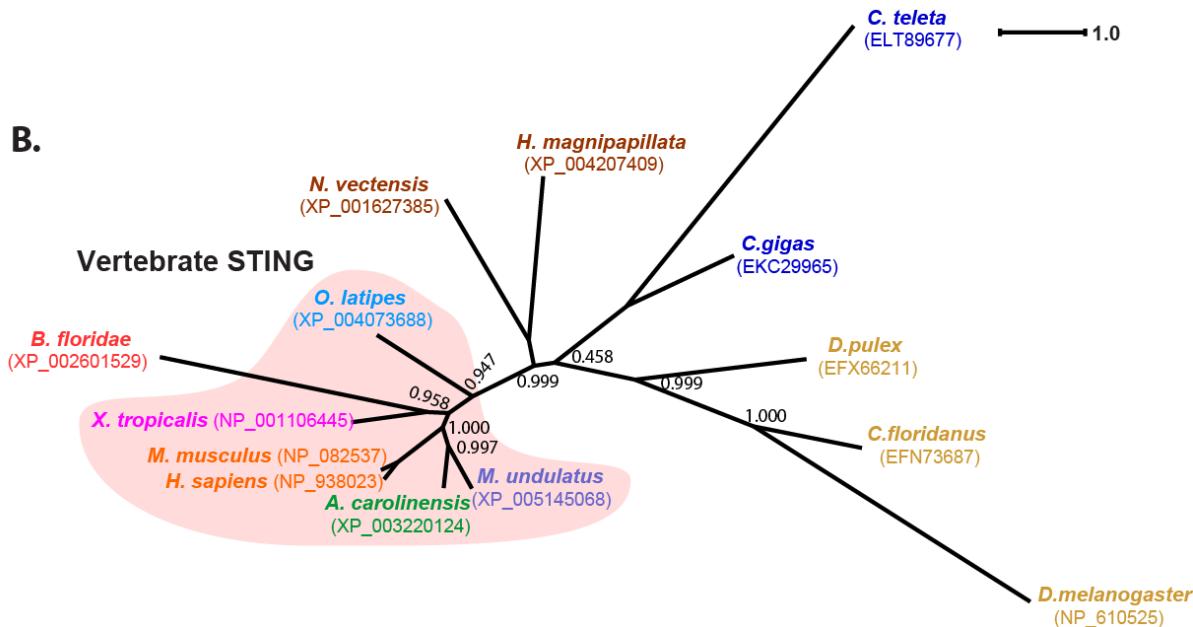
A.**B.**

Figure 24. Phylogenetic analyses of cGAS and STING

cGAS and STING proteins from representative species were obtained by NCBI BlastP search. STING homologs were found throughout metazoa. Although cGAS homologs were identified with certainty in vertebrates (*red*), the invertebrate cGASs are likely to be orthologs of human protein MAB21L2 (*blue*). Multiple sequence alignments of cGAS and STING from different species were generated using ClustalW. Alignments were adjusted manually before loading to the PhyML program for phylogenetic analysis. The evolutionary tree was built with LG as the substitution model, four discrete rate categories for the rate heterogeneity among sites, Nearest Neighbor Interchange for the tree improvement, and SH-like approximate likelihood-ratio test for estimating the branch support. Bootstrap support values are shown for major branches. Trees were drawn using Dendroscope software.

CHAPTER FOUR

Methodology

Cells, Antibodies, Nucleotides and General Methods

cDNA clones for m-cGAS (E330019A16; accession #BC145653) and h-cGAS (C6orf150; accession #BC108714) were purchased from Open Biosystems. Cell lines including L929, HEK293T, and Raw264.7, were cultured in DMEM (Sigma) supplemented with 10% calf serum and antibiotics. THP-1 cells were cultured in RPMI containing 10% fetal bovine serum, antibiotics, and 50 μ M of β -mercaptoethanol. CpG ODN was purchased from InvivoGen. Genomic DNA from *E. coli* and *V. cholerae* was prepared by phenol/chloroform extraction followed by RNase A digestion.

The antibody against human cGAS (C6orf150) was purchased from Sigma. The rabbit polyclonal antibodies against STING were generated and affinity purified as described previously (Tanaka and Chen, 2012). Antibodies against human and murine IRF3 were obtained from Santa Cruz Biotechnology and Invitrogen, respectively. Poly (I:C), herring testis (HT) DNA, Anti-Flag(M2)-agarose and all chemicals were from Sigma unless otherwise indicated. ISD was prepared from equimolar amounts of the sense and antisense DNA oligonucleotide (sense strand sequence: 5'-TACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACA-3'). The oligonucleotides were heated at 95°C for 5 min and cooled to room temperature. All chromatography columns were from GE Healthcare. [35 S]-methionine was from Perkin Elmer. siRNA against indicated genes were from Sigma and transfected into L929 cells using Lipofectamine 2000 according to manufacturer's instruction (Invitrogen). The

lentiviral vector for shRNA was described previously (Tanaka and Chen, 2012); the

targeting sequence for different genes were (only sense strand sequence is shown):

STING: 5'- CCTCATCAGTGGAAATGGAA; m-cGAS-a: 5'-

GGATTGAGCTACAAGAATA, and m-cGAS-b: 5'-GCTGTAACACTTCTTATCA-3';

h-cGAS: 5'-GGAAGGAAATGGTTCCAA.

The procedures for native gel electrophoresis to detect IRF3 dimerization, SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and Western blotting have been described previously (Seth et al., 2005). Digitonin permeabilization was used to deliver cGAMP, c-di-GMP, and c-di-AMP into cultured cells as previously described (Woodward et al., 2010). Reverse transcription and real-time PCR reactions were carried out as described before (Chiu et al., 2009). Briefly, iScript cDNA synthesis kit and iQ SYBR Green Supermix (Bio-Rad) were used according to manufacturer's instructions. Quantitative real time PCR was performed on an Applied Biosystem ViiA7 with the following primers: Mouse IFN- β , sense: TCCGAGCAGAGATCT TCAGGAA, antisense: TGCAACCACCACTCATTCTGAG; Mouse HPRT, sense: CAGTCCCAGCGTCGTGATTAG, anti-sense: AACACACTTTCCAATCCTCGG; Human IFN- β , sense: AGGACAGGATGAACTTGAC, anti-sense: TGATAGACATTAGCCAGGAG; Human TNF α , sense: TGCTTGTCCTCAGCCTCTT, anti-sense: GGTTGCTACAACATGGGCT; Human GAPDH, sense: GAGTCAACGGATT GGCGT, anti-sense: TTGATTGGAGGGATCTCG; Mouse cGAS, sense: ACCGGACAAGCTAAAGAAGGTGCT, anti-sense:

GCAGCAGGCGTTCCACAACTTAT; Human cGAS, sense:

GGGAGCCCTGCTGTAACACTTCTTAT, anti-sense:

CCTTGATGCTGGGTACAAGGT.

Viruses

Sendai virus (Cantell strain, Charles River Laboratories) was used at a final concentration of 50 hemagglutinating units/ml. VSV(ΔM51)-GFP were used at a multiplicity of infection (MOI) of 10. HSV1 WT and ΔICP34.5 strains were propagated and titered by plaque assays on Vero Cells and used at 20 MOI. The d109 and 7134 strains were gifts from Dr. Neal DeLuca (University of Pittsburgh) and Dr. David Knipe (Harvard University) and used at an MOI of 0.5 and 10, respectively (Melroe et al., 2004; Samaniego et al., 1998). Vaccinia virus (VACV, WR strain) was propagated in HeLa cells and used at an MOI of 5 after heat inactivation.

In vitro Assay for STING Activator

To isolate the cellular factor that could activate the STING pathway, L929 cells were transfected with DNA for 4-6 hours and then homogenized by douncing in a hypotonic buffer containing 10mM Tris-HCl, pH7.4, 10mM KCl, 1.5mM MgCl₂. The homogenate was centrifuged at 100,000 rpm for 20 min, and the supernatant (S100) was heated at 95°C for 5 min and centrifuged again at 12,000 rpm for 5 min to remove denatured proteins. The heat-resistant supernatant was mixed with 10⁶ THP-1 or Raw264.7 cells in an 8μl reaction containing 2 mM ATP, 1 U/μl Benzonase and 1.5 ng/μl PFO. The mixture was incubated at 30°C for 1.5 hours. Cells in the reaction were lysed

by adding 0.2% NP40 and subjected to native gel electrophoresis. IRF3 dimerization was detected by immunoblotting with an IRF3 antibody. In some experiments [³⁵S]-IRF3 was added to the mixture for detection of IRF3 dimerization by autoradiography.

To synthesize the STING activator in a cell-free system, S100 from non-stimulated L929 or other cell lines as indicated was mixed with 0.1 mg/ml HT-DNA or ISD, 20 mM Hepes-OH, pH7.2, 5 mM MgCl₂ and 2 mM ATP. After incubation at 37°C for 45 min, the mixture was heated at 95°C for 5 min, and the heat-resistant supernatant was mixed with PFO permeabilized THP1 or Raw264.7 cells to measure IRF3 activation as described above.

Purification of the Small Molecule STING Activator

150 dishes (150mm) of L929 stably expressing STING-Flag were lysed in lysis buffer I [50mM Tris-Cl, pH7.4, 100mM NaCl, 10% Glycerol, 0.5% NP40, 1mM EDTA, 1mM EGTA, 2% protease inhibitor cocktail (Roche), 0.5mM DTT] and centrifuged at 1,000 x g for 5 mins. The supernatant was mixed with 1 ml of Flag (M2)-agarose and rotated at 4°C for 2 hrs. The beads were washed in lysis buffer I and used as immobilized STING affinity resin. To generate the small molecule STING activator, 50 ml reaction containing ~250 mg L929 S100 and 5 mg DNA was carried out at 37°C for 45 min. The heat-resistant supernatant from the reaction was mixed with the STING affinity resin and incubated at 4°C for 1 hour by constant rotation. The beads were washed in a buffer containing 20mM Hepes, pH7.2, 50 mM NaCl, 0.1mM EGTA, and the STING activator was eluted by boiling in water for 5 min. After a brief centrifugation, the supernatant was loaded on a mono Q column (0.1ml bed volume, GE Healthcare) in 20 mM Tris-HCl, pH

7.4, and eluted with a 0-0.5M NaCl linear gradient. Active fractions (0.12-0.15M NaCl) were pooled and loaded onto a Superdex-peptide column (2.4 ml bed volume, GE Healthcare) and eluted in 20 mM Tris-HCl, pH 7.4. Active fractions were applied to a C-18 column (Eclipse Plus 4.6x30 mm, 3.5 μ m, Agilent Technologies) equilibrated with 0.1% formic acid and eluted with a linear gradient of 0-100% methanol. Fractions were dried at 60°C under vacuum and dissolved in water for activity assay and mass spectrometry.

LC-MS/MS of STING Activator

LC-MS/MS analysis of the purified STING activator was performed on a LTQ-XL mass spectrometer (ThermoFisher Scientific) equipped with a HPLC system (Eksigent NanoLC 2D). Analytes were eluted using a segmented gradient of 2–70% of eluent B [0.1% Formic Acid (FA) / 100% Acetonitrile (ACN)] with a constant flow of 200 nl/min in 56 min. The mass spectrometer was operated in data-dependent mode with dynamic exclusion of 30s. Full scan MS spectra were acquired in a mass range from m/z 100 to 1500. The 3 most intense ions were sequentially isolated for CID fragmentation in the linear ion trap to generate MS/MS spectra.

For targeted quantification of cGAMP, a selective reaction monitoring (SRM) assay was developed on the LTQ XL mass spectrometer. In the assay, three transitions of each target ion were monitored (see Table 2). The parent ion was isolated with a mass window of 1.0 m/z units and fragmented (collision energy=35, activation time=30 ms at Q=0.25), while the resulting daughter ion was scanned with a mass window of 1.6 m/z unit. The maximum ion accumulation time was 100ms, and the number of microscans was set to 1.

Precursor	Charge	Product_1	Product_2	Product_3	CE
675.11	1+	312.05	524.06	540.05	35
338.06	2+	136.06	152.06	524.06	35

Table 2. Transition list and collision energies used to monitor cGAMP by SRM. CE: collision energy

Biochemical Fractionation of cGAS Activity From Cell Extracts

Cytosolic extracts (S100) of L929-shSTING cells were used to purify endogenous cGAS activity through three different routes (see figure 9A), each route using approximately 250 mg of starting material [from ~200 plates (15 cm ID) of cultured cells]. In route I, S100 was loaded onto a Heparin column equilibrated with buffer A [20 mM Tris-HCl, pH7.4, 0.02% CHAPS, 0.5 mM DTT, and 0.1 mM PMSF], and eluted sequentially with 0.5 M and 1.0 M NaCl in buffer B. The 1.0 M eluate, which contained the cGAS activity, was further fractionated with a 0.5 M-1.0 M gradient on a heparin column. Active fractions were pooled, concentrated and subjected to gel filtration on a Superose 6 PC 3.2/30 column in buffer B [20mM Tris-Cl, pH8.0, 150mM NaCl, 0.02% CHAPS, 0.5mM DTT, 0.1mM PMSF] using the ETTAN LC system (GE Healthcare). Active fractions were loaded onto a Mono Q PC 1.6/5 column (GE Healthcare) equilibrated with buffer A and eluted with a 0-0.5 M NaCl gradient. In route II, the 1.0 M eluate from the Heparin column was loaded onto a hydroxyapatite column (Bio-Rad)

equilibrated with a buffer containing 10mM KPO₄, pH7.0 (a mixture of K₂HPO₄ and KH₂PO₄), and eluted with 0.125 M KPO₄, pH7.0. The eluate was further fractionated with the Mono Q PC column as described in route I. Active fractions were pooled and loaded onto a Mono S PC 1.6/5 column in buffer C [20 mM Tris-Cl, pH 8.3, 0.02% CHAPS, 0.5 mM DTT, 0.1mM PMSF], and eluted with a 0-1.0 M NaCl gradient in buffer C. In route III, active fractions from mono Q PC of route II were applied to biotin-ISD (Sigma) immobilized on monomeric Avidin UltraLink beads (Pierce), which was equilibrated with buffer D [20mM Tris-Cl, pH7.4, 100 mM NaCl, 0.02% CHAPS]. After extensive washing with the same buffer, bound proteins were eluted with 2 mM biotin in buffer D.

In vitro Assay for cGAS Activity

To measure cGAS activity, cell extracts or fractions from chromatography were mixed with buffer E (20 mM Hepes, pH7.2, 5 mM MgCl₂, 2 mM ATP, 2 mM GTP, 0.1 mM EGTA), in the presence or absence of 0.1 mg/ml herring testis DNA (HT-DNA). After incubation at 37°C for 45 min, the mixture was heated at 95°C for 5 min, centrifuged at 10,000 x g for 3 mins, and then the heat-resistant supernatant was mixed with PFO permeabilized THP1 or Raw264.7 cells to measure IRF3 activation.

Protein Identification by Mass Spectrometry

Fractions from the last step of each purification route were resolved by SDS-PAGE and then visualized by silver staining. For each purification route, one lane from the fraction containing the highest cGAS activity and two adjacent lanes from fractions

containing low cGAS activity were chosen for mass spectrometry analyses. Each lane was cut into 10 gel slices of approximately equal sizes, which were destained and then reduced in 20 mM DTT at 56°C for 30 min followed by alkylation in 55 mM iodoacetamide in the dark for 1 hr. Proteins in the gels were digested *in situ* with sequence grade trypsin (Promega) in 50 mM ammonium bicarbonate at 37°C overnight. Peptides were extracted sequentially with 5% formic acid (FA) /50% acetonitrile (ACN) and 0.1% FA /75% ACN, vacuum dried and then resuspended in 10 µl of 0.1% FA. 5 µl of the peptide sample was loaded via an autosampler (Thermofisher Scientific) onto a homemade analytical column (75 µm ID, 150 mm length) packed with C-18 resin (100 Å, 3 µm, MICHROM Bioresources). The peptides were then eluted with a 78 min gradient generated by a Dionex Ultimate 3000 nanoLC system (Thermofisher Scientific) as follows: 2–30% B in 68 min, 30–35% B in 4 min, 35–40% B in 2 min, 40–60% B in 3 min, and 60–80% B in 1 min (A = 0.1% FA; B = 0.1% FA/ 100% ACN). The eluted peptides were sprayed into to a Quadrupole-Orbitrap Hybrid mass spectrometer (Q Exactive, ThermoFisher Scientific) equipped with a nano-electrospray ion source. MS/MS spectra were acquired in a data-dependent mode whereby the top 10 most abundant parent ions were subjected to further fragmentation by higher energy collision dissociation (HCD). For protein identification, database searches were performed on an in-house Mascot server (Matrix Science) against the IPI mouse database (v3.87). Carbamidomethylcysteine was set as a fixed modification and N-acetylation and methionine oxidation were set as variable modifications.

Quantitative Analysis of Mass Spectrometry Data

For label-free quantification of proteins identified in the fractions from the last step of each purification route, raw data were searched again using the Andromeda search engine and the MaxQuant software package (1.3.0.5) against the IPI mouse database (v3.87) (Cox and Mann, 2008; Cox et al., 2011). The first search tolerance was set at 20 ppm and main search deviation at 6 ppm. The required minimum peptide length was six amino acids. The false discovery rate (FDR) at both peptide and protein levels was set to 0.01. The methods for comparative analysis of the amount of a protein in different fractions were defined manually and coupled to MaxQuant through the experimental design file. To obtain protein candidates that correlated with the cGAS activity in different purification routes, MaxQuant output results were verified manually with a comprehensive consideration of peptide counts, sequence coverage and protein intensities. The Venn diagram showing proteins that overlapped among different purification routes was generated by the “VENNY” online server.

Expression and Purification of Recombinant cGAS Proteins

cDNA encoding full-length human or mouse cGAS was inserted into a modified pET28a vector containing an in-frame His6-SUMO tag. The *E. coli* strain BL21/pLys harboring the plasmid was induced with 0.5 mM IPTG at 18°C overnight. *E. coli* was resuspended and sonicated in lysis buffer II [50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol, 0.2 mM PMSF] and centrifuged at 100,000 x g for 20 min. The supernatant was incubated with Ni-NTA beads (Qiagen) in lysis buffer II. After washing in the same buffer, bound proteins were eluted with buffer F (20 mM TrisCl, pH7.4, 150 mM NaCl, 300 mM Imidazole). The His6-SUMO tag was cleaved by

SUMO protease (Ulp1) and the mixture was loaded onto a 1-ml HiTrap Heparin column (GE Healthcare) in buffer A. After washing with buffer A containing 0.5 M NaCl, the cGAS protein was eluted with 1.0 M NaCl in the same buffer.

GST-tagged cGAS was generated in pLys/BL21 using Gateway cloning according to manufacturer's instruction. After induction with 0.8 mM IPTG at 18°C overnight, the bacteria were resuspended and sonicated in lysis buffer I [50 mM Tris-HCl, pH7.4, 100mM NaCl, 10% glycerol, 0.5% NP40, 0.5 mM EDTA, 0.5 mM EGTA], and centrifuged at 100,000 x g for 20 min. The supernatant was mixed with glutathione Sepharose (GE Healthcare) and washed extensively with lysis buffer I. Bound proteins were eluted with 10 mM reduced glutathione in 50 mM Tris-Cl, pH8.0, and dialysed against buffer G [20mM Tris-Cl, pH. 7.4, 50mM NaCl, 10% Glycerol].

To express recombinant cGAS protein in mammalian cells, HEK293T was transfected with pcDNA3-Flag-h-cGAS for 36 hrs. Cytosolic extracts were prepared from 20 plates (15-cm ID) of cultured cells by douncing in a hypotonic buffer [10mM Tris-Cl, pH7.4, 10mM KCl, 1.5mM MgCl₂, 0.5mM DTT and protease inhibitor cocktail (Roche)], and loaded onto a 5-ml HiTrap Heparin column in buffer A. After washing with 0.5M NaCl in buffer A, Flag-h-cGAS was eluted with 1.0M NaCl and then mixed with agarose beads coupled with anti-Flag (M2) antibody (Sigma). After washing in lysis buffer I, bound proteins were eluted with 0.2 mg/ml FLAG peptide in a buffer containing 50 mM Tris-Cl, pH8.0, and 0.02% CHAPS.

DNA Binding Assay

Recombinant GST-tagged cGAS was incubated with streptavidin UltraLink beads in the presence of ISD or biotin-ISD in lysis buffer I. After three washes in the same buffer, bound proteins were eluted by boiling in SDS sample buffer and detected by immunoblotting using a GST antibody (Covance). A biotin-RNA sequence (ACGGAAAGACCCGU) from 23S rRNA of DH5 α was used as a negative control.

Immunostaining and Confocal Microscopy

Immunostaining with L929 cells stably expressing Flag-tagged m-cGAS was performed as described previously (Tanaka and Chen, 2012). Cells were transfected with Cy3-ISD for the indicated times and fixed in 4% formaldehyde in PBS for 15 min. Anti-Flag M2 (1:500, Sigma) and Alexa 488 conjugated goat anti-mouse IgG (1:500, Invitrogen) were used as the primary and secondary antibodies to detect Flag-cGAS, respectively. Nuclei were labeled by staining with DAPI in the mounting medium (Vectashield). Images of the cells were collected with a Zeiss LSM510 META laser scanning confocal microscope (Carl Zeiss MicroImaging Inc).

Chemical Synthesis of cGAMP Isomers

Chemical syntheses of cGAMP isomers with different phosphodiester-linkage combinations were conducted by Dr. Heping Shi, under the supervision of Dr. Chuo Chen in the Department of Biochemistry. The detailed experimental procedures can be found in our published paper (Zhang et al., 2013).

Enzymatic Synthesis and Purification of cGAMP

To generate natural cGAMP using the enzyme cGAS, a reaction containing 20mM Tris-Cl, pH7.5, 5mM MgCl₂, 10mM CoCl₂, 0.01mg/ml herring testis DNA, 1mM ATP, 1mM GTP, and 0.1μM recombinant SUMO-tagged human cGAS (aa147–522) was incubated at 37°C for 2hrs. The mixture was fractionated on a Hitrap Q column using a linear 0–0.5M NaCl gradient; a UV peak corresponding to cGAMP was collected and loaded onto a C18 column (201TP510, 1cmX25cm, Phenomenex, Hesperia, CA), and eluted with a linear 0–100% methanol gradient.

Preparation of Endogenous cGAMP from Mouse and Human Cells

Endogenous cGAMP was prepared from DNA transfected L929 and THP-1 cells, respectively. After HT-DNA transfection for 4 hours, about 3×10^7 cells were lysed in hypotonic buffer. The lysates were heated at 95°C for 5 min and centrifuged again at 17,000g for 10 min to remove denatured proteins. The heat-resistant supernatant was fractionated on a C-18 column (Eclipse Plus 4.6×30 mm, 3.5μm, Agilent Technologies) equilibrated with 0.1% formic acid and eluted with a linear gradient of 0–100% methanol. The presence of cGAMP in each fraction was monitored by activity assay as described above, and the fraction with peak activity was used for further MS and MS/MS analysis.

MS and MS/MS Analyses of cGAMP Isomers

High resolution LC-MS analysis was performed using a Dionex Ultimate 3000 nanoLC system (Thermofisher Scientific) coupled to a Quadrupole-Orbitrap Hybrid mass spectrometer (Q-Exactive, ThermoFisher Scientific) equipped with a nano-electrospray ion source. Ionization source parameters were set to: positive mode; capillary temperature,

250°C; spray voltage, 2.3 kV. Analytes were separated on a homemade analytical column (75 µm ID, 100 mm length) packed with C-18 resin (100 Å, 5 µm, MICHROM Bioresources). HPLC eluent A: 0.1% FA; eluent B: 0.1% FA /100% ACN; HPLC gradient: 2–30% B in 12 min, 30–70% B in 2 min; 100 nl/min. Full scan MS spectra were acquired in a mass range from m/z 300 to 800 with a resolution of 140,000 in the Orbitrap mass analyzer.

Tandem MS/MS analysis was conducted using the same LC-MS system. Full scan mass spectra were acquired from m/z 300–700 with a resolution of 70,000 at m/z = 200 in the Orbitrap. MS/MS spectra (resolution: 35,000 at m/z = 200) were acquired in a data-dependent mode whereby the top 5 most abundant parent ions were subjected to further fragmentation by higher energy collision dissociation (HCD). The normalized collision energy was set at 30.

Expression and Purification Human STING CTD

The DNA of wild type human STING C-terminal domain (STING CTD, 139–379) was cloned into a pET-SUMO vector (Invitrogen). Overexpression of STING CTD was induced in *E. coli* BL21/pLysS with 0.8 mM IPTG when the cell density reached an OD₆₀₀ of 1.2. After growth at 20°C for 12 hrs, the cells were collected, resuspended in a buffer containing 25mM Hepes, pH 7.8, and 150mM NaCl, 0.2mM PMSF, and disrupted by French Press with 2 passes at 15,000 p.s.i. Cell debris were removed by centrifugation at 27,000g for 50 min. The supernatant was loaded onto a Ni-NTA agarose resin (Qiagen). Subsequently, the resin was rinsed three times with 20 ml buffer containing 25 mM Hepes, pH 7.8, 500 mM NaCl, 20 mM imidazole-HCl, pH 8.0. The SUMO tag was then

removed by digesting the proteins using the SUMO protease at 4°C overnight. The eluate was concentrated to about 15 mg/ml before applying to gel-filtration chromatography (Superdex-200 10/30, GE Healthcare), which was equilibrated in the buffer containing 25 mM Hepes, pH 7.8, 150 mM NaCl. The peak fractions of the protein were collected and concentrated to 0.42–1.4 mg/ml (15–50 µM) for ITC experiments.

Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC) was conducted to measure the binding affinities between STING and cGAMP isomers or c-di-GMP using a VP-ITC microcalorimeter (GE Healthcare). The protein and the ligand concentrations are shown in Table 1. The titrations were performed at 20°C in the buffer containing 25 mM Hepes, pH 7.8, 150 mM NaCl. 32 injections were performed with 4 minutes spacing time. The titration traces were integrated by NITPIC (Keller et al., 2012) and then the curves were fitted by SEDFIT (Houtman et al., 2007). The figures were prepared using GUSSI (<http://biophysics.swmed.edu/MBR/software.html>).

Isolation of Lung Fibroblasts, BMDMs and Dendritic Cells

To isolate primary lung fibroblasts, lung tissues from each mouse were minced and digested twice in 5 ml of DMEM with 0.1% collagenase D (Roche) and 0.2% trypsin (Sigma) at 37°C for 45-60 min. After washing with DMEM containing 10% calf serum, cells were seeded in a 10-cm dish in 12 ml DMEM containing 10% fetal bovine serum (FBS) in addition to 10% calf serum and antibiotics. Bone marrow cells were collected from femurs and tibiae of mice. To obtain BMDM, about 1×10^7 bone marrow cells were

cultured in DMEM containing 10% FBS, antibiotics and conditioned media from L929 cell culture. After 6 to 7 days, mature macrophages were harvested and cultured on 96-well or 12-well plates for experiments.

To obtain dendritic cells, bone marrow cells were cultured in RPMI 1640 containing 10% FBS, 10 mM HEPES (pH 7.4), 50 µM β-mercaptoethanol, 10 ng/ml murine GM-CSF (peproTech) or 100 ng/ml murine Flt3 ligand (peproTech). After 6–8 days, cells were collected and used as GM-CSF-induced BMDCs (cDC) and Flt3L-induced BMDCs (pDC), respectively.

ELISA

The culture media of mouse lung fibroblasts (3×10^5 cells in 1ml), BMDMs (5×10^5 cells in 0.5ml), GM-CSF DCs (5×10^5 cells in 0.5ml) and Flt3L-DCs (1×10^5 cells in 0.5ml) were collected 20 hours after stimulation. ELISA for mouse IFN-α and IFN-β was performed using the VeriKine Kit (PBL) according to the manufacturer's instructions.

Crystallization of Mouse cGAS-DNA Complex

Recombinant mouse cGAS (147-507) was expressed and purified in a similar way as described above. The protein was concentrated to 25mg/ml in buffer containing 25mM Tris-HCl (pH8.0), 300mM NaCl, 2mM DTT. For crystallization trials, a 16bp dsDNA (Gao et al., 2013b) was mixed with mcGAS (147-507) in a 1.2:1 molar ratio followed by incubation at room temperature for 30 min. Crystals were grown at 20 °C by hanging-drop vapor diffusion method. It yielded crystals in a buffer containing 25% Methanol, 0.1M Tris-HCl, pH 8.0, 0.01M MgCl₂. The crystals were cryoprotected with

crystallization buffer plus 30% ethylene glycol before being flash frozen in liquid nitrogen.

X-ray diffraction data were collected at the Structural Biology Center (Beamline 19ID) at Argonne National Laboratory. Diffraction data were processed with the HKL 3000 package (Minor et al., 2006). The structure was solved by molecular replacement using 4K98 as the search model, and then optimized by manual building with Coot (Emsley et al., 2010). Refinement was performed after excluding 10% of the data for the R-free calculation.

Analytical Ultracentrifugation

AUC experiments were carried out with the help of Dr. Chad Brautigam and Dr. Thomas Scheuermann in the UT Southwestern Macromolecular Biophysics Resource core, using a Beckman Coulter Optima XL-I ultracentrifuge. All samples were prepared in a buffer containing 25 mM Tris-HCl pH 8.0 and 25 mM NaCl, and incubated overnight at 4°C before the centrifugation. The concentrations for m-cGAS and DNA used in each sample are indicated in Figure 21C and 21D. Data were acquired using both interference optics and absorbance optics. Detailed experimental and data processing procedures can be found in a paper we published (Zhang et al., 2014a).

Luciferase Reporter Assay for cGAS Mutants

HEK293T-STING-IFN β luciferase reporter cell line was constructed by lentiviral-mediated stable expression of human STING into previously described HEK293-IFN β -luciferase cells (Chiu et al., 2009). To test the IFN β inducing ability of different cGAS

mutants, expression plasmids were transfected into HEK293T-STING-IFN β luciferase reporter cells. 24 hours after transfection, cells were lysed in the passive lysis buffer (Promega). An aliquot of the lysate was used to measure firefly luciferase activity, whereas another aliquot for measuring renilla luciferase activity (Promega). IFN β induction was calculated by normalizing firefly luciferase activity to renilla luciferase activity. All measurements were performed on a FLUOstar Optima Fluorometer (BMG Labtech).

BIBLIOGRAPHY

- Ablasser, A., Hemmerling, I., Schmid-Burgk, J.L., Behrendt, R., Roers, A., and Hornung, V. (2014). TREX1 deficiency triggers cell-autonomous immunity in a cGAS-dependent manner. *Journal of immunology (Baltimore, Md : 1950)* *192*, 5993-5997.
- Ablasser, A., Schmid-Burgk, J.L., Hemmerling, I., Horvath, G.L., Schmidt, T., Latz, E., and Hornung, V. (2013). Cell intrinsic immunity spreads to bystander cells via the intercellular transfer of cGAMP. *Nature* *503*, 530-534.
- Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen Recognition and Innate Immunity. *Cell* *124*, 783-801.
- Alexopoulou, L., Holt, A.C., Medzhitov, R., and Flavell, R.A. (2001). Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* *413*, 732-738.
- Balbo, A., Minor, K.H., Velikovsky, C.A., Mariuzza, R.A., Peterson, C.B., and Schuck, P. (2005). Studying multiprotein complexes by multisignal sedimentation velocity analytical ultracentrifugation. *Proceedings of the National Academy of Sciences of the United States of America* *102*, 81-86.
- Beavo, J.A., and Brunton, L.L. (2002). Cyclic nucleotide research [mdash] still expanding after half a century. *Nat Rev Mol Cell Biol* *3*, 710-718.
- Burckstummer, T., Baumann, C., Bluml, S., Dixit, E., Durnberger, G., Jahn, H., Planyavsky, M., Bilban, M., Colinge, J., Bennett, K.L., et al. (2009). An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. *Nat Immunol* *10*, 266-272.
- Burdette, D.L., Monroe, K.M., Sotelo-Troha, K., Iwig, J.S., Eckert, B., Hyodo, M., Hayakawa, Y., and Vance, R.E. (2011). STING is a direct innate immune sensor of cyclic di-GMP. *Nature* *478*, 515-518.
- Chiu, Y.-H., MacMillan, J.B., and Chen, Z.J. (2009). RNA Polymerase III Detects Cytosolic DNA and Induces Type-I Interferons Through the RIG-I Pathway. *Cell* *138*, 576-591.
- Civril, F., Deimling, T., de Oliveira Mann, C.C., Ablasser, A., Moldt, M., Witte, G., Hornung, V., and Hopfner, K.-P. (2013). Structural mechanism of cytosolic DNA sensing by cGAS. *Nature* *498*, 332-337.

- Coban, C., Tozuka, M., Jounai, N., Kobiyama, K., Takeshita, F., Tang, C.K., and Ishii, K.J. (2014). Chapter 11 - DNA Vaccine: Does it Target the Double Stranded-DNA Sensing Pathway?*. In *Biological DNA Sensor*, K.J.I.K. Tang, ed. (Amsterdam: Academic Press), pp. 257-270.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotech* 26, 1367-1372.
- Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research* 10, 1794-1805.
- Crow, Y.J., Hayward, B.E., Parmar, R., Robins, P., Leitch, A., Ali, M., Black, D.N., van Bokhoven, H., Brunner, H.G., Hamel, B.C., *et al.* (2006). Mutations in the gene encoding the 3[prime]-5[prime] DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus. *Nat Genet* 38, 917-920.
- Davies, B.W., Bogard, R.W., Young, T.S., and Mekalanos, J.J. (2012). Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. *Cell* 149, 358-370.
- Deng, L., Liang, H., Xu, M., Yang, X., Burnette, B., Arina, A., Li, X.D., Mauceri, H., Beckett, M., Darga, T., *et al.* (2014). STING-Dependent Cytosolic DNA Sensing Promotes Radiation-Induced Type I Interferon-Dependent Antitumor Immunity in Immunogenic Tumors. *Immunity* 41, 843-852.
- Dey, B., Dey, R.J., Cheung, L.S., Pokkali, S., Guo, H., Lee, J.-H., and Bishai, W.R. (2015). A bacterial cyclic dinucleotide activates the cytosolic surveillance pathway and mediates innate resistance to tuberculosis. *Nat Med advance online publication*.
- Diebold, S.S., Kaisho, T., Hemmi, H., Akira, S., and Reis e Sousa, C. (2004). Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* 303, 1529-1531.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta crystallographica Section D, Biological crystallography* 66, 486-501.
- Fernandes-Alnemri, T., Yu, J.-W., Datta, P., Wu, J., and Alnemri, E.S. (2009). AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* 458, 509-513.
- Fitzgerald, K.A., McWhirter, S.M., Faia, K.L., Rowe, D.C., Latz, E., Golenbock, D.T., Coyle, A.J., Liao, S.M., and Maniatis, T. (2003). IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway. *Nat Immunol* 4, 491-496.

- Gall, A., Treuting, P., Elkon, K.B., Loo, Y.M., Gale, M., Jr., Barber, G.N., and Stetson, D.B. (2012). Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease. *Immunity* 36, 120-131.
- Gao, D., Wu, J., Wu, Y.T., Du, F., Aroh, C., Yan, N., Sun, L., and Chen, Z.J. (2013a). Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science* 341, 903-906.
- Gao, P., Ascano, M., Wu, Y., Barchet, W., Gaffney, Barbara L., Zillinger, T., Serganov, Artem A., Liu, Y., Jones, Roger A., Hartmann, G., *et al.* (2013b). Cyclic [G(2',5')pA(3',5')p] Is the Metazoan Second Messenger Produced by DNA-Activated Cyclic GMP-AMP Synthase. *Cell* 153, 1094-1107.
- Gehrke, N., Mertens, C., Zillinger, T., Wenzel, J., Bald, T., Zahn, S., Tuting, T., Hartmann, G., and Barchet, W. (2013). Oxidative damage of DNA confers resistance to cytosolic nuclelease TREX1 degradation and potentiates STING-dependent immune sensing. *Immunity* 39, 482-495.
- Goubau, D., Deddouche, S., and Reis e Sousa, C. (2013). Cytosolic Sensing of Viruses. *Immunity* 38, 855-869.
- Hansen, K., Prabakaran, T., Laustsen, A., Jørgensen, S.E., Rahbæk, S.H., Jensen, S.B., Nielsen, R., Leber, J.H., Decker, T., Horan, K.A., *et al.* (2014a). Listeria monocytogenes induces IFN β expression through an IFI16-, cGAS- and STING-dependent pathway. *Embo j* 33, 1654-1666.
- Hansen, K., Prabakaran, T., Laustsen, A., Jørgensen, S.E., Rahbæk, S.H., Jensen, S.B., Nielsen, R., Leber, J.H., Decker, T., Horan, K.A., *et al.* (2014b). Listeria monocytogenes induces IFN β expression through an IFI16-, cGAS- and STING-dependent pathway. *The EMBO Journal* 33, 1654-1666.
- Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H., and Bauer, S. (2004). Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* 303, 1526-1529.
- Hemmi, H., Kaisho, T., Takeuchi, O., Sato, S., Sanjo, H., Hoshino, K., Horiuchi, T., Tomizawa, H., Takeda, K., and Akira, S. (2002). Small anti-viral compounds activate immune cells via the TLR7 MyD88-dependent signaling pathway. *Nat Immunol* 3, 196-200.
- Hemmi, H., Takeuchi, O., Kawai, T., Kaisho, T., Sato, S., Sanjo, H., Matsumoto, M., Hoshino, K., Wagner, H., Takeda, K., *et al.* (2000). A Toll-like receptor recognizes bacterial DNA. *Nature* 408, 740-745.

- Honda, K., and Taniguchi, T. (2006). IRFs: master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors. *Nat Rev Immunol* *6*, 644-658.
- Hornung, V., Ablasser, A., Charrel-Dennis, M., Bauernfeind, F., Horvath, G., Caffrey, D.R., Latz, E., and Fitzgerald, K.A. (2009). AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* *458*, 514-518.
- Hou, F., Sun, L., Zheng, H., Skaug, B., Jiang, Q.X., and Chen, Z.J. (2011). MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell* *146*, 448-461.
- Houtman, J.C., Brown, P.H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L.E., and Schuck, P. (2007). Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein science : a publication of the Protein Society* *16*, 30-42.
- Ishii, K.J., Kawagoe, T., Koyama, S., Matsui, K., Kumar, H., Kawai, T., Uematsu, S., Takeuchi, O., Takeshita, F., Coban, C., *et al.* (2008). TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines. *Nature* *451*, 725-729.
- Ishikawa, H., and Barber, G.N. (2011). The STING pathway and regulation of innate immune signaling in response to DNA pathogens. *Cellular and molecular life sciences : CMLS* *68*, 1157-1165.
- Ishikawa, H., Ma, Z., and Barber, G.N. (2009). STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* *461*, 788-792.
- Jeremiah, N., Neven, B., xE, xE, dicte, Gentili, M., Callebaut, I., Maschalidi, S., Stolzenberg, M.-C., Goudin, N., *et al.* (2014). Inherited STING-activating mutation underlies a familial inflammatory syndrome with lupus-like manifestations. *The Journal of Clinical Investigation* *124*, 5516-5520.
- Jiang, X., Kinch, L.N., Brautigam, C.A., Chen, X., Du, F., Grishin, N.V., and Chen, Z.J. (2012). Ubiquitin-induced oligomerization of the RNA sensors RIG-I and MDA5 activates antiviral innate immune response. *Immunity* *36*, 959-973.
- Kawai, T., and Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol* *11*, 373-384.
- Keller, S., Vargas, C., Zhao, H., Piszczeck, G., Brautigam, C.A., and Schuck, P. (2012). High-precision isothermal titration calorimetry with automated peak-shape analysis. *Analytical chemistry* *84*, 5066-5073.

- Kranzusch, P.J., Lee, A.S., Berger, J.M., and Doudna, J.A. (2013). Structure of human cGAS reveals a conserved family of second-messenger enzymes in innate immunity. *Cell reports* 3, 1362-1368.
- Kranzusch, P.J., Lee, A.S., Wilson, S.C., Solovykh, M.S., Vance, R.E., Berger, J.M., and Doudna, J.A. (2014). Structure-guided reprogramming of human cGAS dinucleotide linkage specificity. *Cell* 158, 1011-1021.
- Kuchta, K., Knizewski, L., Wyrwicz, L.S., Rychlewski, L., and Ginalski, K. (2009). Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. *Nucleic Acids Research* 37, 7701-7714.
- Lahaye, X., Satoh, T., Gentili, M., Cerboni, S., Conrad, C., Hurbain, I., El Marjou, A., Lacabaratz, C., Lelievre, J.D., and Manel, N. (2013). The capsids of HIV-1 and HIV-2 determine immune detection of the viral cDNA by the innate sensor cGAS in dendritic cells. *Immunity* 39, 1132-1142.
- Lam, E., Stein, S., and Falck-Pedersen, E. (2014). Adenovirus detection by the cGAS/STING/TBK1 DNA sensing cascade. *J Virol* 88, 974-981.
- Li, L., Yin, Q., Kuss, P., Maliga, Z., Millán, J.L., Wu, H., and Mitchison, T.J. (2014). Hydrolysis of 2'3'-cGAMP by ENPP1 and design of nonhydrolyzable analogs. *Nat Chem Biol* 10, 1043-1048.
- Li, X., Shu, C., Yi, G., Chaton, Catherine T., Shelton, Catherine L., Diao, J., Zuo, X., Kao, C.C., Herr, Andrew B., and Li, P. (2013a). Cyclic GMP-AMP Synthase Is Activated by Double-Stranded DNA-Induced Oligomerization. *Immunity* 39, 1019-1031.
- Li, X.-D., and Chen, Z.J. (2012). Sequence specific detection of bacterial 23S ribosomal RNA by TLR13, Vol 1.
- Li, X.D., Wu, J., Gao, D., Wang, H., Sun, L., and Chen, Z.J. (2013b). Pivotal roles of cGAS-cGAMP signaling in antiviral defense and immune adjuvant effects. *Science* 341, 1390-1394.
- Liu, S., Chen, J., Cai, X., Wu, J., Chen, X., Wu, Y.-T., Sun, L., and Chen, Z.J. (2013). MAVS recruits multiple ubiquitin E3 ligases to activate antiviral signaling cascades. *eLife* 2, e00785.
- Liu, Y., Jesus, A.A., Marrero, B., Yang, D., Ramsey, S.E., Montealegre Sanchez, G.A., Tenbrock, K., Wittkowski, H., Jones, O.Y., Kuehn, H.S., *et al.* (2014). Activated STING in a Vascular and Pulmonary Syndrome. *New England Journal of Medicine* 371, 507-518.

- Manzanillo, P.S., Shiloh, M.U., Portnoy, D.A., and Cox, J.S. (2012). Mycobacterium tuberculosis activates the DNA-dependent cytosolic surveillance pathway within macrophages. *Cell host & microbe* 11, 469-480.
- Melroe, G.T., DeLuca, N.A., and Knipe, D.M. (2004). Herpes Simplex Virus 1 Has Multiple Mechanisms for Blocking Virus-Induced Interferon Production. *Journal of Virology* 78, 8411-8420.
- Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006). HKL-3000: the integration of data reduction and structure solution - from diffraction images to an initial model in minutes. *Acta Crystallographica Section D* 62, 859-866.
- Mossman, K.L., and Smiley, J.R. (2002). Herpes Simplex Virus ICP0 and ICP34.5 Counteract Distinct Interferon-Induced Barriers to Virus Replication. *Journal of Virology* 76, 1995-1998.
- Negishi, H., Osawa, T., Ogami, K., Ouyang, X., Sakaguchi, S., Koshiba, R., Yanai, H., Seko, Y., Shitara, H., Bishop, K., et al. (2008). A critical link between Toll-like receptor 3 and type II interferon signaling pathways in antiviral innate immunity. *Proceedings of the National Academy of Sciences of the United States of America* 105, 20446-20451.
- Oldenburg, M., Krüger, A., Ferstl, R., Kaufmann, A., Nees, G., Sigmund, A., Bathke, B., Lauterbach, H., Suter, M., Dreher, S., et al. (2012). TLR13 Recognizes Bacterial 23S rRNA Devoid of Erythromycin Resistance-Forming Modification. *Science* 337, 1111-1115.
- Ouyang, S., Song, X., Wang, Y., Ru, H., Shaw, N., Jiang, Y., Niu, F., Zhu, Y., Qiu, W., Parvatiyar, K., et al. (2012). Structural analysis of the STING adaptor protein reveals a hydrophobic dimer interface and mode of cyclic di-GMP binding. *Immunity* 36, 1073-1086.
- Padrick, S.B., Deka, R.K., Chuang, J.L., Wynn, R.M., Chuang, D.T., Norgard, M.V., Rosen, M.K., and Brautigam, C.A. (2010). Determination of Protein Complex Stoichiometry Through Multisignal Sedimentation Velocity Experiments. *Analytical biochemistry* 407, 89-103.
- Peisley, A., Wu, B., Xu, H., Chen, Z.J., and Hur, S. (2014). Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I. *Nature* 509, 110-114.
- Platani, L.C. (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat Rev Immunol* 5, 375-386.
- Rossjohn, J., Polekhina, G., Feil, S.C., Morton, C.J., Tweten, R.K., and Parker, M.W. (2007). Structures of Perfringolysin O Suggest a Pathway for Activation of Cholesterol-dependent Cytolysins. *Journal of molecular biology* 367, 1227-1236.

- Samaniego, L.A., Neiderhiser, L., and DeLuca, N.A. (1998). Persistence and Expression of the Herpes Simplex Virus Genome in the Absence of Immediate-Early Proteins. *Journal of Virology* 72, 3307-3320.
- Sauer, J.-D., Sotelo-Troha, K., von Moltke, J., Monroe, K.M., Rae, C.S., Brubaker, S.W., Hyodo, M., Hayakawa, Y., Woodward, J.J., Portnoy, D.A., *et al.* (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. *Infection and Immunity* 79, 688-694.
- Schoggins, J.W., MacDuff, D.A., Imanaka, N., Gainey, M.D., Shrestha, B., Eitson, J.L., Mar, K.B., Richardson, R.B., Ratushny, A.V., Litvak, V., *et al.* (2014). Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity. *Nature* 505, 691-695.
- Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* 472, 481-485.
- Seth, R.B., Sun, L., Ea, C.K., and Chen, Z.J. (2005). Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* 122, 669-682.
- Sharma, S., DeOliveira, Rosane B., Kalantari, P., Parroche, P., Goutagny, N., Jiang, Z., Chan, J., Bartholomeu, Daniella C., Lauw, F., Hall, J.P., *et al.* (2011). Innate Immune Recognition of an AT-Rich Stem-Loop DNA Motif in the Plasmodium falciparum Genome. *Immunity* 35, 194-207.
- Sharma, S., tenOever, B.R., Grandvaux, N., Zhou, G.P., Lin, R., and Hiscott, J. (2003). Triggering the interferon antiviral response through an IKK-related pathway. *Science* 300, 1148-1151.
- Shu, C., Yi, G., Watts, T., Kao, C.C., and Li, P. (2012). Structure of STING bound to cyclic di-GMP reveals the mechanism of cyclic dinucleotide recognition by the immune system. *Nature structural & molecular biology* 19, 722-724.
- Silverman, N., and Maniatis, T. (2001). NF-kappaB signaling pathways in mammalian and insect innate immunity. *Genes & development* 15, 2321-2342.
- Silvin, A., and Manel, N. (2015). Innate immune sensing of HIV infection. *Current Opinion in Immunology* 32, 54-60.
- Stetson, D.B., Ko, J.S., Heidmann, T., and Medzhitov, R. (2008). Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* 134, 587-598.

- Storek, K.M., Gertsvolf, N.A., Ohlson, M.B., and Monack, D.M. (2015). cGAS and Ifi204 Cooperate To Produce Type I IFNs in Response to Francisella Infection. *Journal of immunology* (Baltimore, Md : 1950).
- Sun, Q., Sun, L., Liu, H.-H., Chen, X., Seth, R.B., Forman, J., and Chen, Z.J. (2006). The Specific and Essential Role of MAVS in Antiviral Innate Immune Responses. *Immunity* 24, 633-642.
- Takaoka, A., Wang, Z., Choi, M.K., Yanai, H., Negishi, H., Ban, T., Lu, Y., Miyagishi, M., Kodama, T., Honda, K., *et al.* (2007). DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* 448, 501-505.
- Tanaka, Y., and Chen, Z.J. (2012). STING Specifies IRF3 Phosphorylation by TBK1 in the Cytosolic DNA Signaling Pathway, Vol 5.
- Unterholzner, L., Keating, S.E., Baran, M., Horan, K.A., Jensen, S.B., Sharma, S., Sirois, C.M., Jin, T., Latz, E., Xiao, T.S., *et al.* (2010). IFI16 is an innate immune sensor for intracellular DNA. *Nat Immunol* 11, 997-1004.
- West, A.P., Khouri-Hanold, W., Staron, M., Tal, M.C., Pineda, C.M., Lang, S.M., Bestwick, M., Duguay, B.A., Raimundo, N., MacDuff, D.A., *et al.* (2015). Mitochondrial DNA stress primes the antiviral innate immune response. *Nature*.
- Woo, S.R., Fuertes, M.B., Corrales, L., Spranger, S., Furdyna, M.J., Leung, M.Y., Duggan, R., Wang, Y., Barber, G.N., Fitzgerald, K.A., *et al.* (2014). STING-dependent cytosolic DNA sensing mediates innate immune recognition of immunogenic tumors. *Immunity* 41, 830-842.
- Woodward, J.J., Iavarone, A.T., and Portnoy, D.A. (2010). c-di-AMP secreted by intracellular Listeria monocytogenes activates a host type I interferon response. *Science* 328, 1703-1705.
- Wu, J., and Chen, Z.J. (2014). Innate immune sensing and signaling of cytosolic nucleic acids. *Annual review of immunology* 32, 461-488.
- Yan, N., Regalado-Magdos, A.D., Stiggebout, B., Lee-Kirsch, M.A., and Lieberman, J. (2010). The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nat Immunol* 11, 1005-1013.
- Yin, Q., Tian, Y., Kabaleeswaran, V., Jiang, X., Tu, D., Eck, M.J., Chen, Z.J., and Wu, H. (2012). Cyclic di-GMP sensing via the innate immune signaling protein STING. *Mol Cell* 46, 735-745.
- Yoneyama, M., and Fujita, T. (2009). RNA recognition and signal transduction by RIG-I-like receptors. *Immunological reviews* 227, 54-65.

- Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and 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., and Chen, Z.J. (2010). Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell* 141, 315-330.
- Zhang, X., Shi, H., Wu, J., Zhang, X., Sun, L., Chen, C., and Chen, Z.J. (2013). Cyclic GMP-AMP Containing Mixed Phosphodiester Linkages Is An Endogenous High Affinity Ligand for STING. *Molecular cell* 51, 10.1016/j.molcel.2013.1005.1022.
- Zhang, X., Wu, J., Du, F., Xu, H., Sun, L., Chen, Z., Brautigam, C.A., Zhang, X., and Chen, Z.J. (2014a). The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. *Cell reports* 6, 421-430.
- Zhang, Y., Yeruva, L., Marinov, A., Prantner, D., Wyrick, P.B., Lupashin, V., and Nagarajan, U.M. (2014b). The DNA sensor, cyclic GMP-AMP synthase, is essential for induction of IFN-beta during Chlamydia trachomatis infection. *Journal of immunology* (Baltimore, Md : 1950) 193, 2394-2404.
- Zhang, Z., Yuan, B., Bao, M., Lu, N., Kim, T., and Liu, Y.J. (2011). The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. *Nat Immunol* 12, 959-965.