HHS Public Access

Author manuscript

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

Published in final edited form as:

Nat Immunol. 2015 October ; 16(10): 1025–1033. doi:10.1038/ni.3267.

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

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

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

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

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

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

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

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

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

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

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

Abstract

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

Correspondence should be addressed to A.-M.H. (aherzner@uni-bonn.de) or M.S. (martin.schlee@uni-bonn.de).
[10]Present addresses: Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts, USA (C.A.H.), and Microsynth, Balgach, Switzerland (D.A.).
[11]These authors contributed equally to this work.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

M.S., A.-M.H., J.L., C.A.H. and G.H., conceptualization; M.S., A.-M.H., C.A.H., M.G., S. Wolter, K.K., T.G., L.A., K.-P.H., T.Z., C.M., T.J., T.S.X. and D.A., methodology; A.-M.H., C.A.H., M.G., D.A., T.J., T.S.X. and M.S., formal analysis; A.-M.H., M.S., C.A.H., M.G., S. Wolter, T.G., L.A., T.Z., C.M. and T.J., investigation; M.S., A.-M.H., E.B., C.A.H., V.H., W.B. and G.H., writing of the original draft; A.-M.H., M.S., E.B., W.B., C.C. and G.H., writing (review and editing); M.S., G.H., V.H., W.B., C.C., E.B., K.-P.H. and T.G., funding acquisition; T.J., T.S.X., S. Wittmann, T.G., L.A. and K.-P.H., resources; and M.S., T.G., K.-P.H., C.C., W.B. and G.H., supervision.

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

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

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

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

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

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

RESULTS

Detection of unpaired guanosines in HIV cDNA stem loops

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

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

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

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

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

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

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

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

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

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

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

indicated G quadruplex–mediated oligomerization (sequences and structures, Fig. 3b), while G_2-YSD or G_3-YSD migrated as a single, low-molecular-weight band (Fig. 3c). Since strong electric fields may interrupt weak interactions during electrophoresis, we investigated the molecular interactions in solution. Anti-parallel G quadruplexes are arranged chirally and are optically active at 295 nm (ref. 39). Their formation can be detected by circular dichroism spectroscopy, which measures the difference in the absorption of left-circularly polarized light and that of right-circularly polarized light (ellipticity) of the analyte at 30–70 °C. G_4-YSD and G_5-YSD exhibited a substantial positive ellipticity at 295 nm that was lost at high temperatures (≥70 °C) (Fig. 3d), which indicated G-quadruplex formation at physiological temperatures. In contrast, G_2-YSD or G_3-YSD, like the negative controls A_3-YSD and a blunt-ended 30-nucleotide DNA, did not show considerable optical activity at 295 nm (Fig. 3d), which indicated that G_2-YSD and G_3-YSD were present as monomeric duplexes. Notably, the quadruplex-forming YSDs did not induce larger amounts of IFN-α than the prototype G_3-YSDs did (Fig. 3e). To exclude the intracellular formation of G quadruplexes, we made use of 26-nucleotide G_2-YSD, which was as active as 20-nucleotide G_3-YSD (Fig. 3e). Here, we replaced the terminal guanosines with 7-deaza-guanosine or inosine. Substitution of nitrogen N7 of the guanine by a C-H group (7-deaza-guanosine) and the missing amine group at carbon C3 of inosine disables formation of the hydrogen bonds essential for the G-quadruplex arrangement[40]. Substitution with 7-deaza-guanosine as well as substitution with inosine did not substantially alter the immunostimulatory activity (Fig. 3f). We concluded that intermolecular G-quadruplex interactions were not involved in the recognition of G-ended YSD but that the monomeric duplex was detected and induced IFN-α/β.

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

Of all human cell lines we tested, only the human monocytic cell line THP-1 was responsive to G_3-YSD (data not shown), in line with what has been described for dsDNA with a random base composition[41]. The RNA helicase RIG-I has been linked to the detection of long AT-rich DNA[7,8]. To exclude the possibility of involvement of signaling via RIG-I and the signaling adaptor MAVS in the recognition of YSD in this cell line, we performed RNA-mediated interference (RNAi) with small interfering (siRNA) targeting MAVS or STING. Knockdown of STING substantially repressed the induction of IFN-α/β by G_3-YSD and plasmid DNA, but knockdown of MAVS did not; instead, knockdown of MAVS repressed the response to the RIG-I ligand 3P-dsRNA (triphosphorylated double-stranded RNA) but not to any of the DNA ligands (Fig. 4a and **Supplementary Fig. 4a**). These results demonstrated that G_3-YSD indeed induced secretion of IFN-α/β via a STING-dependent pathway.

To identify the receptor responsible for the recognition of G_3-YSD, we expressed Flag-tagged candidate receptors in HEK293T human embryonic kidney cells and performed bead-coupled co-precipitation experiments with immobilized DNA ligands (Fig. 4b–d). IFI16, ZBP1, DDX41 and cGAS have been described as IFN-α/β-inducing cytosolic receptors for dsDNA in fibroblasts, monocytes, macrophages or conventional dendritic cells[18,19,29,42,43]. Among those candidates, IFI16 and cGAS bound to both G_3-YSD and C_3-YSD, while ZBP1, DDX41 and the negative control RIG-I did not detectably precipitate

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

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

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

The interferon response correlates with viral ssDNA content

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

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

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

Long ssDNA is highly immunostimulatory

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

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

cGAS senses unpaired G in human retroelement sequences

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

DISCUSSION

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

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

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

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

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

ONLINE METHODS

Ethics statement

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

Stimulatory nucleic acids

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

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

Cell culture and stimulation

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

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

PAGE of nucleic acids

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

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

Circular dichroism spectroscopy

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

Detection of cytosolic 5′ IRD-700–labeled DNA

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

DNA interaction assays

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

cGAS *in vitro* activity assay

For *in vitro* reactions of cGAS in the presence of varying dsDNA stimuli, 2 μM recombinant human cGAS (truncated version: residues 155–522) was mixed with 88 ng/μl hybridized DNA, 0.1 mM ATP and 0.1 mM GTP in buffer A (100 mM NaCl, 40 mM Tris, pH 7.5, and 10 mM $MgCl_2$). After 90 min of incubation at 37 °C, the reaction mixture was analyzed by reverse-phase HPLC. Samples were prepared in 300 mM triethylammonium acetate, then were applied to a 4.1- × 250-mm PRP-1 column (Hamilton) and separated for 18 min at a flow rate of 1 ml/min in an isocratic gradient of 100 mM ammonium acetate.

RNAi

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

Quantitative PCR (RNA detection)

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

Lentiviral particles

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

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

Infection of monocyte-derived macrophages and THP-1 cells

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

Strand-specific detection of HIV reverse transcripts

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

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

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

Preparation of ssDNA from PCR products

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

Statistical analysis

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

Acknowledgments

We thank C. Siering for help with circular dichroism spectroscopy, and S. Schmitt for discussions. Supported by Deutsche Forschungsgemeinschaft (SFB670 to M.S., W.B., V.H. and G.H.; DFG SCHL1930/1–1 to M.S.; SFB704 to G.H., V.H. and W.B.; and SFB832 and KFO177 to C.C. and G.H.), the Deutsche Forschungsgemeinschaft Excellence Cluster ImmunoSensation (G.H., M.S., V.H., E.B. and W.B.), BONFOR of the University of Bonn (E.B.) and the German Center of Infectious Disease (G.H., V.H. and W.B.).

References

1. Krieg AM, et al. CpG motifs in bacterial DNA trigger direct B-cell activation. Nature. 1995; 374:546–549. [PubMed: 7700380]
2. Hemmi H, et al. A Toll-like receptor recognizes bacterial DNA. Nature. 2000; 408:740–745. [PubMed: 11130078]

3. Hartmann G, et al. Rational design of new CpG oligonucleotides that combine B cell activation with high IFN-alpha induction in plasmacytoid dendritic cells. Eur J Immunol. 2003; 33:1633–1641. [PubMed: 12778481]
4. Hornung V, et al. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. Nature. 2009; 458:514–518. [PubMed: 19158675]
5. Fernandes-Alnemri T, Yu JW, Datta P, Wu J, Alnemri ES. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. Nature. 2009; 458:509–513. [PubMed: 19158676]
6. Bürckstümmer T, et al. An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. Nat Immunol. 2009; 10:266–272. [PubMed: 19158679]
7. Ablasser A, et al. RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. Nat Immunol. 2009; 10:1065–1072. [PubMed: 19609254]
8. Chiu YH, Macmillan JB, Chen ZJ. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. Cell. 2009; 138:576–591. [PubMed: 19631370]
9. Schlee M, et al. Recognition of 5′triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. Immunity. 2009; 31:25–34. [PubMed: 19576794]
10. Cavlar T, Ablasser A, Hornung V. Induction of type I IFNs by intracellular DNA-sensing pathways. Immunol Cell Biol. 2012; 90:474–482. [PubMed: 22450802]
11. Stetson DB, Medzhitov R. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. Immunity. 2006; 24:93–103. [PubMed: 16413926]
12. Yang P, et al. The cytosolic nucleic acid sensor LRRFIP1 mediates the production of type I interferon via a beta-catenin-dependent pathway. Nat Immunol. 2010; 11:487–494. [PubMed: 20453844]
13. Bagashev A, et al. Leucine-rich repeat (in Flightless I) interacting protein-1 regulates a rapid type I interferon response. J Interferon Cytokine Res. 2010; 30:843–852. [PubMed: 20586614]
14. Zhang X, et al. Cutting edge: Ku70 is a novel cytosolic DNA sensor that induces type III rather than type I IFN. J Immunol. 2011; 186:4541–4545. [PubMed: 21398614]
15. Kim T, et al. Aspartate-glutamate-alanine-histidine box motif (DEAH)/RNA helicase A helicases sense microbial DNA in human plasmacytoid dendritic cells. Proc Natl Acad Sci USA. 2010; 107:15181–15186. [PubMed: 20696886]
16. Ishikawa H, Ma Z, Barber GN. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. Nature. 2009; 461:788–792. [PubMed: 19776740]
17. Zhong B, et al. The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. Immunity. 2008; 29:538–550. [PubMed: 18818105]
18. Unterholzner L, et al. IFI16 is an innate immune sensor for intracellular DNA. Nat Immunol. 2010; 11:997–1004. [PubMed: 20890285]
19. Sun L, Wu J, Du F, Chen X, Chen ZJ. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. Science. 2013; 339:786–791. [PubMed: 23258413]
20. Jakobsen MR, Paludan SR. IFI16: At the interphase between innate DNA sensing and genome regulation. Cytokine Growth Factor Rev. 2014; 25:649–655. [PubMed: 25027602]
21. Wu J, et al. Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. Science. 2013; 339:826–830. [PubMed: 23258412]
22. Gao P, et al. Cyclic [G(2′,5′)pA(3′,5′)p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. Cell. 2013; 153:1094–1107. [PubMed: 23647843]
23. Ablasser A, et al. cGAS produces a 2–5′′-linked cyclic dinucleotide second messenger that activates STING. Nature. 2013; 498:380–384. [PubMed: 23722158]
24. Civril F, et al. Structural mechanism of cytosolic DNA sensing by cGAS. Nature. 2013; 498:332–337. [PubMed: 23722159]
25. Kato K, et al. Structural and functional analyses of DNA-sensing and immune activation by human cGAS. PLoS ONE. 2013; 8:e76983. [PubMed: 24116191]

26. Kranzusch PJ, Lee AS, Berger JM, Doudna JA. Structure of human cGAS reveals a conserved family of second-messenger enzymes in innate immunity. Cell Rep. 2013; 3:1362–1368. [PubMed: 23707061]
27. Gao D, et al. Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. Science. 2013; 341:903–906. [PubMed: 23929945]
28. Zhang X, et al. The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. Cell Rep. 2014; 6:421–430. [PubMed: 24462292]
29. Li X, et al. Cyclic GMP-AMP synthase is activated by double-stranded DNA-induced oligomerization. Immunity. 2013; 39:1019–1031. [PubMed: 24332030]
30. Li XD, et al. Pivotal roles of cGAS-cGAMP signaling in antiviral defense and immune adjuvant effects. Science. 2013; 341:1390–1394. [PubMed: 23989956]
31. Lam E, Stein S, Falck-Pedersen E. Adenovirus detection by the cGAS/STING/TBK1 DNA sensing cascade. J Virol. 2014; 88:974–981. [PubMed: 24198409]
32. Lahaye X, et al. The capsids of HIV-1 and HIV-2 determine immune detection of the viral cDNA by the innate sensor cGAS in dendritic cells. Immunity. 2013; 39:1132–1142. [PubMed: 24269171]
33. Rasaiyaah J, et al. HIV-1 evades innate immune recognition through specific cofactor recruitment. Nature. 2013; 503:402–405. [PubMed: 24196705]
34. Suzuki K, et al. Activation of target-tissue immune-recognition molecules by double-stranded polynucleotides. Proc Natl Acad Sci USA. 1999; 96:2285–2290. [PubMed: 10051633]
35. Ishii KJ, et al. A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. Nat Immunol. 2006; 7:40–48. [PubMed: 16286919]
36. Jakobsen MR, et al. From the Cover: IFI16 senses DNA forms of the lentiviral replication cycle and controls HIV-1 replication. Proc Natl Acad Sci USA. 2013; 110:E4571–E4580. [PubMed: 24154727]
37. Galvis AE, Fisher HE, Nitta T, Fan H, Camerini D. Impairment of HIV-1 cDNA Synthesis by DBR1 Knockdown. J Virol. 2014; 88:7054–7069. [PubMed: 24672043]
38. Gehrke N, et al. Oxidative damage of DNA confers resistance to cytosolic nuclease TREX1 degradation and potentiates STING-dependent immune sensing. Immunity. 2013; 39:482–495. [PubMed: 23993650]
39. Huppert JL. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. Chem Soc Rev. 2008; 37:1375–1384. [PubMed: 18568163]
40. Murchie AI, Lilley DM. Retinoblastoma susceptibility genes contain 5′ sequences with a high propensity to form guanine-tetrad structures. Nucleic Acids Res. 1992; 20:49–53. [PubMed: 1738603]
41. Hagmann CA, et al. RIG-I detects triphosphorylated RNA of Listeria monocytogenes during infection in non-immune cells. PLoS ONE. 2013; 8:e62872. [PubMed: 23653683]
42. Takaoka A, et al. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. Nature. 2007; 448:501–505. [PubMed: 17618271]
43. Zhang Z, et al. The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. Nat Immunol. 2011; 12:959–965. [PubMed: 21892174]
44. Fisher TS, Darden T, Prasad VR. Mutations proximal to the minor groove-binding track of human immunodeficiency virus type 1 reverse transcriptase differentially affect utilization of RNA versus DNA as template. J Virol. 2003; 77:5837–5845. [PubMed: 12719577]
45. Mankan AK, et al. Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. EMBO J. 2014; 33:2937–2946. [PubMed: 25425575]
46. Stetson DB, Ko JS, Heidmann T, Medzhitov R. Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell. 2008; 134:587–598. [PubMed: 18724932]
47. Ablasser A, et al. TREX1 deficiency triggers cell-autonomous immunity in a cGAS-dependent manner. J Immunol. 2014; 192:5993–5997. [PubMed: 24813208]
48. Yoh SM, et al. PQBP1 Is a Proximal Sensor of the cGAS-Dependent Innate Response to HIV-1. Cell. 2015; 161:1293–1305. [PubMed: 26046437]

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

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

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

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

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

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

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

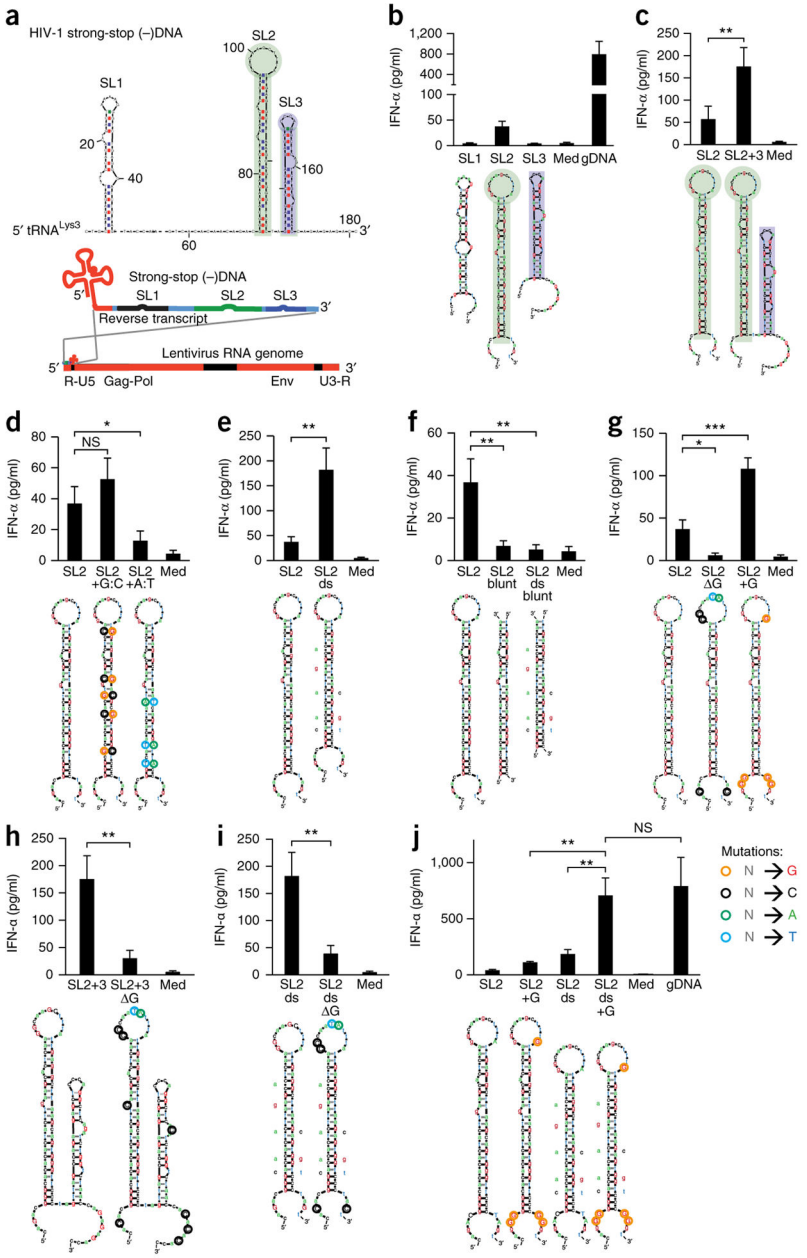

Figure 1.

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

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

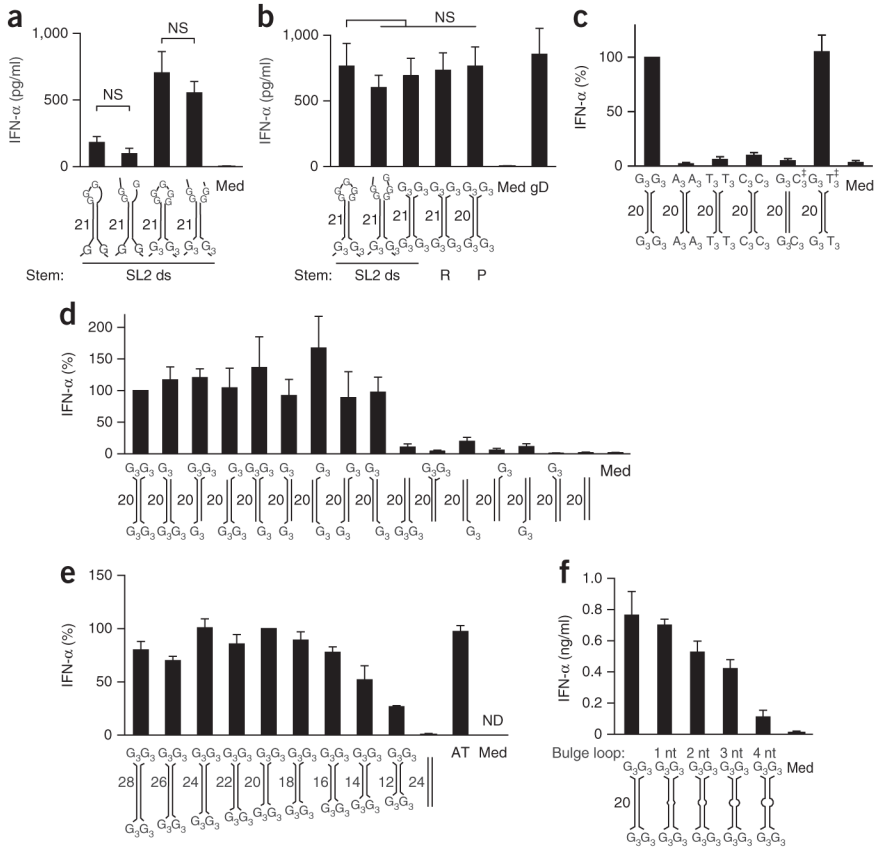

Figure 2.

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

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

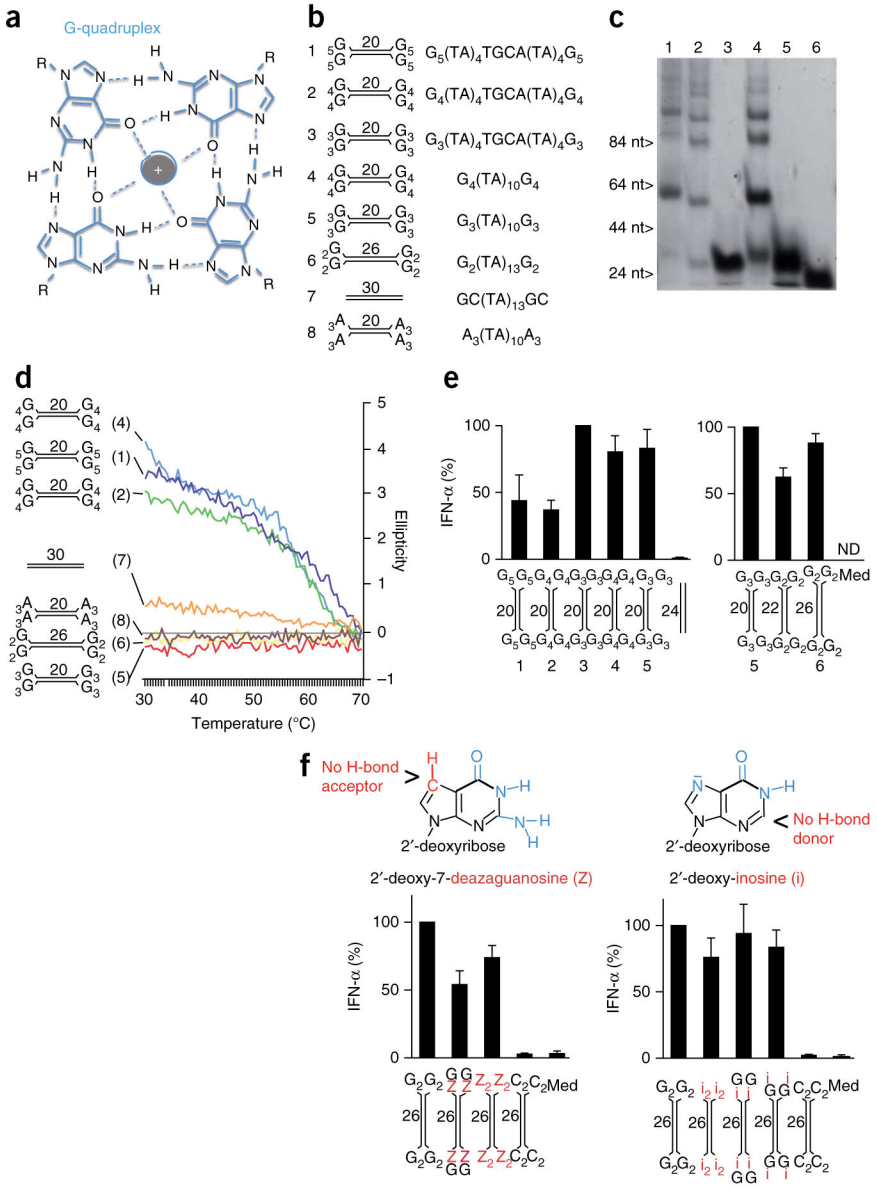

Figure 3.

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

substitution of terminal guanosines (G) with deazaguanosine (Z; left) or inosine (i; right); top, chemical structure of 2′-deoxy-7-deazaguanosine and 2′-deoxy-inosine. Numbers adjacent to stems (**b**,**d**–**f**) indicate stem length (sequences of DNA structures, **Supplementary Table 1**). Data are representative of two experiments (**c**,**d**) or are pooled from two (**e**, **f**, right) or three (**f**, right) experiments with one or two biological replicates in each (**e**,**f**; mean and s.e.m. of $n = 3$ donors (**e**), $n = 6$ donors (**f**, left) or $n = 4$ donors (**f**, right)).

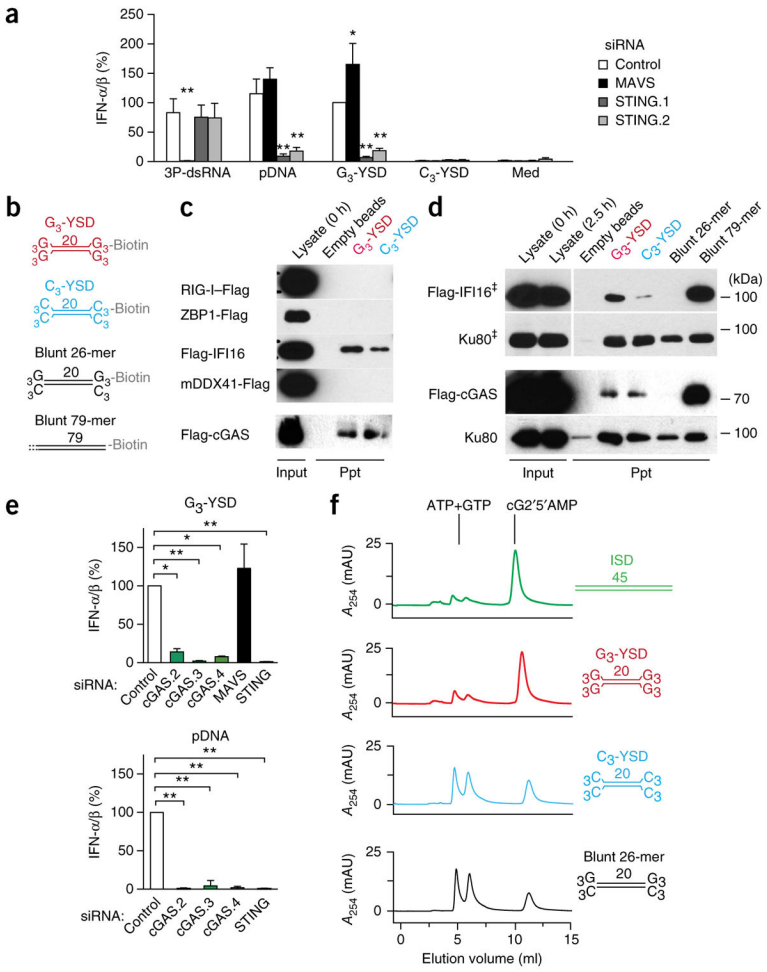

Figure 4.

G-YSD activates the cGAS-STING axis. (**a**) IFN-α/β activity in supernatants of THP-1 cells treated with control siRNA or siRNA targeting MAVS or STING (either of two target sequences (STING.1 or STING.2)) (key) and then, 72 h later, stimulated for 20 h with the RIG-I ligand 3P-dsRNA, plasmid DNA (pDNA), G_3-YSD or C_3-YSD; results are presented relative to those of THP-1 cells treated with control siRNA and stimulated with G_3-YSD, set as 100%. (**b**) Secondary structures of biotinylated DNA duplexes used in interaction assays in **c**,**d**: G_3-YSD or C_3-YSD, or short, 26-nucleotide (26-mer) or long, 79-nucleotide (79-mer) blunt dsDNA. (**c**) Immunoblot analysis of Flag-tagged receptor candidates (left margin) with lysate alone (0.1 volume of the lysate used in the precipitation assays; Input) at the beginning (0 h) of the assay (far left lane) or after precipitation (Ppt) with empty beads (middle left lane) or with G_3-YSD or C_3-YSD (above lanes; as in **b**). (**d**) Immunoblot analysis of Flag-tagged IFI16 or cGAS or of the nonspecific DNA-binding control Ku80 with lysate alone (Input) at the beginning (0 h) or end (2.5 h) of the experiment, or after precipitation with empty beads, G_3-YSD, C_3-YSD (as in **c**), or with short (26-nucleotide) or long (79-nucleotide) blunt dsDNA (as in **b**) (above lanes). ‡, lysate and precipitate detection are from the same blot with one empty lane removed. (**e**) IFN-α/β in supernatants of THP-1 cells treated with control siRNA or siRNA targeting cGAS (one of three target sequences

(cGAS.2, cGAS.3 or cGAS.4)), MAVS or STING (horizontal axes) and then, 48 h later, stimulated for 20 h with G_3-YSD (top) or plasmid DNA (bottom); results are presented relative to those of THP-1 cells treated with control siRNA, set as 100%. (**f**) HPLC detecting the conversion of ATP and GTP to cG2′5′AMP (cyclic [G(2′,5′)pA(3′,5′)p]) by cGAS in the presence of ISD, G_3-YSD, C_3-YSD or the 26-nucleotide blunt DNA (without biotin) in **b**, presented (in milli-absorption units (mAU)) as absorption at 254 nm (A_{254}). Sequences of DNA structures, **Supplementary Table 1**. *$P \leq 0.01$ and **$P \leq 0.001$, presented only for results significantly different from control (two-way ANOVA followed by Bonferroni's post-hoc test (**a**) or one-way ANOVA followed by Tukey's post-hoc test (**e**)). Data are pooled from three (**a**,**e**, bottom) or four (**e**, top) independent experiments with biological replicates (mean and s.e.m.) or are representative of three experiments (**c**,**d**,**f**).

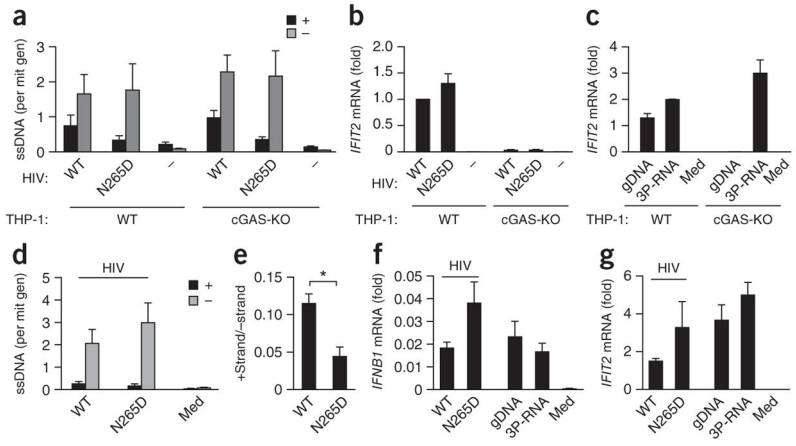

Figure 5.

The interferon response induced by HIV-1 infection correlates with the presence of (−)-strand DNA but not with the presence of (+)-strand DNA. (**a**) Strand-specific quantitative detection of the HIV-1 (+) strand and (−) strand (key) in cytosol-enriched nucleic acid preparations of wild-type (WT) or cGAS-deficient (cGAS-KO) THP-1 cells 4 h after infection with HIV-1 particles containing wild-type (HIV: WT) or mutant (HIV: N265D) reverse transcriptase or without infection (HIV: −); results are presented as ssDNA copies per mitochondrial genome (mit gen). (**b**,**c**) *IFIT2* mRNA in THP-1 cells 4 h after treatment as in **a** (**b**) or after transfection of genomic DNA or 3P-dsRNA or treatment with medium alone (Med) (**c**); results are presented as copies of *IFIT2* mRNA per copy of control *GAPDH* mRNA (**c**), or that value relative to the results of wild-type THP-1 cells infected with wild-type particles (**b**). (**d**) Strand-specific quantitative detection of the HIV-1 (+) and (−) strand (key) in cytosol-enriched nucleic acid preparations of monocyte-derived macrophages 4 h after no infection (Med) or infection with HIV-1 particles as in **a** (presented as in **a**). (**e**) Ratio of (+) strand to (−) strand in **d**. (**f**,**g**) *IFNB1* mRNA (**f**) and *IFIT2* mRNA (**g**) in monocyte-derived macrophages 4 h after stimulation with HIV-1 particles as in **a** or transfection of genomic DNA or 3P-dsRNA; results are presented as copies of *IFNB1* mRNA (**f**) or *IFIT2* mRNA (**g**) per copy of control *GAPDH* mRNA. **P* ≤ 0.01 (ratio-paired *t*-test). Data are from three (**a**,**b**) i or two (**c**) independent experiments (mean and s.e.m.) or are pooled from two experiments with two biological replicates in each (**d**–**g**; mean and s.e.m. of *n* = 4 donors).

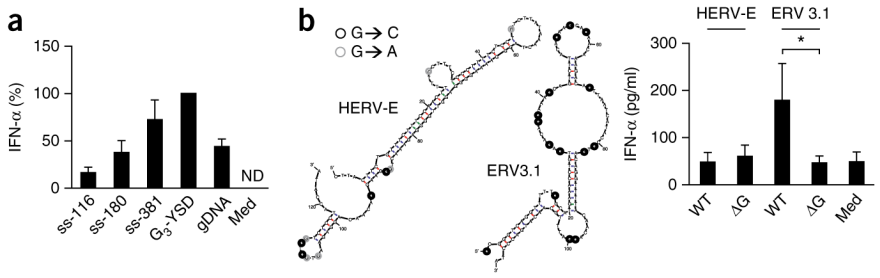

Figure 6.

Long ssDNA comprising the 5′-terminal HIV-1 (−)-strand sequence is highly immunostimulatory, and the recognition of endogenous retroelement-derived ssDNA depends on unpaired guanosines. (**a**) ELISA of IFN-α in supernatants of monocyte-derived macrophages 36 h after treatment with medium alone (Med) or transfection of genomic DNA, G_3-YSD or ssDNA species of various lengths (116 nucleotides (ss-116; this is SL2 plus SL3 as in Fig. 1c), 180 nucleotides (ss-180) or 381 nucleotides (ss-381)); results are presented relative to those of cells transfected with G_3-YSD, set as 100%. (**b**) IFN-α in supernatants of chloroquine-treated PBMCs 24 h after treatment with medium alone (Med) or transfection of wild-type (WT) or mutant (ΔG) HERV-E or ERV3.1 (right), and mFOLD-derived models of the secondary structures of HERV-E or ERV3.1 (left); circles indicate mutation of guanosine (black, G to C; gray: G to A). *$P \leq 0.05$ (one-way ANOVA followed by Tukey's post-hoc test). Data are pooled from two experiments with one or two biological replicates in each (mean and s.e.m. of $n = 3$ donors (**a**) or $n = 4$ donors (**b**)).