COMPOSITIONS AND METHODS FOR ALTERING SECOND MESSENGER SIGNALING

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present invention claims priority to United States provisional patent application nos. 61/817,269, filed April 29, 2013, and 61/819,369, filed May 3, 2013, the entire contents of each of which are hereby incorporated by reference herein.

BACKGROUND OF THE INVENTION

[0002] The importance of cyclic dinucleotides as bacterial second messengers is well established, with cyclic di-GMP (c-di-GMP) now acknowledged as a universal bacterial second messenger. This versatile molecule has been shown to play key roles in cell cycle and differentiation, motility and virulence, as well as in the regulation of biofilm formation and dispersion. Advances in our understanding of c-di-GMP has emerged with the identification, structural characterization, and mechanistic understanding of the catalytic activities of the bacterial enzymes responsible for the synthesis and degradation of this second messenger. Crystal structures of c-di-GMP in the free state and when bound to enzymes responsible for its synthesis and degradation have shown that this second messenger can adopt either monomeric or a dimeric bis-intercalated folds. It appears that formation of c-(3',5')-di-GMP from two molecules of GTP occurs via a two-step reaction and formation of 3',5'-phosphodiester linkages, with two molecules of pyrophosphate as byproducts of the cyclization reaction. Moreover, multiple receptors targeted by c-(3',5')-di-GMP and the diverse ways bacteria signal through this second messenger have been identified. Indeed, the field of c-di-GMP study as a second messenger has grown immensely and yielded major advances in our understanding of the physiology and mechanisms of bacterial cyclic dinucleotide signaling over the last two and a half decades. In parallel studies, c-(3',5')-di-GMP-specific riboswitches have also been identified, including ones that are involved in cyclic dinucleotide-induced RNA splicing.

[0003] There is much interest currently towards gaining a molecular and functional understanding of innate immunity sensors of higher metazoans that recognize nucleic acids in the cytoplasm and trigger type I interferon induction. Cytoplasmic dsDNA of pathogenic bacterial or viral origin, and perhaps also displaced nuclear or mitochondrial DNA following cellular stress, represent such a trigger. These events involving self-nucleic acid recognition in turn could trigger autoimmune diseases such as systemic lupus erythematosus and Sjögren syndrome. Indeed, in recent years many cytoplasmic DNA sensors have been identified, including DAI (DNA-dependent activator of IFN-regulatory factor), LRRFIP1 (leucine-rich repeat and flightless I interacting protein 1), DDX41 (DEAD box polypeptide 41), and members of the HIN-200 (hematopoietic interferon-inducible nuclear proteins) family such as AIM2 (absent in melanoma 2) and IFI16 (interferon-inducible protein 16). Molecular information is available on the HIN domain family as reflected in structures of their complexes with dsDNA. A requirement for multiple sensors may be a reflection of distinctive cell-type specific activities. Cytoplasmic detection of dsDNA activates stimulator of interferon genes (STING) in the cytoplasm, which in turn initiates a cascade of events by first activating kinases IKK (IκB kinase) and TBK1 (TANK-binding kinase 1), leading to phosphorylation and activation of the transcription factors NF-κB (nuclear factor κB) and IRF3 (interferon regulatory factor). These phosphorylated transcription factors translocate to the nucleus to target immune and inflammatory genes leading to the production of cytokines and type I interferons, thereby triggering the host immune response. Therefore, there is a need for therapeutic agents to modulate the induction of interferon and other relevant components in these pathways.

SUMMARY OF THE INVENTION

[0004] The present invention provides, among other things, novel cyclic-GMP-AMP (cGAMP) analogs, mimics, mimetics and variants as described in more detail below. These cGAMP compounds and compositions are, among other things, useful in the design of research tools, as a research tool, and as therapeutice modalities such as enzyme modulators including agonists and antagonists of cGAS. The present invention also provides crystallographic data for cyclic-GMP-AMP synthase (cGAS). These crystallographic data provide the basis for which to

design modulators (agonists and antagonists) such as cGAMP compounds or small molecules, which are useful in the fields of research, therapeutics and/or diagnostics.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] **Figure 1A-H**. Structures of cGAMP Synthase (cGAS) in the Free State and Bound to dsDNA. (**A**) 2.0 Å crystal structure of cGAS in the free state. The backbone of the protein in a ribbon representation is colored in light gray. (**B**) 2.1 Å crystal structure of cGAS bound to a complementary 16-bp DNA duplex (with one base 5'-overhang at each end). The protein and DNA are colored in dark gray in the binary complex. (**C**) A schematic of intermolecular hydrogen bonds in the binary cGAS-DNA complex. (**D**) Superposed structures of cGAS in the free state (light gray) and in the cGAS-DNA complex (dark gray). (**E, F**) Large changes within the β-sheet (panel E) and catalytic pocket (panel F) segments on proceeding from cGAS in the free state (light gray) to the binary complex with bound DNA (dark gray). (**G**) Narrow entrance to the catalytic pocket in the structure of cGAS in the free state, with the protein in an electrostatic representation. (**H**) Widened entrance to the catalytic pocket in the structure of the binary cGAS-DNA complex.

[0006] **Figure 2A-H**. Structure of the ternary complex of cGAS, dsDNA and ATP. (**A**) 2.4 Å crystal structure of the ternary complex of cGAS bound to dsDNA and ATP. The protein and dsDNA are shown in ribbon, with bound ATP in a space-filling representation. (**B**) Superposed structures of the binary complex of cGAS and DNA and the ternary complex with added ATP.(**C, D**) Absence of changes in the backbone within the β-sheet (panel C) and catalytic pocket (panel D) segments on proceeding from the binary cGAS and dsDNA complex to the ternary complex with added ATP. (**E, F**) Two alternate views of intermolecular contacts between ATP and catalytic pocket residues in the ternary complex. Two cations are shown as spheres, with hydrogen bonds shown by dashed lines. (**G**) 2Fo-Fc electron density map contoured at 1.2σ (light gray) and Fo-Fc map contoured at 3.0σ (dark gray) of bound ATP, pair of cations and coordinating residues in the catalytic pocket. This map contains some weak unaccounted for

electron density (dark gray). (**H**) View of bound ATP in a space-filling representation within the catalytic pocket, with the protein in an electrostatic representation.

[0007] **Figure 3A-H**. Structure of the Ternary Complex of cGAS, dsDNA with Bound Products 5'-pppG(2',5')pG and 5'-pG(2',5')pA. (**A**) 1.9 Å crystal structure of the ternary complex of cGAS bound to dsDNA and 5'-pppG(2',5')pG. The protein and DNA are shown in ribbon, with bound 5'-pppG(2',5')pG in a space-filling representation. (**B, C**) Two alternate views of intermolecular contacts between 5'-pppG(2',5')pG and catalytic pocket residues in the ternary complex. Two cations are shown as spheres, with hydrogen bonds shown by dashed lines. (**D**) 2Fo-Fc electron density map contoured at 1.2σ of bound 5'-pppG(2',5')pG in the catalytic pocket of the ternary complex.(**E**) View of bound 5'-pppG(2',5')pG in a space-filling representation within the catalytic pocket, with the protein in an electrostatic representation.(**F, G**) Two alternate views of intermolecular contacts between 5'-pG(2',5')pA and catalytic pocket residues in the 2.3 Å ternary complex of cGAS, dsDNA and GMP + ATP. (**H**) Superposition of structures of bound 5'-pppG(2',5')pG (dark gray) and 5'-pG(2',5')pA (light gray).

[0008] **Figure 4A-H**. Structure of the Ternary Complex of cGAS, DNA with Bound Product c[G(2',5')pA(3',5')p]. (**A**) 2.3 Å crystal structure of the ternary complex of cGAS bound to dsDNA and product c[G(2',5')pA(3',5')p]. The protein and DNA are shown in ribbon, with bound product c[G(2',5')pA(3',5')p] in a space-filling representation. (**B, C**) Two alternate views of intermolecular contacts between product c[G(2',5')pA(3',5')p] and catalytic pocket residues in the ternary complex. (**D**) 2Fo-Fc electron density map contoured at 1.2σ of bound c[G(2',5')pA(3',5')p] in the catalytic pocket of the ternary complex. (**E**) View of bound c[G(2',5')pA(3',5')p] in a space-filling representation positioned towards on end of the catalytic pocket, with the protein in an electrostatic representation. (**F**) A view of c[G(2',5')pA(3',5')p] highlighting the 2',5' linkage at the GpA step and the 3',5' linkage at the ApG step. (**G**) Stacking of the G residue of 5'-pG(2',5')pA on Tyr 421 in its ternary complex with cGAS and dsDNA. (**H**) Stacking of the A residue of c[G(2',5')pA(3',5')p] on Tyr 421 in its ternary complex with cGAS and dsDNA.

[0009] **Figure 5A-D**. Characterization of c[G(2',5')pA(3',5')p] Formation by cGAS. Generation of c[G(2',5')pA(3',5')p] and linear products and intermediates were assayed by thin-layer chromatography (TLC) using purified recombinant truncated (A panel, amino acids 147-

507, used in crystallization studies) and full-length cGAS (B-D panels, amino acids 1-507). Long- and short-dashed lines indicate the origin and solvent fronts, respectively. (**A**) A 45-nt dsDNA was incubated with cGAS (x-y) in reaction buffer containing indicated divalent cation (or EDTA) and α^{32}p-ATP and -GTP. Chemically synthesized cGAMP containing both 3',5' linkages was co-spotted in every sample and its migration, visualized by UV, is indicated (dashed outlines). (**B**) cGAS was incubated with single (ss) or double (ds) stranded DNA, RNA, DNA/RNA duplex, or 8-oxoguanine (8-O-G) modified DNA of similar sequence and c[G(2',5')pA(3',5')p] formation was monitored using α^{32}p-ATP. (**C**) Mono- and di-phosphorylated adenosine and guanosine were used as substrates to determine order of c[G(2',5')pA(3',5')p] formation. Slow-migrating 2',5'-linked intermediate species when cGAS and dsDNA is incubated with α^{32}p-ATP and GMP (5'-pGpA) or GDP (5'-ppGpA). (**D**) dsDNA-dependent cGAMP reaction intermediates were visualized by using 2' or 3' dATP and dGTP. Slow migrating intermediate species, corresponding to pppGpA (lane 1) or pppGpdA (lanes 2 and 3), are seen by changing TLC mobile phase composition. Intermediate species were confirmed using γ^{32}p-GTP.

[0010] **Figure 6A-C**. Definitive Identification of c[G(2',5')pA(3',5')p] as the Enzymatic Product of cGAS. (**A**) UV 260 nm chromatographs of GTP, ATP, c[G(2',5')pA(2',5')p], c[G(3',5')pA(3',5')p], c[G(2',5')pA(3',5')p] and cGAS reaction (rxn, asterisk) solutions from reverse-phase HPLC analyses. cGAS reaction samples were injected alone or with addition of indicated reference standards. Shaded region shows the retention time corresponding to the elution of c[G(2',5')pA(3',5')p]. (**B**) UV 260 nm chromatographs from HPLC analysis of the cGAS product obtained from dissolved crystals when injected alone (top trace), or co-injected with c[G(2',5')pA(2',5')p] reference compound (middle trace). Additional unidentified peaks were present in the dissolved crystal solution, but elute later. The three reference cGAMP compounds were co-injected due to a change (0.5 sec) in the retention time of c[G(2',5')pA(3',5')p] as a result of applying the dissolved crystal solution to the column. (**C**) NMR spectra of the sugar H1' proton region of three chemically synthesized cGAMP reference compounds with the cGAS rxn in 99.9% D_2O in 10 mM K_2HPO_4-KH_2PO_4 (pH 6.6) buffer. The NMR spectrum of the cGAS rxn corresponds to c[G(2',5')pA(3',5')p] reference compound. The H1' proton is a doublet ($^3J_{HH}$ = 9 Hz) when the phosphate is attached to the 2'-position, but a

singlet when the phosphate is attached to the 3'-position, reflecting the different puckers of the five-membered sugar ring dependent on the position of the attached phosphate group.

[0011] **Figure 7A-D**. Functional analysis of cGAS Mutants and the Model for Two-step Generation of c[G(2',5')pA(3',5')p]. (**A**) Levels of c[G(2',5')pA(3',5')p] formation by cGAS full-length wt and indicated mutants were compared by TLC analysis. Long- and short-dashed lines indicate the origin and solvent fronts, respectively. (**B, C**) Expression vectors of murine cGAS WT, or carrying single and multiple alanine mutations of DNA binding (panel B) and catalytic (panel C) residues were transiently transfected into HEK 293 cells together with an IFN-β Gluc reporter, and constitutive STING and Firefly luc expression plasmids. In this setting expressed cGAS is engaged in the cytosol by the co-transfected DNA plasmids. Gluc values were determined in triplicate, 36 h after transfection, normalized to Firefly luc, and are shown as fold induction over control plasmid (as mean ± s.e.m). Data in panels B and C are representative of 3 - 5 independent experiments for each mutant. (**D**) A schematic representation of a proposed model associated with a two-step generation of c[G(2',5')pA(3',5')p] within the single catalytic pocket of cGAS. In this model, the first step involves formation of a 5'-pppGpA intermediate followed by formation of c[G(2',5')pA(3',5')p]. Note, also that the bound ligand is believed to undergo two flip-overs on the pathway to c[G(2',5')pA(3',5')p] formation.

[0012] **Figure S1A-C**. Sequence Alignment and Crystal Structure of cGAS in the Free State and Comparison with Human OAS1. (**A**) Sequence alignment of cGAS from mouse and human (construct used for structural studies) spanning amino acids 147 to 507 (C-terminus). The putative catalytic residues are indicated in boxes. (**B**) Two alternate views of the structure of cGAS in the free state. The backbone of the protein is shown in a ribbon representation and colored in light gray. (**C**) Stereo view of superposed structures of cGAS (light gray) and human oligoadenylate synthetase 1 (OAS1) (black; PDB: 1PX5) in the free state. The r.m.s.d between structures is 4.1 Å.

[0013] **Figure S2A-F**. Molecular Recognition Features in the Structure of cGAS Bound to dsDNA and Comparison with hOAS1 Bound to dsRNA and 2'-Datp. (**A, B**) Examples of intermolecular contacts between cGAS and dsDNA. Water molecules are shown as black spheres, with hydrogen bonds are indicated by dashed lines. We observe one sequence-specific hydrogen bond between the side chain of Arg161 and the O2 carbonyl of T8 as shown in panel

B. (**C, D, E**) Examples of conformational shifts on proceeding from cGAS in the free state (light gray) to the binary complex with bound dsDNA (gray). A shift of 5.1 Å is observed in the β-sheet segment on complex formation (panel C). A long α-helix breaks into two segments, with one segment moving towards the dsDNA on complex formation, including the side chain of Arg161, which moves by 9.2 Å (panel D). Several Tyr and Lys residues within loop segments shift between 6.7 and 17.6 Å on complex formation (panel E). (**F**) Stereo view of the superposed structures of the protein components of cGAS in the dsDNA bound state (light gray) and OAS1 in the dsRNA bound state plus 2'-dATP (black, PDB: 4IG8). The r.m.s.d between structures is 3.2 Å. The dsDNA bound to cGAS and dsRNA bound to OAS1 are omitted from depiction for clarity.

[0014] **Figure S3A-I**. Structure of cGAS with 5'-pppG(2',5')pG in the Catalytic Pocket of its Ternary Complex Formed upon Crystallization with GTP. (**A**) Superposed structures of the binary complex of cGAS with DNA (gray) and the ternary complex with bound 5'-pppG(2',5')pG intermediate product (dark gray). (**B, C**) Minimal changes are observed in the backbone within the β-sheet (panel B) and catalytic pocket (panel C) segments on proceeding from the binary complex to the ternary complex with bound 5'-pppG(2',5')pG. (**D**) Two alternate views of the bound 5'-pppG(2',5')pG in the catalytic pocket of the ternary complex. Mg^{2+} cations are shown as spheres. Note that the alignment of bound ligand is 5'-pppG(*syn*)p(2',5')pG(*anti*). (**E**) Two alternate views of the omit Fo-Fc omit electron density map contoured at 3.0σ of bound 5'-pppG(2',5')pG in the catalytic pocket of the ternary complex. (**F**) Two alternate views of the superposed structures of bound 5'-pppG(2',5')pG (gray) and ATP (dark gray) in their respective ternary complexes with cGAS and dsDNA. (**G**) Omit map recorded at 4σ identifying two bound cations in the structure of the ternary complex. (**H, I**) Octahedral coordination geometry around the two bound cations in the structure of the ternary complex.

[0015] **Figure S4A-F**. Structure of cGAS and c[G(2',5')pA(3',5')p] Bound in the Catalytic Pocket of the Ternary Complex Formed upon Crystallization with GTP + ATP. (**A**) Superposed structures of the binary complex of cGAS and DNA (gray) and the ternary complex with added GTP+ATP for which the bound product is c[G(2',5')pA(3',5')p] (dark gray) obtained from crystallization with ATP and GTP. (**B, C**) No conformational changes occurred in the

backbone within the β-sheet (panel B) and catalytic pocket (panel C) segments on proceeding from the binary complex to the ternary complex with bound c[G(2',5')pA(3',5')p]. (**D**) Two alternate views of the bound product cGAMP in the catalytic pocket of the ternary complex. Note that the bound ligand c[G(2',5')pA(3',5')p] revealed a 2',5' phosphodiester linkage within the GpA step. Based on HPLC comparison, the structure of c[G(2',5')pA(3',5')p] is shown with a 3',5' linkage at the ApG step. Both G and A residues adopt *anti* alignments at their glycosidic bonds. (**E**) Two alternate views of the Fo-Fc omit electron density map contoured at 3.0σ of bound c[G(2',5')pA(3',5')p] in the catalytic pocket of the ternary complex. (**F**) Two alternate views of the superposed structures of bound c[G(2',5')pA(3',5')p] (gray) and ATP (dark gray) in their respective ternary complexes with cGAS and dsDNA.

[0016] **Figure S5A-C**. Thin-layer Chromatography (TLC) Conditions for Monitoring Formation of Cgamp. (**A, B**) Indicated nucleotides were spotted on high-performance silica gel TLC plates, resolved by various solvent systems, and visualized by UV. Two mobile phase conditions were used (A and B). Solvent system 1 was used in the majority of experiments for detection of c[G(2',5')pA(3',5')p], whereas solvent 2 was used for a better separation of the mono and tri-phosphorylated linear intermediates. Dashed lines indicate the solvent fronts. (**C**) Calculated Rf values.

[0017] **Figure S6A-E**. dsDNA-length and Nucleotide Requirements of cGAS Activity. (**A**) Full-length cGAS was incubated with equimolar or mass-normalized quantities of 16-, 36-, or 45-nt dsDNA then assayed for cGAMP formation. Long- and short-dashed lines in panels A-C, indicate the origin and solvent fronts, respectively. (**B**) Truncated (tr) and full-length (fl) cGAS was incubated with 45 bp dsDNA in reaction buffer containing the various indicated nucleotides. cGAS (tr) exhibits activity, albeit less than cGAS (fl). c[G(2',5')pA(3',5')p] forms using 2'-dATP, when 2'-dATP or GTP was radiolabelled, but not at all when 2'-dGTP was used. 2'-dATP with 2'-dGTP yielded no c[G(2',5')pA(3',5')p], indicating that blockage of the 2' OH positions in adenosine, and more importantly guanosine, prevented c[G(2',5')pA(3',5')p] production. Asterisks (*) denote which nucleotides were supplemented with an α^{32}p-radiolabelled form. dNTP indicates the triphosphorylated 2'-deoxynucleotide. (**C**) Full-length cGAS was incubated in reaction buffer containing dsDNA and the indicated combination of ribonucleotides, then analyzed by TLC. Trace amounts of cyclic product were formed upon

incubation of α^{32}p-ATP with UTP, or α^{32}p-GTP with GTP, CTP, and UTP. Optimal product formation requires GTP and ATP. The low level of cyclic product formation with UTP and ATP, but no ATP alone, suggests that UTP can be accommodated at the GTP binding site but reduced in affinity and/or activity. The migrations of all products are consistent with formation of cyclical dinucleotides. (**D**) HPLC analysis of dsDNA-dependent cGAS generation of c[G(2',5')pA(3',5')p] over time. A single cGAS reaction was initiated and samples were analyzed by HPLC at indicated times. (**E**) Highly conserved residues G198 and S199 were mutated to alanine, or G198 to proline to reduce steric flexibility. Expression plasmids for mutant and WT cGAS were transiently transfected into HEK 293 cells together with an IFN-β Gluc reporter, and constitutive STING and Firefly luc expression plasmids. Gluc values were determined in triplicate, 36 h after transfection, normalized to Firefly luc, and are shown as fold induction over Control plasmid (as mean ±s.e.m). Data are representative of 3 independent experiments for each mutant.

[0018] **Figure S7**. Syntheses of cGAMP isomers. Synthesis of cGAMP containing 2',5' linkages at both GpA and ApG steps (6) (top panel). Synthesis of cGAMP containing 2',5' at GpA step and 3',5' at ApG step (11) (middle panel). Synthesis of cGAMP containing 3',5' linkages at both GpA and ApG steps (15) (bottom panel).

[0019] **Figure S8A-D**. Resonance assignments of c[G(2',5')pA(3',5')p] from HMBC, COSY, and HSQC two-dimensional NMR spectra. (**A**) HMBC spectrum showing correlations between aromatics and the sugar C1'-H1'. (**B**) HMBC spectrum showing correlations within sugar rings. In (A) and (B), correlations within the guanine base are connected by solid lines and assignments are specified on the upper and left edges for protons and carbons respectively, while correlations within the adenine base are connected by dashed lines and assignments are specified on the lower and right edges for protons and carbons respectively; The large unsuppressed 1-bond C-H couplings are indicated by blue lines connecting the coupled pairs of signals. (**C**) Double quantum filtered COSY spectrum. Guanine correlations are connected by solid lines and resonances are labeled on the cross peaks above the diagonal; Adenine correlations are connected by dashed lines and resonances are labeled on the cross peaks below the diagonal. (**D**) Aliphatic HSQC spectrum summarizing sugar proton and carbon assignments. See also Figure 6C and Table S4.

[0020] **Figure S9.** STING-dependent induction of murine alpha-interferon and human CXCL10 by cGAMP compounds. The dose-dependent biological activities of indicated cGAMP isomers were measured by enzyme-linked immunosorbent assay (ELISA), quantifying for the induction of endogenous murine α-interferon (m-Ifna) or human CXCL10 (h-CXCL10) proteins. (A-B), Mouse bone marrow derived macrophages (BMDM) were either treated first with Digitonin (Dig) to permeabilize plasma membranes prior to cGAMP addition (A) or cGAMP isomers were passively delivered to cells by addition in culture medium (B). (C), cGAMP activation was also measured in human THP-1 cells. Data are representative of 2 independent experiments, each done in triplicate (error bars, s.e.m.). (D), The half maximal effective concentration (EC_{50}) values were estimated based on 4-parametric sigmoidal dose-response curves; 95% confidence interval ranges (CI) are provided.

[0021] **Figure S10** is an exemplary block diagram of a computing device and a mobile computing device.

[0022] **Figure S11** is an exemplary block diagram of a network environment for establishing a multi-channel context aware communication environment.

DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS

[0023] It has been shown that VSP-1 (*Vibrio* 7th pandemic island-1) genes encode a novel class of dinucleotide cyclases, which preferentially synthesize a cyclic-GMP-AMP (designated cGAMP) molecule, thereby expanding our horizon to cyclic GA-dinucleotides (Davies et al. 2012). More recently, cyclic GMP-AMP synthase (cGAS, official human gene symbol MB21D1) was identified as a cytoplasmic DNA sensor that activated the type I interferon pathway by synthesizing the second messenger cGAMP (Sun et al. 2013; Wu et al. 2013). cGAS was shown to be a member of the nucleotidyltransferase family, and to be capable of generating a cGAMP *in vitro* from GTP and ATP in the presence of dsDNA (but not dsRNA), while chemically synthesized cGAMP containing a pair of 3',5' linkages was shown to stimulate the production of interferon in THP1 and Raw264.7 cells at concentrations as low as 10 nM. The authors also demonstrated through experiments involving either overexpression or knockdown of

cGAS, that the synthetic cGAMP bound to and activated STING, resulting in the activation of transcription factor IRF3 and subsequent induction of interferon (Sun et al. 2013; Wu et al. 2013).

[0024] A critical assumption in these studies on cGAS was that cGAMP contained a pair of 3',5' linkages (Sun et al. 2013; Wu et al. 2013), in line with those reported previously for c-di-GMP in bacterial systems as outlined above. The present invention encompasses the recognition that the previously assigned structure of cGAMP by Sun and Wu was incorrect. Thus, one aspect of the present invention is the identification of the previously unknown problem of misidentification of the structure of cGAMP. The present disclosure combines structural, chemical, *in vitro* biochemical and *in vivo* cellular assays to establish unequivocally that this second messenger unexpectedly contains 2',5' linkage at the GpA step and 3',5' linkage at the ApG step {designated c[G(2',5')pA(3',5')p]}, thus identifying correctly and for the first time, the founding member of a new family of metazoan second messengers regulating type I interferon induction in response to cytoplasmic DNA.

[0025] In certain embodiments, the present invention provides compounds comprising cyclic GA-dinucleotides [c[G(2',5')pA(3',5')p]] containing a 2',5' linkage (at the GpA step). In some embodiments, such compounds are useful for the study of cellular signaling and immune surveillance in metazoans. In some embodiments, such compounds are useful in the treatment, diagnosis or prophylaxis of disorders, diseases or conditions in medicine. In some embodiments, such compounds are useful to modulate targets involved in immune response. In some embodiments, the compounds and/or compositions of the invention are useful as research tools and/or reagents, particuluarly in kits and assays for biological or chemical research.

[0026] The present invention also provides crystallographic data useful in the design of modulators of cGAS. In certain embodiments, the invention provides modulators of cGAS that comprises features to form appropriate binding interactions with cGAS. In some embodiments, such modulators comprise features that form appropriate binding interactions with targets that bind to cGAMP.

Definitions

[0027] Compounds of this invention include those described generally above, and are further illustrated by the classes, subclasses, and species disclosed herein. As used herein, the following definitions shall apply unless otherwise indicated. For purposes of this invention, the chemical elements are identified in accordance with the Periodic Table of the Elements, CAS version, Handbook of Chemistry and Physics, 75th Ed. Additionally, general principles of organic chemistry are described in "Organic Chemistry", Thomas Sorrell, University Science Books, Sausalito: 1999, and "March's Advanced Organic Chemistry", 5th Ed., Ed.: Smith, M.B. and March, J., John Wiley & Sons, New York: 2001, the entire contents of which are hereby incorporated by reference.

[0028] The abbreviations used herein have their conventional meaning within the chemical and biological arts. The chemical structures and formulae set forth herein are constructed according to the standard rules of chemical valency known in the chemical arts.

[0029] Unless otherwise stated, structures depicted herein are also meant to include all isomeric (e.g., enantiomeric, diastereomeric, and geometric (or conformational)) forms of the structure; for example, the R and S configurations for each asymmetric center, Z and E double bond isomers, and Z and E conformational isomers. Therefore, single stereochemical isomers as well as enantiomeric, diastereomeric, and geometric (or conformational) mixtures of the present compounds are within the scope of the invention. Unless otherwise stated, all tautomeric forms of the compounds of the invention are within the scope of the invention. Additionally, unless otherwise stated, structures depicted herein are also meant to include compounds that differ only in the presence of one or more isotopically enriched atoms. For example, compounds having the present structures including the replacement of hydrogen by deuterium or tritium, or the replacement of a carbon by a ^{13}C- or ^{14}C-enriched carbon are within the scope of this invention. Such compounds are useful, for example, as analytical tools, as probes in biological assays, or as therapeutic agents in accordance with the present invention.

[0030] Provided compounds may comprise one or more saccharide moieties. Unless otherwise specified, both D- and L-configurations, and mixtures thereof, are within the scope of

the disclosure. Unless otherwise specified, both α- and β-linked embodiments, and mixtures thereof, are contemplated by the present disclosure.

[0031] If, for instance, a particular enantiomer of a compound of the present disclosure is desired, it may be prepared by asymmetric synthesis, chiral chromatography, or by derivation with a chiral auxiliary, where the resulting diastereomeric mixture is separated and the auxiliary group cleaved to provide the pure desired enantiomers. Alternatively, where the molecule contains a basic functional group, such as amino, or an acidic functional group, such as carboxyl, diastereomeric salts are formed with an appropriate optically-active acid or base, followed by resolution of the diastereomers thus formed by fractional crystallization or chromatographic means well known in the art, and subsequent recovery of the pure enantiomers.

[0032] The term "acyl," as used herein, represents a hydrogen or an alkyl group (e.g., a haloalkyl group), as defined herein, that is attached to the parent molecular group through a carbonyl group, as defined herein, and is exemplified by formyl (i.e., a carboxyaldehyde group), acetyl, propionyl, butanoyl and the like. Exemplary unsubstituted acyl groups include from 1 to 7, from 1 to 11, or from 1 to 21 carbons. In some embodiments, the alkyl group is further substituted with 1, 2, 3, or 4 substituents as described herein.

[0033] The term "aliphatic" or "aliphatic group", as used herein, means a straight-chain (i.e., unbranched) or branched, substituted or unsubstituted hydrocarbon chain that is completely saturated or that contains one or more units of unsaturation, or a monocyclic hydrocarbon or bicyclic hydrocarbon that is completely saturated or that contains one or more units of unsaturation, but which is not aromatic (also referred to herein as "carbocycle," "cycloaliphatic" or "cycloalkyl"), that has a single point of attachment to the rest of the molecule. Unless otherwise specified, aliphatic groups contain 1-6 aliphatic carbon atoms. In some embodiments, aliphatic groups contain 1-5 aliphatic carbon atoms. In some embodiments, aliphatic groups contain 1-4 aliphatic carbon atoms. In some embodiments, aliphatic groups contain 1-3 aliphatic carbon atoms, and in yet other embodiments, aliphatic groups contain 1-2 aliphatic carbon atoms. In some embodiments, "cycloaliphatic" (or "carbocycle" or "cycloalkyl") refers to a monocyclic C_3-C_6 hydrocarbon that is completely saturated or that contains one or more units of unsaturation, but which is not aromatic, that has a single point of attachment to the rest of the molecule. Suitable aliphatic groups include, but are not limited to, linear or branched,

substituted or unsubstituted alkyl, alkenyl, alkynyl groups and hybrids thereof such as (cycloalkyl)alkyl, (cycloalkenyl)alkyl or (cycloalkyl)alkenyl.

[0034] The term "heteroatom" means one or more of oxygen, sulfur, nitrogen, phosphorus, or silicon (including, any oxidized form of nitrogen, sulfur, phosphorus, or silicon; the quaternized form of any basic nitrogen or; a substitutable nitrogen of a heterocyclic ring, for example N (as in 3,4-dihydro-2*H*-pyrrolyl), NH (as in pyrrolidinyl) or NR^+ (as in N-substituted pyrrolidinyl)).

[0035] The term "unsaturated," as used herein, means that a moiety has one or more units of unsaturation.

[0036] The term "alkyl," as used herein, refers to saturated, straight– or branched–chain hydrocarbon radicals derived from an aliphatic moiety containing between one and six carbon atoms by removal of a single hydrogen atom. Unless otherwise specified, alkyl groups contain 1–12 carbon atoms. In certain embodiments, alkyl groups contain 1–8 carbon atoms. In certain embodiments, alkyl groups contain 1–6 carbon atoms. In some embodiments, alkyl groups contain 1–5 carbon atoms, in some embodiments, alkyl groups contain 1–4 carbon atoms, in some embodiments alkyl groups contain 1–3 carbon atoms, and in some embodiments alkyl groups contain 1–2 carbon atoms. Examples of alkyl radicals include, but are not limited to, methyl, ethyl, n–propyl, isopropyl, n–butyl, iso–butyl, sec–butyl, sec–pentyl, iso–pentyl, tert–butyl, n–pentyl, neopentyl, n–hexyl, sec–hexyl, n–heptyl, n–octyl, n–decyl, n–undecyl, dodecyl, and the like.

[0037] The term "alkenyl," as used herein, denotes a monovalent group derived from a straight– or branched–chain aliphatic moiety having at least one carbon–carbon double bond by the removal of a single hydrogen atom. Unless otherwise specified, alkenyl groups contain 2–12 carbon atoms. In certain embodiments, alkenyl groups contain 2–8 carbon atoms. In certain embodiments, alkenyl groups contain 2–6 carbon atoms. In some embodiments, alkenyl groups contain 2–5 carbon atoms, in some embodiments, alkenyl groups contain 2–4 carbon atoms, in some embodiments alkenyl groups contain 2–3 carbon atoms, and in some embodiments alkenyl groups contain 2 carbon atoms. Alkenyl groups include, for example, ethenyl, propenyl, butenyl, 1–methyl–2–buten–1–yl, and the like.

[0038] The term "alkynyl," as used herein, refers to a monovalent group derived from a straight– or branched–chain aliphatic moiety having at least one carbon–carbon triple bond by the removal of a single hydrogen atom. Unless otherwise specified, alkynyl groups contain 2–12 carbon atoms. In certain embodiments, alkynyl groups contain 2–8 carbon atoms. In certain embodiments, alkynyl groups contain 2–6 carbon atoms. In some embodiments, alkynyl groups contain 2–5 carbon atoms, in some embodiments, alkynyl groups contain 2–4 carbon atoms, in some embodiments alkynyl groups contain 2–3 carbon atoms, and in some embodiments alkynyl groups contain 2 carbon atoms. Representative alkynyl groups include, but are not limited to, ethynyl, 2–propynyl (propargyl), 1–propynyl, and the like.

[0039] The term "alkylene" refers to a bivalent alkyl group. An "alkylene chain" is a polymethylene group, i.e., $–(CH_2)_n–$, wherein n is a positive integer, preferably from 1 to 6, from 1 to 4, from 1 to 3, from 1 to 2, or from 2 to 3. A substituted alkylene chain is a polymethylene group in which one or more methylene hydrogen atoms are replaced with a substituent. Suitable substituents include those described below for a substituted aliphatic group.

[0040] The term "alkenylene" refers to a bivalent alkenyl group. A substituted alkenylene chain is a polymethylene group containing at least one double bond in which one or more hydrogen atoms are replaced with a substituent. Suitable substituents include those described below.

[0041] The term "halo," as used herein, represents a halogen selected from bromine, chlorine, iodine, or fluorine

[0042] The term "halogen" means F, Cl, Br, or I.

[0043] The term "haloalkoxy," as used herein, represents an alkoxy group, as defined herein, substituted by a halogen group (i.e., F, Cl, Br, or I). A haloalkoxy may be substituted with one, two, three, or, in the case of alkyl groups of two carbons or more, four halogens. Haloalkoxy groups include perfluoroalkoxys (e.g., $-OCF_3$), $-OCHF_2$, $-OCH_2F$, $-OCCl_3$, $-OCH_2CH_2Br$, $-OCH_2CH(CH_2CH_2Br)CH_3$, and $-OCHICH_3$. In some embodiments, the haloalkoxy group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for alkyl groups.

[0044] The term "haloalkyl," as used herein, represents an alkyl group, as defined herein, substituted by a halogen group (i.e., F, Cl, Br, or I). A haloalkyl may be substituted with one, two, three, or, in the case of alkyl groups of two carbons or more, four halogens. Haloalkyl groups include perfluoroalkyls (e.g., $-CF_3$), $-CHF_2$, $-CH_2F$, $-CCl_3$, $-CH_2CH_2Br$, $-CH_2CH(CH_2CH_2Br)CH_3$, and $-CHICH_3$. In some embodiments, the haloalkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for alkyl groups.

[0045] The term "aryl" used alone or as part of a larger moiety as in "aralkyl," "aralkoxy," or "aryloxyalkyl," refers to monocyclic and bicyclic ring systems having a total of five to 10 ring members, wherein at least one ring in the system is aromatic and wherein each ring in the system contains three to seven ring members. The term "aryl" may be used interchangeably with the term "aryl ring". In some embodiments, an 8-10 membered bicyclic aryl group is an optionally substituted naphthyl ring. In certain embodiments of the present invention, "aryl" refers to an aromatic ring system which includes, but not limited to, phenyl, biphenyl, naphthyl, anthracyl and the like, which may bear one or more substituents. Also included within the scope of the term "aryl," as it is used herein, is a group in which an aromatic ring is fused to one or more non–aromatic rings, such as indanyl, phthalimidyl, naphthimidyl, phenanthridinyl, or tetrahydronaphthyl, and the like.

[0046] The terms "heteroaryl" and "heteroar–," used alone or as part of a larger moiety, e.g., "heteroaralkyl," or "heteroaralkoxy," refer to groups having 5 to 10 ring atoms, preferably 5, 6, or 9 ring atoms; having 6, 10, or 14 π electrons shared in a cyclic array; and having, in addition to carbon atoms, from one to five heteroatoms. Heteroaryl groups include, without limitation, thienyl, furanyl, pyrrolyl, imidazolyl, pyrazolyl, triazolyl, tetrazolyl, oxazolyl, isoxazolyl, oxadiazolyl, thiazolyl, isothiazolyl, thiadiazolyl, pyridyl, pyridazinyl, pyrimidinyl, pyrazinyl, indolizinyl, purinyl, naphthyridinyl, and pteridinyl. The terms "heteroaryl" and "heteroar–", as used herein, also include groups in which a heteroaromatic ring is fused to one or more aryl, cycloaliphatic, or heterocyclyl rings, where the radical or point of attachment is on the heteroaromatic ring. Nonlimiting examples include indolyl, isoindolyl, benzothienyl, benzofuranyl, dibenzofuranyl, indazolyl, benzimidazolyl, benzthiazolyl, quinolyl, isoquinolyl, cinnolinyl, phthalazinyl, quinazolinyl, quinoxalinyl, 4*H*–quinolizinyl, carbazolyl, acridinyl, phenazinyl, phenothiazinyl, phenoxazinyl, tetrahydroquinolinyl, tetrahydroisoquinolinyl, and

pyrido[2,3–b]–1,4–oxazin–3(4H)–one. A heteroaryl group may be mono– or bicyclic. The term "heteroaryl" may be used interchangeably with the terms "heteroaryl ring," "heteroaryl group," or "heteroaromatic," any of which terms include rings that are optionally substituted. The term "heteroaralkyl" refers to an alkyl group substituted by a heteroaryl, wherein the alkyl and heteroaryl portions independently are optionally substituted.

[0047] As used herein, the terms "heterocycle," "heterocyclyl," "heterocyclic radical," and "heterocyclic ring" are used interchangeably and refer to a 5-, 6- or 7-membered ring, unless otherwise specified, containing one, two, three, or four heteroatoms independently selected from the group consisting of nitrogen, oxygen, and sulfur. The 5-membered ring has zero to two double bonds, and the 6- and 7-membered rings have zero to three double bonds. Exemplary unsubstituted heterocyclyl groups are of 1 to 12 (e.g., 1 to 11, 1 to 10, 1 to 9, 2 to 12, 2 to 11, 2 to 10, or 2 to 9) carbons. The term "heterocyclyl" also represents a heterocyclic compound having a bridged multicyclic structure in which one or more carbons and/or heteroatoms bridges two non-adjacent members of a monocyclic ring, e.g., a quinuclidinyl group. The term "heterocyclyl" includes bicyclic, tricyclic, and tetracyclic groups in which any of the above heterocyclic rings is fused to one, two, or three carbocyclic rings, e.g., an aryl ring, a cyclohexane ring, a cyclohexene ring, a cyclopentane ring, a cyclopentene ring, or another monocyclic heterocyclic ring, such as indolyl, quinolyl, isoquinolyl, tetrahydroquinolyl, benzofuryl, benzothienyl and the like. Examples of fused heterocyclyls include tropanes and 1,2,3,5,8,8a-hexahydroindolizine. Heterocyclics include pyrrolyl, pyrrolinyl, pyrrolidinyl, pyrazolyl, pyrazolinyl, pyrazolidinyl, imidazolyl, imidazolinyl, imidazolidinyl, pyridyl, piperidinyl, homopiperidinyl, pyrazinyl, piperazinyl, pyrimidinyl, pyridazinyl, oxazolyl, oxazolidinyl, isoxazolyl, isoxazolidiniyl, morpholinyl, thiomorpholinyl, thiazolyl, thiazolidinyl, isothiazolyl, isothiazolidinyl, indolyl, indazolyl, quinolyl, isoquinolyl, quinoxalinyl, dihydroquinoxalinyl, quinazolinyl, cinnolinyl, phthalazinyl, benzimidazolyl, benzothiazolyl, benzoxazolyl, benzothiadiazolyl, furyl, thienyl, thiazolidinyl, isothiazolyl, triazolyl, tetrazolyl, oxadiazolyl (e.g., 1,2,3-oxadiazolyl), purinyl, thiadiazolyl (e.g., 1,2,3-thiadiazolyl), tetrahydrofuranyl, dihydrofuranyl, tetrahydrothienyl, dihydrothienyl, dihydroindolyl, dihydroquinolyl, tetrahydroquinolyl, tetrahydroisoquinolyl, dihydroisoquinolyl, pyranyl, dihydropyranyl, dithiazolyl, benzofuranyl, isobenzofuranyl, benzothienyl, and the like, including

dihydro and tetrahydro forms thereof, where one or more double bonds are reduced and replaced with hydrogens. Still other exemplary heterocyclyls include: 2,3,4,5-tetrahydro-2-oxo-oxazolyl; 2,3-dihydro-2-oxo-1H-imidazolyl; 2,3,4,5-tetrahydro-5-oxo-1H-pyrazolyl (e.g., 2,3,4,5-tetrahydro-2-phenyl-5-oxo-1H-pyrazolyl); 2,3,4,5-tetrahydro-2,4-dioxo-1H-imidazolyl (e.g., 2,3,4,5-tetrahydro-2,4-dioxo-5-methyl-5-phenyl-1H-imidazolyl); 2,3-dihydro-2-thioxo-1,3,4-oxadiazolyl (e.g., 2,3-dihydro-2-thioxo-5-phenyl-1,3,4-oxadiazolyl); 4,5-dihydro-5-oxo-1*H*-triazolyl (e.g., 4,5-dihydro-3-methyl-4-amino 5-oxo-1*H*-triazolyl); 1,2,3,4-tetrahydro-2,4-dioxopyridinyl (e.g., 1,2,3,4-tetrahydro-2,4-dioxo-3,3-diethylpyridinyl); 2,6-dioxo-piperidinyl (e.g., 2,6-dioxo-3-ethyl-3-phenylpiperidinyl); 1,6-dihydro-6-oxopyridiminyl; 1,6-dihydro-4-oxopyrimidinyl (e.g., 2-(methylthio)-1,6-dihydro-4-oxo-5-methylpyrimidin-1-yl); 1,2,3,4-tetrahydro-2,4-dioxopyrimidinyl (e.g., 1,2,3,4-tetrahydro-2,4-dioxo-3-ethylpyrimidinyl); 1,6-dihydro-6-oxo-pyridazinyl (e.g., 1,6-dihydro-6-oxo-3-ethylpyridazinyl); 1,6-dihydro-6-oxo-1,2,4-triazinyl (e.g., 1,6-dihydro-5-isopropyl-6-oxo-1,2,4-triazinyl); 2,3-dihydro-2-oxo-1*H*-indolyl (e.g., 3,3-dimethyl-2,3-dihydro-2-oxo-1*H*-indolyl and 2,3-dihydro-2-oxo-3,3′-spiropropane-1*H*-indol-1-yl); 1,3-dihydro-1-oxo-2*H*-iso-indolyl; 1,3-dihydro-1,3-dioxo-2*H*-iso-indolyl; 1*H*-benzopyrazolyl (e.g., 1-(ethoxycarbonyl)- 1*H*-benzopyrazolyl); 2,3-dihydro-2-oxo-1*H*-benzimidazolyl (e.g., 3-ethyl-2,3-dihydro-2-oxo-1*H*-benzimidazolyl); 2,3-dihydro-2-oxo-benzoxazolyl (e.g., 5-chloro-2,3-dihydro-2-oxo-benzoxazolyl); 2,3-dihydro-2-oxo-benzoxazolyl; 2-oxo-2H-benzopyranyl; 1,4-benzodioxanyl; 1,3-benzodioxanyl; 2,3-dihydro-3-oxo,4*H*-1,3-benzothiazinyl; 3,4-dihydro-4-oxo-3*H*-quinazolinyl (e.g., 2-methyl-3,4-dihydro-4-oxo-3*H*-quinazolinyl); 1,2,3,4-tetrahydro-2,4-dioxo-3*H*-quinazolyl (e.g., 1-ethyl-1,2,3,4-tetrahydro-2,4-dioxo-3*H*-quinazolyl); 1,2,3,6-tetrahydro-2,6-dioxo-7*H*-purinyl (e.g., 1,2,3,6-tetrahydro-1,3-dimethyl-2,6-dioxo-7 *H* -purinyl); 1,2,3,6-tetrahydro-2,6-dioxo-1 *H* –purinyl (e.g., 1,2,3,6-tetrahydro-3,7-dimethyl-2,6-dioxo-1 *H* -purinyl); 2-oxobenz[*c,d*]indolyl; 1,1-dioxo-2H-naphth[1,8-*c,d*]isothiazolyl; and 1,8-naphthylenedicarboxamido. Additional heterocyclics include 3,3a,4,5,6,6a-hexahydro-pyrrolo[3,4-b]pyrrol-(2H)-yl, and 2,5-diazabicyclo[2.2.1]heptan-2-yl, homopiperazinyl (or diazepanyl), tetrahydropyranyl, dithiazolyl, benzofuranyl, benzothienyl, oxepanyl, thiepanyl, azocanyl, oxecanyl, and thiocanyl.

[0048] As used herein, the term "partially unsaturated" refers to a ring moiety that includes at least one double or triple bond. The term "partially unsaturated" is intended to

encompass rings having multiple sites of unsaturation, but is not intended to include aryl or heteroaryl moieties, as herein defined.

[0049] As described herein, compounds of the invention may, when specified, contain "optionally substituted" moieties. In general, the term "substituted," whether preceded by the term "optionally" or not, means that one or more hydrogens of the designated moiety are replaced with a suitable substituent. Unless otherwise indicated, an "optionally substituted" group may have a suitable substituent at each substitutable position of the group, and when more than one position in any given structure may be substituted with more than one substituent selected from a specified group, the substituent may be either the same or different at every position. Combinations of substituents envisioned by this invention are preferably those that result in the formation of stable or chemically feasible compounds. The term "stable," as used herein, refers to compounds that are not substantially altered when subjected to conditions to allow for their production, detection, and, in certain embodiments, their recovery, purification, and use for one or more of the purposes disclosed herein.

[0050] Suitable monovalent substituents on a substitutable carbon atom of an "optionally substituted" group are independently halogen; $-(CH_2)_{0-4}R^\circ$; $-(CH_2)_{0-4}OR^\circ$; $-O(CH_2)_{0-4}R^\circ$, $-O-(CH_2)_{0-4}C(O)OR^\circ$; $-(CH_2)_{0-4}CH(OR^\circ)_2$; $-(CH_2)_{0-4}SR^\circ$; $-(CH_2)_{0-4}Ph$, which may be substituted with R°; $-(CH_2)_{0-4}O(CH_2)_{0-1}Ph$ which may be substituted with R°; $-CH=CHPh$, which may be substituted with R°; $-(CH_2)_{0-4}O(CH_2)_{0-1}$-pyridyl which may be substituted with R°; $-NO_2$; $-CN$; $-N_3$; $-(CH_2)_{0-4}N(R^\circ)_2$; $-(CH_2)_{0-4}N(R^\circ)C(O)R^\circ$; $-N(R^\circ)C(S)R^\circ$; $-(CH_2)_{0-4}N(R^\circ)C(O)NR^\circ_2$; $-N(R^\circ)C(S)NR^\circ_2$; $-(CH_2)_{0-4}N(R^\circ)C(O)OR^\circ$; $-N(R^\circ)N(R^\circ)C(O)R^\circ$; $-N(R^\circ)N(R^\circ)C(O)NR^\circ_2$; $-N(R^\circ)N(R^\circ)C(O)OR^\circ$; $-(CH_2)_{0-4}C(O)R^\circ$; $-C(S)R^\circ$; $-(CH_2)_{0-4}C(O)OR^\circ$; $-(CH_2)_{0-4}C(O)SR^\circ$; $-(CH_2)_{0-4}C(O)OSiR^\circ_3$; $-(CH_2)_{0-4}OC(O)R^\circ$; $-OC(O)(CH_2)_{0-4}SR-$, $SC(S)SR^\circ$; $-(CH_2)_{0-4}SC(O)R^\circ$; $-(CH_2)_{0-4}C(O)NR^\circ_2$; $-C(S)NR^\circ_2$; $-C(S)SR^\circ$; $-SC(S)SR^\circ$, $-(CH_2)_{0-4}OC(O)NR^\circ_2$; $-C(O)N(OR^\circ)R^\circ$; $-C(O)C(O)R^\circ$; $-C(O)CH_2C(O)R^\circ$; $-C(NOR^\circ)R^\circ$; $-(CH_2)_{0-4}SSR^\circ$; $-(CH_2)_{0-4}S(O)_2R^\circ$; $-(CH_2)_{0-4}S(O)_2OR^\circ$; $-(CH_2)_{0-4}OS(O)_2R^\circ$; $-S(O)_2NR^\circ_2$; $-(CH_2)_{0-4}S(O)R^\circ$; $-N(R^\circ)S(O)_2NR^\circ_2$; $-N(R^\circ)S(O)_2R^\circ$; $-N(OR^\circ)R^\circ$; $-C(NH)NR^\circ_2$; $-P(O)_2R^\circ$; $-P(O)R^\circ_2$; $-OP(O)R^\circ_2$; $-OP(O)(OR^\circ)_2$; SiR°_3; $-(C_{1-4}$ straight or branched alkylene$)O-N(R^\circ)_2$; or $-(C_{1-4}$ straight or branched alkylene$)C(O)O-N(R^\circ)_2$,

wherein each R° may be substituted as defined below and is independently hydrogen, C_{1-6} aliphatic, $-CH_2Ph$, $-O(CH_2)_{0-1}Ph$, $-CH_2$-(5-6 membered heteroaryl ring), or a 5–6–membered saturated, partially unsaturated, or aryl ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur, or, notwithstanding the definition above, two independent occurrences of R°, taken together with their intervening atom(s), form a 3–12–membered saturated, partially unsaturated, or aryl mono– or bicyclic ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur, which may be substituted as defined below.

[0051] Suitable monovalent substituents on R° (or the ring formed by taking two independent occurrences of R° together with their intervening atoms), are independently halogen, $-(CH_2)_{0-2}R^{\bullet}$, $-(haloR^{\bullet})$, $-(CH_2)_{0-2}OH$, $-(CH_2)_{0-2}OR^{\bullet}$, $-(CH_2)_{0-2}CH(OR^{\bullet})_2$; $-O(haloR^{\bullet})$, $-CN$, $-N_3$, $-(CH_2)_{0-2}C(O)R^{\bullet}$, $-(CH_2)_{0-2}C(O)OH$, $-(CH_2)_{0-2}C(O)OR^{\bullet}$, $-(CH_2)_{0-2}SR^{\bullet}$, $-(CH_2)_{0-2}SH$, $-(CH_2)_{0-2}NH_2$, $-(CH_2)_{0-2}NHR^{\bullet}$, $-(CH_2)_{0-2}NR^{\bullet}_2$, $-NO_2$, $-SiR^{\bullet}_3$, $-OSiR^{\bullet}_3$, $-C(O)SR^{\bullet}$, $-(C_{1-4}$ straight or branched alkylene$)C(O)OR^{\bullet}$, or $-SSR^{\bullet}$ wherein each $R^{\bullet}$ is unsubstituted or where preceded by "halo" is substituted only with one or more halogens, and is independently selected from C_{1-4} aliphatic, $-CH_2Ph$, $-O(CH_2)_{0-1}Ph$, or a 5–6–membered saturated, partially unsaturated, or aryl ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur. Suitable divalent substituents on a saturated carbon atom of R° include =O and =S.

[0052] Suitable divalent substituents on a saturated carbon atom of an "optionally substituted" group include the following: =O, =S, $=NNR^*_2$, $=NNHC(O)R^*$, $=NNHC(O)OR^*$, $=NNHS(O)_2R^*$, $=NR^*$, $=NOR^*$, $-O(C(R^*_2))_{2-3}O-$, or $-S(C(R^*_2))_{2-3}S-$, wherein each independent occurrence of R^* is selected from hydrogen, C_{1-6} aliphatic which may be substituted as defined below, or an unsubstituted 5–6–membered saturated, partially unsaturated, or aryl ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur. Suitable divalent substituents that are bound to vicinal substitutable carbons of an "optionally substituted" group include: $-O(CR^*_2)_{2-3}O-$, wherein each independent occurrence of R^* is selected from hydrogen, C_{1-6} aliphatic which may be substituted as defined below, or an unsubstituted 5–6–membered saturated, partially unsaturated, or aryl ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur.

[0053] Suitable substituents on the aliphatic group of R^* include halogen, –$R^\bullet$, -(halo$R^\bullet$), -OH, –$OR^\bullet$, –O(halo$R^\bullet$), –CN, –C(O)OH, –C(O)$OR^\bullet$, –NH_2, –$NHR^\bullet$, –$NR^\bullet_2$, or –NO_2, wherein each $R^\bullet$ is unsubstituted or where preceded by "halo" is substituted only with one or more halogens, and is independently C_{1-4} aliphatic, –CH_2Ph, –$O(CH_2)_{0-1}Ph$, or a 5–6–membered saturated, partially unsaturated, or aryl ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur.

[0054] Suitable substituents on a substitutable nitrogen of an "optionally substituted" group include –$R^\dagger$, –$NR^\dagger_2$, –C(O)$R^\dagger$, –C(O)$OR^\dagger$, –C(O)C(O)$R^\dagger$, –$C(O)CH_2C(O)R^\dagger$, –$S(O)_2R^\dagger$, -$S(O)_2NR^\dagger_2$, –C(S)$NR^\dagger_2$, –C(NH)$NR^\dagger_2$, or –$N(R^\dagger)S(O)_2R^\dagger$; wherein each $R^\dagger$ is independently hydrogen, C_{1-6} aliphatic which may be substituted as defined below, unsubstituted –OPh, or an unsubstituted 5–6–membered saturated, partially unsaturated, or aryl ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur, or, notwithstanding the definition above, two independent occurrences of $R^\dagger$, taken together with their intervening atom(s) form an unsubstituted 3–12–membered saturated, partially unsaturated, or aryl mono– or bicyclic ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur.

[0055] Suitable substituents on the aliphatic group of $R^\dagger$ are independently halogen, –$R^\bullet$, -(halo$R^\bullet$), –OH, –$OR^\bullet$, –O(halo$R^\bullet$), –CN, –C(O)OH, –C(O)$OR^\bullet$, –NH_2, –$NHR^\bullet$, –$NR^\bullet_2$, or -NO_2, wherein each $R^\bullet$ is unsubstituted or where preceded by "halo" is substituted only with one or more halogens, and is independently C_{1-4} aliphatic, –CH_2Ph, –$O(CH_2)_{0-1}Ph$, or a 5–6–membered saturated, partially unsaturated, or aryl ring having 0–4 heteroatoms independently selected from nitrogen, oxygen, or sulfur.

[0056] In another aspect, the present disclosure provides "pharmaceutically acceptable" compositions, which comprise a therapeutically effective amount of one or more of the compounds described herein, formulated together with one or more pharmaceutically acceptable carriers (additives) and/or diluents. As described in detail, the pharmaceutical compositions of the present disclosure may be specially formulated for administration in solid or liquid form, including those adapted for the following: oral administration, for example, drenches (aqueous or non-aqueous solutions or suspensions), tablets, *e.g.*, those targeted for buccal, sublingual, and systemic absorption, boluses, powders, granules, pastes for application to the tongue; parenteral administration, for example, by subcutaneous, intramuscular, intravenous or epidural injection

as, for example, a sterile solution or suspension, or sustained-release formulation; topical application, for example, as a cream, ointment, or a controlled-release patch or spray applied to the skin, lungs, or oral cavity; intravaginally or intrarectally, for example, as a pessary, cream or foam; sublingually; ocularly; transdermally; or nasally, pulmonary and to other mucosal surfaces.

[0057] The phrase "pharmaceutically acceptable" is employed herein to refer to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

[0058] The phrase "pharmaceutically acceptable carrier" as used herein means a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, or solvent encapsulating material, involved in carrying or transporting the subject compound from one organ, or portion of the body, to another organ, or portion of the body. Each carrier must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the patient. Some examples of materials which can serve as pharmaceutically-acceptable carriers include: sugars, such as lactose, glucose and sucrose; starches, such as corn starch and potato starch; cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; powdered tragacanth; malt; gelatin; talc; excipients, such as cocoa butter and suppository waxes; oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; glycols, such as propylene glycol; polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; esters, such as ethyl oleate and ethyl laurate; agar; buffering agents, such as magnesium hydroxide and aluminum hydroxide; alginic acid; pyrogen-free water; isotonic saline; Ringer's solution; ethyl alcohol; pH buffered solutions; polyesters, polycarbonates and/or polyanhydrides; and other non-toxic compatible substances employed in pharmaceutical formulations.

[0059] As used herein, the term "pharmaceutically acceptable salt" refers to those salts which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of humans and lower animals without undue toxicity, irritation, allergic response and the like, and are commensurate with a reasonable benefit/risk ratio. Pharmaceutically acceptable

salts are well known in the art. For example, S. M. Berge et al., describe pharmaceutically acceptable salts in detail in J. Pharmaceutical Sciences, 1977, 66, 1–19, incorporated herein by reference. Pharmaceutically acceptable salts of the compounds of this invention include those derived from suitable inorganic and organic acids and bases. Examples of pharmaceutically acceptable, nontoxic acid addition salts are salts of an amino group formed with inorganic acids such as hydrochloric acid, hydrobromic acid, phosphoric acid, sulfuric acid and perchloric acid or with organic acids such as acetic acid, oxalic acid, maleic acid, tartaric acid, citric acid, succinic acid or malonic acid or by using other methods used in the art such as ion exchange. Other pharmaceutically acceptable salts include adipate, alginate, ascorbate, aspartate, benzenesulfonate, benzoate, bisulfate, borate, butyrate, camphorate, camphorsulfonate, citrate, cyclopentanepropionate, digluconate, dodecylsulfate, ethanesulfonate, formate, fumarate, glucoheptonate, glycerophosphate, gluconate, hemisulfate, heptanoate, hexanoate, hydroiodide, 2–hydroxy–ethanesulfonate, lactobionate, lactate, laurate, lauryl sulfate, malate, maleate, malonate, methanesulfonate, 2–naphthalenesulfonate, nicotinate, nitrate, oleate, oxalate, palmitate, pamoate, pectinate, persulfate, 3–phenylpropionate, phosphate, pivalate, propionate, stearate, succinate, sulfate, tartrate, thiocyanate, p–toluenesulfonate, undecanoate, valerate salts, and the like.

[0060] Salts derived from appropriate bases include alkali metal, alkaline earth metal, ammonium and $N^+(C_{1-4}alkyl)_4$ salts. Representative alkali or alkaline earth metal salts include sodium, lithium, potassium, calcium, magnesium, and the like. Further pharmaceutically acceptable salts include, when appropriate, nontoxic ammonium, quaternary ammonium, and amine cations formed using counterions such as halide, hydroxide, carboxylate, sulfate, phosphate, nitrate, loweralkyl sulfonate and aryl sulfonate.

[0061] In certain embodiments, neutral forms of the compounds are regenerated by contacting the salt with a base or acid and isolating the parent compound in the conventional manner. In some embodiments, the parent form of the compound differs from the various salt forms in certain physical properties, such as solubility in polar solvents.

[0062] One of ordinary skill in the art will appreciate that the synthetic methods, as described herein, utilize a variety of protecting groups. By the term "protecting group," as used herein, it is meant that a particular functional moiety, *e.g.*, O, S, or N, is masked or blocked,

permitting, if desired, a reaction to be carried out selectively at another reactive site in a multifunctional compound. Suitable protecting groups are well known in the art and include those described in detail in *Protecting Groups in Organic Synthesis*, T. W. Greene and P. G. M. Wuts, 3rd edition, John Wiley & Sons, 1999, the entirety of which is incorporated herein by reference. In certain embodiments, a protecting group reacts selectively in good yield to give a protected substrate that is stable to the projected reactions; the protecting group is preferably selectively removable by readily available, preferably non-toxic reagents that do not attack the other functional groups; the protecting group forms a separable derivative (more preferably without the generation of new stereogenic centers); and the protecting group will preferably have a minimum of additional functionality to avoid further sites of reaction. As detailed herein, oxygen, sulfur, nitrogen, and carbon protecting groups may be utilized. By way of non-limiting example, hydroxyl protecting groups include methyl, methoxylmethyl (MOM), methylthiomethyl (MTM), *t*-butylthiomethyl, (phenyldimethylsilyl)methoxymethyl (SMOM), benzyloxymethyl (BOM), *p*-methoxybenzyloxymethyl (PMBM), (4-methoxyphenoxy)methyl (*p*-AOM), guaiacolmethyl (GUM), *t*-butoxymethyl, 4-pentenyloxymethyl (POM), siloxymethyl, 2-methoxyethoxymethyl (MEM), 2,2,2-trichloroethoxymethyl, bis(2-chloroethoxy)methyl, 2-(trimethylsilyl)ethoxymethyl (SEMOR), tetrahydropyranyl (THP), 3-bromotetrahydropyranyl, tetrahydrothiopyranyl, 1-methoxycyclohexyl, 4-methoxytetrahydropyranyl (MTHP), 4-methoxytetrahydrothiopyranyl, 4-methoxytetrahydrothiopyranyl S,S-dioxide, 1-[(2-chloro-4-methyl)phenyl]-4-methoxypiperidin-4-yl (CTMP), 1,4-dioxan-2-yl, tetrahydrofuranyl, tetrahydrothiofuranyl, 2,3,3a,4,5,6,7,7a-octahydro-7,8,8-trimethyl-4,7-methanobenzofuran-2-yl, 1-ethoxyethyl, 1-(2-chloroethoxy)ethyl, 1-methyl-1-methoxyethyl, 1-methyl-1-benzyloxyethyl, 1-methyl-1-benzyloxy-2-fluoroethyl, 2,2,2-trichloroethyl, 2-trimethylsilylethyl, 2-(phenylselenyl)ethyl, *t*-butyl, allyl, *p*-chlorophenyl, *p*-methoxyphenyl, 2,4-dinitrophenyl, benzyl, *p*-methoxybenzyl, 3,4-dimethoxybenzyl, *o*-nitrobenzyl, *p*-nitrobenzyl, *p*-halobenzyl, 2,6-dichlorobenzyl, *p*-cyanobenzyl, *p*-phenylbenzyl, 2-picolyl, 4-picolyl, 3-methyl-2-picolyl *N*-oxido, diphenylmethyl, *p,p'*-dinitrobenzhydryl, 5-dibenzosuberyl, triphenylmethyl, α-naphthyldiphenylmethyl, *p*-methoxyphenyldiphenylmethyl, di(*p*-methoxyphenyl)phenylmethyl, tri(*p*-methoxyphenyl)methyl, 4-(4'-bromophenacyloxyphenyl)diphenylmethyl, 4,4',4''-tris(4,5-dichlorophthalimidophenyl)methyl, 4,4',4''-tris(levulinoyloxyphenyl)methyl, 4,4',4''-

tris(benzoyloxyphenyl)methyl, 3-(imidazol-1-yl)bis(4',4''-dimethoxyphenyl)methyl, 1,1-bis(4-methoxyphenyl)-1'-pyrenylmethyl, 9-anthryl, 9-(9-phenyl)xanthenyl, 9-(9-phenyl-10-oxo)anthryl, 1,3-benzodithiolan-2-yl, benzisothiazolyl S,S-dioxido, trimethylsilyl (TMS), triethylsilyl (TES), triisopropylsilyl (TIPS), dimethylisopropylsilyl (IPDMS), diethylisopropylsilyl (DEIPS), dimethylthexylsilyl, *t*-butyldimethylsilyl (TBDMS), *t*-butyldiphenylsilyl (TBDPS), tribenzylsilyl, tri-*p*-xylylsilyl, triphenylsilyl, diphenylmethylsilyl (DPMS), *t*-butylmethoxyphenylsilyl (TBMPS), formate, benzoylformate, acetate, chloroacetate, dichloroacetate, trichloroacetate, trifluoroacetate, methoxyacetate, triphenylmethoxyacetate, phenoxyacetate, *p*-chlorophenoxyacetate, 3-phenylpropionate, 4-oxopentanoate (levulinate), 4,4-(ethylenedithio)pentanoate (levulinoyldithioacetal), pivaloate, adamantoate, crotonate, 4-methoxycrotonate, benzoate, *p*-phenylbenzoate, 2,4,6-trimethylbenzoate (mesitoate), alkyl methyl carbonate, 9-fluorenylmethyl carbonate (Fmoc), alkyl ethyl carbonate, alkyl 2,2,2-trichloroethyl carbonate (Troc), 2-(trimethylsilyl)ethyl carbonate (TMSEC), 2-(phenylsulfonyl) ethyl carbonate (Psec), 2-(triphenylphosphonio) ethyl carbonate (Peoc), alkyl isobutyl carbonate, alkyl vinyl carbonate alkyl allyl carbonate, alkyl *p*-nitrophenyl carbonate, alkyl benzyl carbonate, alkyl *p*-methoxybenzyl carbonate, alkyl 3,4-dimethoxybenzyl carbonate, alkyl *o*-nitrobenzyl carbonate, alkyl *p*-nitrobenzyl carbonate, alkyl *S*-benzyl thiocarbonate, 4-ethoxy-1-napththyl carbonate, methyl dithiocarbonate, 2-iodobenzoate, 4-azidobutyrate, 4-nitro-4-methylpentanoate, *o*-(dibromomethyl)benzoate, 2-formylbenzenesulfonate, 2-(methylthiomethoxy)ethyl, 4-(methylthiomethoxy)butyrate, 2-(methylthiomethoxymethyl)benzoate, 2,6-dichloro-4-methylphenoxyacetate, 2,6-dichloro-4-(1,1,3,3-tetramethylbutyl)phenoxyacetate, 2,4-bis(1,1-dimethylpropyl)phenoxyacetate, chlorodiphenylacetate, isobutyrate, monosuccinoate, (*E*)-2-methyl-2-butenoate, *o*-(methoxycarbonyl)benzoate, α-naphthoate, nitrate, alkyl *N,N,N',N'*-tetramethylphosphorodiamidate, alkyl *N*-phenylcarbamate, borate, dimethylphosphinothioyl, alkyl 2,4-dinitrophenylsulfenate, sulfate, methanesulfonate (mesylate), benzylsulfonate, and tosylate (Ts). For protecting 1,2- or 1,3-diols, the protecting groups include methylene acetal, ethylidene acetal, 1-*t*-butylethylidene ketal, 1-phenylethylidene ketal, (4-methoxyphenyl)ethylidene acetal, 2,2,2-trichloroethylidene acetal, acetonide, cyclopentylidene ketal, cyclohexylidene ketal, cycloheptylidene ketal, benzylidene acetal, *p*-methoxybenzylidene acetal, 3,4-dimethoxybenzylidene acetal, 2-nitrobenzylidene acetal, methoxymethylene acetal,

ethoxymethylene acetal, α-methoxybenzylidene ortho ester, α-(*N,N'*-dimethylamino)benzylidene derivative, 2-oxacyclopentylidene ortho ester, di-*t*-butylsilylene group (DTBS), 1,3-(1,1,3,3-tetraisopropyldisiloxanylidene) derivative (TIPDS), cyclic carbonates, cyclic boronates, ethyl boronate, and phenyl boronate. Exemplary protecting groups are detailed herein, however, it will be appreciated that the present invention is not intended to be limited to these protecting groups; rather, a variety of additional equivalent protecting groups can be readily identified using the above criteria and utilized in the method of the present invention. Additionally, a variety of protecting groups are described by Greene and Wuts (*supra*).

[0063] The symbol "〰", except when used as a bond to depict unknown or mixed stereochemistry, denotes the point of attachment of a chemical moiety to the remainder of a molecule or chemical formula.

[0064] As used herein, the term "isolated" refers to a substance and/or entity that has been (1) separated from at least some of the components with which it was associated when initially produced (whether in nature and/or in an experimental setting), and/or (2) designed, produced, prepared, and/or manufactured by the hand of man. Isolated substances and/or entities may be separated from about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% of the other components with which they were initially associated. In some embodiments, isolated agents are about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% pure. As used herein, a substance is "pure" if it is substantially free of other components. In some embodiments, as will be understood by those skilled in the art, a substance may still be considered "isolated" or even "pure", after having been combined with certain other components such as, for example, one or more carriers or excipients (*e.g.*, buffer, solvent, water, etc.); in such embodiments, percent isolation or purity of the substance is calculated without including such carriers or excipients. In some embodiments, isolation involves or requires disruption of covalent bonds (e.g., to isolate a polypeptide domain from a longer polypeptide and/or to isolate a nucleotide sequence element from a longer oligonucleotide or nucleic acid).

[0065] The term "modulator" is used to refer to an entity whose presence in a system in which an activity of interest is observed correlates with a change in level and/or nature of that activity as compared with that observed under otherwise comparable conditions when the modulator is absent. In some embodiments, a modulator is an activator or agonist, in that activity is increased in its presence as compared with that observed under otherwise comparable conditions when the modulator is absent. In some embodiments, a modulator is an inhibitor or antagonist, in that activity is reduced in its presence as compared with otherwise comparable conditions when the modulator is absent. In some embodiments, a modulator interacts directly with a target entity whose activity is of interest. In some embodiments, a modulator interacts indirectly (i.e., directly with an intermediate agent that interacts with the target entity) with a target entity whose activity is of interest. In some embodiments, a modulator affects level of a target entity of interest; alternatively or additionally, in some embodiments, a modulator affects activity of a target entity of interest without affecting level of the target entity. In some embodiments, a modulator affects both level and activity of a target entity of interest, so that an observed difference in activity is not entirely explained by or commensurate with an observed difference in level. As used herein, an "activity" is any process, carried out by a molecule, compound, cell, tissue or organ. Such processes may be catalytic or non-catalytic. For example, the cGAS molecules of the present invention may act as enzymes and as such may have enzymatic activity.

[0066] The term "nucleic acid," in its broadest sense, includes any compound and/or substance that comprise a polymer of nucleotides. These polymers are often referred to as polynucleotides. Exemplary nucleic acids or polynucleotides of the invention include, but are not limited to, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs, including LNA having a β- D-ribo configuration, α-LNA having an α-L-ribo configuration (a diastereomer of LNA), 2′-amino-LNA having a 2′-amino functionalization, and 2′-amino- α-LNA having a 2′-amino functionalization) or hybrids thereof.

[0067] The present disclosure provides for modified nucleosides and nucleotides. As described herein "nucleoside" is defined as a compound containing a sugar molecule (e.g., a pentose or ribose) or a derivative thereof in combination with an organic base (e.g., a purine or

pyrimidine) or a derivative thereof (also referred to herein as "nucleobase"). As described herein, "nucleotide" is defined as a nucleoside including a phosphate group. The modified nucleotides may by synthesized by any useful method, as described herein (e.g., chemically, enzymatically, or recombinantly to include one or more modified or non-natural nucleosides).

[0068] The modified nucleotide base pairing encompasses not only the standard adenosine-thymine, adenosine-uracil, or guanosine-cytosine base pairs, but also base pairs formed between nucleotides and/or modified nucleotides comprising non-standard or modified bases, wherein the arrangement of hydrogen bond donors and hydrogen bond acceptors permits hydrogen bonding between a non-standard base and a standard base or between two complementary non-standard base structures. One example of such non-standard base pairing is the base pairing between the modified nucleotide inosine and adenine, cytosine or uracil.

[0069] The modified nucleosides and nucleotides can include a modified nucleobase. Examples of nucleobases found in RNA include, but are not limited to, adenine, guanine, cytosine, and uracil. Examples of nucleobase found in DNA include, but are not limited to, adenine, guanine, cytosine, and thymine.

[0070] As will be clear from context, in some embodiments, "nucleic acid" refers to individual nucleic acid residues (*e.g.*, nucleotides and/or nucleosides); in some embodiments, "nucleic acid" refers to an oligonucleotide chain comprising individual nucleic acid residues. In some embodiments, a "nucleic acid" is or comprises RNA; in some embodiments, a "nucleic acid" is or comprises DNA. In some embodiments, a nucleic acid is, comprises, or consists of one or more natural nucleic acid residues. In some embodiments, a nucleic acid is, comprises, or consists of one or more nucleic acid analogs. In some embodiments, a nuclic acid analog differs from a nucleic acid in that it does not utilize a phosphodiester backbone. For example, in some embodiments, a nucleic acid is, comprises, or consists of one or more "peptide nucleic acids", which are known in the art and have peptide bonds instead of phosphodiester bonds in the backbone, are considered within the scope of the present invention. Alternatively or additionally, in some embodiments, a nucleic acid has one or more phosphorothioate and/or 5'-N-phosphoramidite linkages rather than phosphodiester bonds. In some embodiments, a nucleic acid is, comprises, or consists of one or more natural nucleosides (e.g., adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and

deoxycytidine). In some embodiments, a nucleic acid is, comprises, or consists of one or more nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, C-5 propynyl-cytidine, C-5 propynyl-uridine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, O(6)-methylguanine, 2-thiocytidine, methylated bases, intercalated bases, and combinations thereof). In some embodiments, a nucleic acid comprises one or more modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, arabinose, and hexose) as compared with those in natural nucleic acids. In some embodiments, a nucleic acid has a nucleotide sequence that encodes a functional gene product such as an RNA or protein. In some embodiments, a nucleic acid includes one or more introns. In some embodiments, nucleic acids are prepared by one or more of isolation from a natural source, enzymatic synthesis by polymerization based on a complementary template (*in vivo* or *in vitro*), reproduction in a recombinant cell or system, and chemical synthesis. In some embodiments, a nucleic acid is at least 2 (dinucleotide or dinucleoside), 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 20, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 or more residues long.

[0071] The term "polypeptide", as used herein, generally has its art-recognized meaning of a polymer of at least three amino acids, linked to one another by peptide bonds. In some embodiments, the term is used to refer to specific functional classes of polypeptides, such as, for example, receptors, enzymes, signaling proteins, structural proteins, autoantigen polypeptides, nicotinic acetylcholine receptor polypeptides, alloantigen polypeptides, etc. For each such class, the present specification provides several examples of amino acid sequences of known exemplary polypeptides within the class; in some embodiments, such known polypeptides are reference polypeptides for the class. In some instances the polypeptide encoded is smaller than about 50 amino acids and the polypeptide is then termed a peptide. If the polypeptide is a peptide, it will be at least about 2, 3, 4, or at least 5 amino acid residues long. Thus, polypeptides include gene products, naturally occurring polypeptides, synthetic polypeptides, homologs, orthologs, paralogs, fragments and other equivalents, variants, and analogs of the foregoing. A

polypeptide may be a single molecule or may be a multi-molecular complex such as a dimer, trimer, or tetramer. They may also comprise single chain or multichain polypeptides such as antibodies or insulin and may be associated or linked. Most commonly disulfide linkages are found in multichain polypeptides. The term polypeptide may also apply to amino acid polymers in which one or more amino acid residues are an artificial chemical analogue of a corresponding naturally occurring amino acid. In such embodiments, the term "polypeptide" refers to any member of the class that shows significant sequence homology or identity with a relevant reference polypeptide. In many embodiments, such member also shares significant activity with the reference polypeptide. For example, in some embodiments, a member polypeptide shows an overall degree of sequence homology or identity with a reference polypeptide that is at least about 30-40%, and is often greater than about 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more and/or includes at least one region (i.e., a conserved region, often including a characteristic sequence element) that shows very high sequence identity, often greater than 90% or even 95%, 96%, 97%, 98%, or 99%. Such a conserved region usually encompasses at least 3-4 and often up to 20 or more amino acids; in some embodiments, a conserved region encompasses at least one stretch of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more contiguous amino acids. In some embodiments, a useful polypeptide as described herein may comprise or consist of a fragment of a parent polypeptide. In some embodiments, a useful polypeptide as described herein may comprise or consist of a plurality of fragments, each of which is found in the same parent polypeptide in a different spatial arrangement relative to one another than is found in the polypeptide of interest (e.g., fragments that are directly linked in the parent may be spatially separated in the polypeptide of interest or vice versa, and/or fragments may be present in a different order in the polypeptide of interest than in the parent), so that the polypeptide of interest is a derivative of its parent polypeptide.

[0072] The term "polypeptide variant" refers to molecules which differ in their amino acid sequence from a native or reference sequence. The amino acid sequence variants may possess substitutions, deletions, and/or insertions at certain positions within the amino acid sequence, as compared to a native or reference sequence. Ordinarily, variants will possess at least about 50% identity (homology) to a native or reference sequence, and preferably, they will

be at least about 80%, more preferably at least about 90% identical (homologous) to a native or reference sequence.

[0073] In some embodiments "variant mimics" are provided. As used herein, the term "variant mimic" is one which contains one or more amino acids which would mimic an activated sequence. For example, glutamate may serve as a mimic for phosphoro-threonine and/or phosphoro-serine. Alternatively, variant mimics may result in deactivation or in an inactivated product containing the mimic, e.g., phenylalanine may act as an inactivating substitution for tyrosine; or alanine may act as an inactivating substitution for serine.

[0074] "Homology" as it applies to amino acid sequences is defined as the percentage of residues in the candidate amino acid sequence that are identical with the residues in the amino acid sequence of a second sequence after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology. Methods and computer programs for the alignment are well known in the art. It is understood that homology depends on a calculation of percent identity but may differ in value due to gaps and penalties introduced in the calculation.

[0075] By "homologs" as it applies to polypeptide sequences means the corresponding sequence of other species having substantial identity to a second sequence of a second species.

[0076] "Analogs" in the context of polypeptides is meant to include polypeptide variants which differ by one or more amino acid alterations, e.g., substitutions, additions or deletions of amino acid residues that still maintain one or more of the properties of the parent or starting polypeptide.

[0077] As used herein, the term "protein" refers to a polypeptide (*i.e.*, a string of at least two amino acids linked to one another by peptide bonds). Proteins may include moieties other than amino acids (*e.g.*, may be glycoproteins, proteoglycans, etc.) and/or may be otherwise processed or modified. Those of ordinary skill in the art will appreciate that a "protein" can be a complete polypeptide chain as produced by a cell (with or without a signal sequence), or can be a characteristic portion thereof. Those of ordinary skill will appreciate that a protein can sometimes include more than one polypeptide chain, for example linked by one or more disulfide bonds or associated by other means. Polypeptides may contain L-amino acids, D-amino acids, or both and may contain any of a variety of amino acid modifications or analogs known in the art.

Useful modifications include, *e.g.*, terminal acetylation, amidation, methylation, etc. In some embodiments, proteins may comprise natural amino acids, non-natural amino acids, synthetic amino acids, and combinations thereof. The term "peptide" is generally used to refer to a polypeptide having a length of less than about 100 amino acids, less than about 50 amino acids, less than 20 amino acids, or less than 10 amino acids. In some embodiments, proteins are antibodies, antibody fragments, biologically active portions thereof, and/or characteristic portions thereof.

[0078] The phrases "parenteral administration" and "administered parenterally" as used herein means modes of administration other than enteral and topical administration, usually by injection, and includes, without limitation, intravenous, intramuscular, intraarterial, intrathecal, intracapsular, intraorbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticulare, subcapsular, subarachnoid, intraspinal, and intrasternal injection and infusion.

[0079] The phrases "systemic administration," "administered systemically," "peripheral administration" and "administered peripherally" as used herein mean the administration of a compound, drug or other material other than directly into the central nervous system, such that it enters the patient's system and, thus, is subject to metabolism and other like processes, for example, subcutaneous administration.

[0080] The term "palliative" refers to treatment that is focused on the relief of symptoms of a disease and/or side effects of a therapeutic regimen, but is not curative.

[0081] The term "therapeutic agent" or "therapeutic modality" refers to any agent that, when administered to a subject, has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect.

[0082] As used herein, the term "therapeutically effective amount" means an amount of a substance (*e.g.*, a therapeutic agent, composition, and/or formulation) that elicits a desired biological response when administered as part of a therapeutic regimen. In some embodiments, a therapeutically effective amount of a substance is an amount that is sufficient, when administered to a subject suffering from or susceptible to a disease, disorder, and/or condition, to treat the disease, disorder, and/or condition. As will be appreciated by those of ordinary skill in this art,

the effective amount of a substance may vary depending on such factors as the desired biological endpoint, the substance to be delivered, the target cell or tissue, *etc.* For example, the effective amount of compound in a formulation to treat a disease, disorder, and/or condition is the amount that alleviates, ameliorates, relieves, inhibits, prevents, delays onset of, reduces severity of and/or reduces incidence of one or more symptoms or features of the disease, disorder, and/or condition. In some embodiments, a therapeutically effective amount is administered in a single dose; in some embodiments, multiple unit doses are required to deliver a therapeutically effective amount.

[0083] As used herein, the term "treat," "treatment," or "treating" refers to any method used to partially or completely alleviate, ameliorate, relieve, inhibit, prevent, delay onset of, reduce severity of and/or reduce incidence of one or more symptoms or features of a disease, disorder, and/or condition. Treatment may be administered to a subject who does not exhibit signs of a disease, disorder, and/or condition. In some embodiments, treatment may be administered to a subject who exhibits only early signs of the disease, disorder, and/or condition for the purpose of decreasing the risk of developing pathology associated with the disease, disorder, and/or condition.

[0084] The expression "unit dose" as used herein refers to a physically discrete unit of a formulation appropriate for a subject to be treated. It will be understood, however, that the total daily usage of a formulation of the present invention will be decided by the attending physician within the scope of sound medical judgment. The specific effective dose level for any particular subject or organism may depend upon a variety of factors including the disorder being treated and the severity of the disorder; activity of specific active compound employed; specific composition employed; age, body weight, general health, sex and diet of the subject; time of administration, and rate of excretion of the specific active compound employed; duration of the treatment; drugs and/or additional therapies used in combination or coincidental with specific compound(s) employed, and like factors well known in the medical arts. A particular unit dose may or may not contain a therapeutically effective amount of a therapeutic agent.

[0085] As used herein, the term "patient" or "subject" refers to a human or any non-human animal (e.g., mouse, rat, rabbit, dog, cat, cattle, swine, sheep, horse or primate) to whom therapy is administered. In many embodiments, a patient is a human being. A human includes

pre and post natal forms. In some embodiments, a patient is a human presenting to a medical provider for diagnosis or treatment of a disease, disorder or condition. In some embodiments, a patient displays one or more symptoms or characteristics of a disease, disorder or condition. In some embodiments, a patient does not display any symptom or characteristic of a disease, disorder, or condition. In some embodiments, a patient is someone with one or more features characteristic of susceptibility to or risk of a disease, disorder, or condition.

[0086] As used herein, the term "sample" or "biological sample" refers to a subset of its tissues, cells or component parts (e.g. body fluids, including but not limited to blood, mucus, lymphatic fluid, synovial fluid, cerebrospinal fluid, saliva, amniotic fluid, amniotic cord blood, urine, vaginal fluid and semen). A sample further may include a homogenate, lysate or extract prepared from a whole organism or a subset of its tissues, cells or component parts, or a fraction or portion thereof, including but not limited to, for example, plasma, serum, spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, blood cells, tumors, organs. A sample further refers to a medium, such as a nutrient broth or gel, which may contain cellular components, such as proteins or nucleic acid molecule.

[0087] As used herein "stable" refers to a compound or molecule that is sufficiently robust to survive isolation to a useful degree of purity from a reaction mixture, and preferably capable of formulation into an efficacious therapeutic agent. Alternatively a compound or molecule may be said to be stable if it is sufficiently robust to withstand any treatment, insult or utilization without undergoing substantial degradation prior to a selected timepoint, event or localization.

[0088] As used herein, the term "stabilize", "stabilized," "stabilized region" means to make or become stable.

[0089] An individual who is "suffering from" a disease, disorder, and/or condition has been diagnosed with and/or displays one or more symptoms of the disease, disorder, and/or condition.

[0090] An individual who is "susceptible to" a disease, disorder, and/or condition has not been diagnosed with the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition may exhibit symptoms of

the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition may not exhibit symptoms of the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition will develop the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition will not develop the disease, disorder, and/or condition.

[0091] The term "computer-readable medium", as used herein, refers to non-volatile (i.e. secondary storage) computer data storage and/or memory to retain digital data even when not powered. Examples of computer-readable medium include, but are not limited to hard disk, floppy disk, flash memory(i.e. solid state memory), Ferroelectric RAM (F-RAM), Magnetoresistive RAM (MRAM), optical disc, standalone RAM disks, ZIP drives, magenetic tape and holographic memory.

[0092] The term "computer system" or "computer", as used herein, refers to a computing device that can be used to implement the techniques described in this disclosure. An exemplary computing device 2500 and a mobile computing device are shown in FIG. S11.

[0093] As used herein, the term "crystal structure" of a composition shall mean a computer readable medium in which is stored a representation of three dimensional positional information (i.e. coordinates) for atoms of the composition.

[0094] As used herein, the term "docking" refers to orienting, rotating, translating a chemical entity in the binding pocket, domain, molecule or molecular complex or portion thereof based on distance geometry or energy. Docking may be performed by distance geometry methods that find sets of atoms of a chemical entity that match sets of sphere centers of the binding pocket, domain, molecule or molecular complex or portion thereof. See Meng et al. J. Comp. Chem. 4: 505-524 (1992). Sphere centers are generated by providing an extra radius of given length from the atoms (excluding hydrogen atoms) in the binding pocket, domain, molecule or molecular complex or portion thereof. Real-time interaction energy calculations, energy minimizations or rigid-body minimizations (Gschwend et al., J. Mol. Recognition 9:175-186 (1996)) can be performed while orienting the chemical entity to facilitate docking. For example, interactive docking experiments can be designed to follow the path of least resistance.

If the user in an interactive docking experiment makes a move to increase the energy, the system will resist that move. However, if that user makes a move to decrease energy, the system will favor that move by increased responsiveness. (Cohen et al., J. Med. Chem. 33:889-894 (1990)). Docking can also be performed by combining a Monte Carlo search technique with rapid energy evaluation using molecular affinity potentials. See Goodsell and Olson, Proteins: Structure, Function and Genetics 8:195-202 (1990). Software programs that carry out docking functions include but are not limited to MATCHMOL (Cory et al., J. Mol. Graphics 2: 39 (1984); MOLFIT (Redington, Comput. Chem. 16: 217 (1992)) and DOCK (Meng et al., supra).

[0095] As used herein, the term "designed" refers to an agent (i) whose structure is or was selected by the hand of man; (ii) that is produced by a process requiring the hand of man; and/or (iii) that is distinct from natural substances and other known agents.

[0096] As used herein, the term "storage environment" comprises any environment comprising secondary storage, i.e. long-term persistent storage. In some embodiments, a storage environment comprises computer-readable medium. In some embodiments, a storage environment comprises a network environment for establishing a multi-channel context aware communication environment (i.e. cloud computing). For example, Figure S11 is a block diagram of a network environment for establishing a multi-channel context aware communication environment.

[0097] As used herein, the term "substantially" refers to the qualitative condition of exhibiting total or near-total extent or degree of a characteristic or property of interest. One of ordinary skill in the biological arts will understand that biological and chemical phenomena rarely, if ever, go to completion and/or proceed to completeness or achieve or avoid an absolute result. The term "substantially" is therefore used herein to capture the potential lack of completeness inherent in many biological and chemical phenomena.

Compounds and Compositions of the Invention

[0098] It has surprisingly been discovered, contrary to the disclosures in the art, that the founding member of a family of metazoan cyclic dinucleotide second messengers regulating type

I interferon induction in response to cytoplasmic DNA comprises unique features heretofor not known. These insights now afford the opportunity to design analogs, mimics, and/or mimetics of the molecule identified and to further use the parent molecule or molecules designed based on the structure of the parent molecules in research and development.

cGAMP analogs, mimics, mimetics and modifications

[0099] The family of second messengers, termed cGAMPs or cyclic GAMPs, includes one or more of the cyclic structures defined as having at least one guanine (G) and one adenine (A) nucleotide and being linked in a cyclic fashion. The linkages between the two nucleotides involve sugar to backbone bond formation. According to the present invention, there are four primary parent members of the cGAMP family. These include [c[G(2',5')pA(3',5')p]], [c[G(2',5')pA(2',5')p]], [c[G(3',5')pA(2',5')p]], [c[G(3',5')pA(3',5')p]]. Furthermore, each of the nucleotides of the pairs may adopt either *syn* or *anti* glycosidic torsion orientations.

[0100] As used herein, the term "cGAMP" refers to any of the parent molecules of the family as well as any of the possible torsion orientations. Individual members of the family may be referred to by their sugar-backbone linkage form, e.g., the newly discovered second messenger, [c[G(2',5')pA(3',5')p]], may be referred to as the "2-prime-3'prime" isomer, referencing the position on the sugar ring forming the sugar-backbone bond. The other isomers may be named likewise. Collectively, the compounds of the invention which are wild type, analogs, mimics, mimetics or modified versions of the cGAMP family, are referred to as the group of "cGMP compounds."

[0101] With the teachings provided herein, one of skill may now design analogs, mimics, or modifications to any of the parent molecules for use as a modulator, either agonist or antagonist, of the cGAS enzyme, and ultimately as a modulator of downstream physiologic events associated with interferon signaling. Linear versions may also be designed as either agonists, antagonists, or competitive inhibitors of cGAS or downstream signaling events associated with either cGAS enzyme activity or interferon signaling.

[0102] The present invention contemplates several types of compounds or compositions which are nucleic acid based including variants and derivatives. These include substitutional, insertional, deletion and covalent variants and derivatives. The term "derivative" is used

synonymously with the term "variant" but generally refers to a molecule that has been modified and/or changed in any way relative to a reference molecule or starting molecule.

[0103] As used herein, an "analog" is meant to include cGAMP variants which differ by one or more alterations, e.g., substitutions, additions or deletions that still maintain one or more of the properties of the parent or starting molecule. Analogs are typically designed using structure-activity relationships (SAR) such as those described herein.

cGAS proteins, variants, derivatives and mutants

[0104] Having now in hand several protein crystal structures in native and varying binding states, it is possible to exploit these protein structures by designing variants, derivatives or mutants of the cGAS enzyme. As such, cGAS enzyme polypeptides, including their variants, derivatives and mutants are considered compounds of the invention. These variants, derivatives and mutants are useful as research tools, for example in kits or assays or as the source of a therapeutic modality. To this end, fragments or portions of the cGAS polypeptide or the variants, derivatives and mutants may be used as antigens for the production of antibodies, or where the fragment maintains a structural element associated with activity, whether binding, catalysis, or transport may also be used as a modulator of the enzyme itself or as a surrogate for dsDNA binding. Collectively, the compounds of the invention which are wild type, variants, derivatives or mutants of cGAS, are referred to as the group of "cGAS molecules." cGAS molecules may comprise any portion or fragment of a cGAS molecule or may comprise mixed domains or fragments from cGAS molecules arising from different structures as defined by the crystal structures disclosed herein.

[0105] The present invention contemplates several types of compositions which are polypeptide based including variants and derivatives. These include substitutional, insertional, deletion and covalent variants and derivatives. The term "derivative" is used synonymously with the term "variant" but generally refers to a molecule that has been modified and/or changed in any way relative to a reference molecule or starting molecule.

[0106] As such, cGAS encoding polypeptides containing substitutions, insertions, and/or additions, deletions and covalent modifications with respect to reference sequences, in particular the polypeptide sequences disclosed herein, are included within the scope of this invention. For example, sequence tags or amino acids, such as one or more lysines, can be added to the peptide sequences of the invention (e.g., at the N-terminal or C-terminal ends). Sequence tags can be used for peptide purification or localization. Lysines can be used to increase peptide solubility or to allow for biotinylation. Alternatively, amino acid residues located at the carboxy and amino terminal regions of the amino acid sequence of a peptide or protein may optionally be deleted providing for truncated sequences. Certain amino acids (e.g., C-terminal or N-terminal residues) may alternatively be deleted depending on the use of the sequence, as for example, expression of the sequence as part of a larger sequence which is soluble, or linked to a solid support.

[0107] "Substitutional variants" when referring to polypeptides are those that have at least one amino acid residue in a native or starting sequence removed and a different amino acid inserted in its place at the same position. The substitutions may be single, where only one amino acid in the molecule has been substituted, or they may be multiple, where two or more amino acids have been substituted in the same molecule.

[0108] As used herein the term "conservative amino acid substitution" refers to the substitution of an amino acid that is normally present in the sequence with a different amino acid of similar size, charge, or polarity. Examples of conservative substitutions include the substitution of a non-polar (hydrophobic) residue such as isoleucine, valine and leucine for another non-polar residue. Likewise, examples of conservative substitutions include the substitution of one polar (hydrophilic) residue for another such as between arginine and lysine, between glutamine and asparagine, and between glycine and serine. Additionally, the substitution of a basic residue such as lysine, arginine or histidine for another, or the substitution of one acidic residue such as aspartic acid or glutamic acid for another acidic residue are additional examples of conservative substitutions. Examples of non-conservative substitutions include the substitution of a non-polar (hydrophobic) amino acid residue such as isoleucine, valine, leucine, alanine, methionine for a polar (hydrophilic) residue such as cysteine, glutamine, glutamic acid or lysine and/or a polar residue for a non-polar residue.

[0109] "Insertional variants" when referring to polypeptides are those with one or more amino acids inserted immediately adjacent to an amino acid at a particular position in a native or starting sequence. "Immediately adjacent" to an amino acid means connected to either the alpha-carboxy or alpha-amino functional group of the amino acid.

[0110] "Deletional variants" when referring to polypeptides are those with one or more amino acids in the native or starting amino acid sequence removed. Ordinarily, deletional variants will have one or more amino acids deleted in a particular region of the molecule.

[0111] "Covalent derivatives" when referring to polypeptides include modifications of a native or starting protein with an organic proteinaceous or non-proteinaceous derivatizing agent, and/or post-translational modifications. Covalent modifications are traditionally introduced by reacting targeted amino acid residues of the protein with an organic derivatizing agent that is capable of reacting with selected side-chains or terminal residues, or by harnessing mechanisms of post-translational modifications that function in selected recombinant host cells. The resultant covalent derivatives are useful in programs directed at identifying residues important for biological activity, for immunoassays, or for the preparation of anti-protein antibodies for immunoaffinity purification of the recombinant glycoprotein. Such modifications are within the ordinary skill in the art and are performed without undue experimentation.

[0112] Certain post-translational modifications are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues may be present in the polypeptides produced in accordance with the present invention.

[0113] Other post-translational modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (T. E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)) the contents of which are incorporated by reference it its entirety.

[0114] "Features" when referring to polypeptides are defined as distinct amino acid sequence-based components of a molecule. Features of the cGAS polypeptides encoded by the

present invention include surface manifestations, local conformational shape, folds, loops, half-loops, domains, half-domains, sites, termini or any combination thereof.

[0115] As used herein when referring to polypeptides the term "surface manifestation" refers to a polypeptide based component of a protein appearing on an outermost surface.

[0116] As used herein when referring to polypeptides the term "local conformational shape" means a polypeptide based structural manifestation of a protein which is located within a definable space of the protein.

[0117] As used herein when referring to polypeptides the term "fold" refers to the resultant conformation of an amino acid sequence upon energy minimization. A fold may occur at the secondary or tertiary level of the folding process. Examples of secondary level folds include beta sheets and alpha helices. Examples of tertiary folds include domains and regions formed due to aggregation or separation of energetic forces. Regions formed in this way include hydrophobic and hydrophilic pockets, and the like.

[0118] As used herein the term "turn" as it relates to protein conformation means a bend which alters the direction of the backbone of a peptide or polypeptide and may involve one, two, three or more amino acid residues.

[0119] As used herein when referring to polypeptides the term "loop" refers to a structural feature of a polypeptide which may serve to reverse the direction of the backbone of a peptide or polypeptide. Where the loop is found in a polypeptide and only alters the direction of the backbone, it may comprise four or more amino acid residues. Oliva et al. have identified at least 5 classes of protein loops (J. Mol Biol 266 (4): 814-830; 1997). Loops may be open or closed. Closed loops or "cyclic" loops may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids between the bridging moieties. Such bridging moieties may comprise a cysteine-cysteine bridge (Cys-Cys) typical in polypeptides having disulfide bridges or alternatively bridging moieties may be non-protein based such as the dibromozylyl agents used herein.

[0120] As used herein when referring to polypeptides the term "half-loop" refers to a portion of an identified loop having at least half the number of amino acid resides as the loop from which it is derived. It is understood that loops may not always contain an even number of amino acid residues. Therefore, in those cases where a loop contains or is identified to comprise

an odd number of amino acids, a half-loop of the odd-numbered loop will comprise the whole number portion or next whole number portion of the loop (number of amino acids of the loop/2+/-0.5 amino acids). For example, a loop identified as a 7 amino acid loop could produce half-loops of 3 amino acids or 4 amino acids (7/2=3.5+/-0.5 being 3 or 4).

[0121] As used herein when referring to polypeptides the term "domain" refers to a motif of a polypeptide having one or more identifiable structural or functional characteristics or properties (e.g., binding capacity, serving as a site for protein-protein interactions).

[0122] As used herein when referring to polypeptides the term "half-domain" means a portion of an identified domain having at least half the number of amino acid resides as the domain from which it is derived. It is understood that domains may not always contain an even number of amino acid residues. Therefore, in those cases where a domain contains or is identified to comprise an odd number of amino acids, a half-domain of the odd-numbered domain will comprise the whole number portion or next whole number portion of the domain (number of amino acids of the domain/2+/-0.5 amino acids). For example, a domain identified as a 7 amino acid domain could produce half-domains of 3 amino acids or 4 amino acids (7/2=3.5+/-0.5 being 3 or 4). It is also understood that sub-domains may be identified within domains or half-domains, these subdomains possessing less than all of the structural or functional properties identified in the domains or half domains from which they were derived. It is also understood that the amino acids that comprise any of the domain types herein need not be contiguous along the backbone of the polypeptide (i.e., nonadjacent amino acids may fold structurally to produce a domain, half-domain or subdomain).

[0123] As used herein when referring to polypeptides the terms "site" as it pertains to amino acid based embodiments is used synonymously with "amino acid residue" and "amino acid side chain." A site represents a position within a peptide or polypeptide that may be modified, manipulated, altered, derivatized or varied within the polypeptide based molecules of the present invention.

[0124] As used herein the terms "termini" or "terminus" when referring to polypeptides refers to an extremity of a peptide or polypeptide. Such extremity is not limited only to the first or final site of the peptide or polypeptide but may include additional amino acids in the terminal

regions. The polypeptide based molecules of the present invention may be characterized as having both an N-terminus (terminated by an amino acid with a free amino group (NH2)) and a C-terminus (terminated by an amino acid with a free carboxyl group (COOH)). Proteins of the invention are in some cases made up of multiple polypeptide chains brought together by disulfide bonds or by non-covalent forces (multimers, oligomers). These sorts of proteins will have multiple N- and C-termini. Alternatively, the termini of the polypeptides may be modified such that they begin or end, as the case may be, with a non-polypeptide based moiety such as an organic conjugate.

[0125] Once any of the features have been identified or defined as a desired component of a polypeptide of the invention, any of several manipulations and/or modifications of these features may be performed by moving, swapping, inverting, deleting, randomizing or duplicating. Furthermore, it is understood that manipulation of features may result in the same outcome as a modification to the molecules of the invention. For example, a manipulation which involved deleting a domain would result in the alteration of the length of a molecule just as modification of a nucleic acid to encode less than a full length molecule would.

[0126] Modifications and manipulations can be accomplished by methods known in the art such as, but not limited to, site directed mutagenesis. The resulting modified molecules may then be tested for activity using *in vitro* or *in vivo* assays such as those described herein or any other suitable screening assay known in the art.

[0127] According to the present invention, the polypeptides may comprise a consensus sequence which is discovered through rounds of experimentation. As used herein a "consensus" sequence is a single sequence which represents a collective population of sequences allowing for variability at one or more sites.

[0128] As recognized by those skilled in the art, protein fragments, functional protein domains, and homologous proteins are also considered to be within the scope of polypeptides of interest of this invention. For example, provided herein is any protein fragment (meaning a polypeptide sequence at least one amino acid residue shorter than a reference polypeptide sequence but otherwise identical) of a reference protein 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or greater than 100 amino acids in length. In another example, any protein that includes a stretch of

about 20, about 30, about 40, about 50, or about 100 amino acids which are about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, or about 100% identical to any of the sequences described herein can be utilized in accordance with the invention. In certain embodiments, a polypeptide to be utilized in accordance with the invention includes 2, 3, 4, 5, 6, 7, 8, 9, 10, or more mutations as shown in any of the sequences provided or referenced herein.

Antibodies

[0129] In some embodiments, the cGAMP compounds or cGAS molecules may be used to generate antibodies. As such, the antibodies so generated are considered further compounds and compositions of the present invention. As used herein, term "antibody" includes monoclonal antibodies (including full length antibodies which have an immunoglobulin Fc region), antibody compositions with polyepitopic specificity, multispecific antibodies (e.g., bispecific antibodies, diabodies, and single-chain molecules), as well as antibody fragments. The term "immunoglobulin" (Ig) is used interchangeably with "antibody" herein. As used herein, the term "monoclonal antibody" refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations and/or post-translation modifications (e.g., isomerizations, amidations) that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site.

[0130] The monoclonal antibodies herein specifically include "chimeric" antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is(are) identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity. Chimeric antibodies of interest herein include, but are not limited to, "primatized" antibodies comprising variable domain antigen-binding sequences derived from a non-human primate (e.g., Old World Monkey, Ape etc.) and human constant region sequences.

[0131] An "antibody fragment" comprises a portion of an intact antibody, preferably the antigen binding and/or the variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', $F(ab')_2$ and Fv fragments; diabodies; linear antibodies; nanobodies; single-chain antibody molecules and multispecific antibodies formed from antibody fragments.

[0132] Any of the five classes of immunoglobulins, IgA, IgD, IgE, IgG and IgM, may be generated by the compounds or molecules of the invention, including the heavy chains designated alpha, delta, epsilon, gamma and mu, respectively.

[0133] While not wishing to be bound by theory, it is believed that antibodies generated using the cGAMP compounds or cGAS molecules disclosed herein will result in improved therapeutic efficacy.

[0134] Antibodies of the invention may be utilized to treat conditions or diseases in many therapeutic areas such as, but not limited to, blood, cardiovascular, CNS, poisoning (including antivenoms), dermatology, endocrinology, gastrointestinal, medical imaging, musculoskeletal, oncology, immunology, inflammation, respiratory, sensory and anti-infective.

[0135] In one embodiment, variants of antibodies may also include, but are not limited to, substitutional variants, conservative amino acid substitution, insertional variants, deletional variants and/or covalent derivatives.

Vaccines

[0136] As used herein, a "vaccine" is a biological preparation that improves immunity to a particular disease or infectious agent. According to the present invention and while not wishing to be bound by theory, it is believed that utilization of the cGAMP compounds or cGAS molecules of the invention may be used as a vaccine or as vaccine adjuvant.

[0137] Vaccines of the invention may be utilized to treat conditions or diseases in many therapeutic areas such as, but not limited to, cardiovascular, CNS, dermatology, endocrinology, oncology, immunology and autoimmunity, inflammation, respiratory, and anti-infective.

ALB Compounds

[0138] In some embodiments, the present invention provides a modulator of a polypeptide that binds cGAMP having a structure comprising the following features:

A ----- L ----- B

wherein:

A is or comprises a moiety that fits in the cGAS adenosine binding site;

B is or comprises a moiety that fits in the cGAS guanosine binding site; and

optionally, L is a linker moiety linking A and B in a manner which allows A and B to adopt appropriate interactions to bind cGAS.

[0139] In some embodiments, the polypeptide that binds cGAMP is cGAS. In some embodiments, the polypeptide that binds cGAMP is STING.

[0140] In some embodiments, A is Ring A as defined below and described in classes and subclasses herein, both singly and in combination. In some embodiments, A optionally makes one or more interactions with cGAS at one or more sites selected from the group consisting of Ser199, Ser420, Lys402, Glu211, Asp213, Asp307, Tyr421, Arg364, and combinations thereof. In some embodiments, A optionally makes one or more interactions with cGAS at one or more sites selected from the group consisting of Tyr421, Asp213, Asp307, Arg364, and combinations thereof.

[0141] In some embodiments, B is Ring B as defined below and described in classes and subclasses herein, both singly and in combination. In some embodiments, B optionally makes one or more interactions with cGAS at one or more sites selected from the group consisting of Tyr421, Thr197, Ser366, Ser368, Arg364, and combinations thereof.

[0142] In some embodiments, a linker moiety is a linker suitable to covalently link A and B and which allows A and B to adopt appropriate interactions to bind cGAS. In some embodiments, a linker together with A and/or B comprises a nucleoside optionally containing one or more phosphate groups. In some embodiments, a linker together with A and B comprises

a cyclic dinucleoside optionally containing one or more phosphate groups. In some embodiments, a modulator is a cyclic-GMP-AMP analog.

[0143] In some embodiments, a linker moiety comprises one or more ribose or phosphate groups. In some embodiments, such ribose and phosphate groups, along with Ring A or B, form a ribonucletide. In some embodiments, a modulator comprises one or more modified ribonucleotides. Modified ribonucleotides are well known in the art, and include modifications to a phosphate group, ribose group, nucleotide base group, and combinations thereof. The present invention contemplates all possible modified ribonucleotides for modulators and compound described herein. In some embodiments, these modifications enhance compound stability *in vivo*. In some embodiments, modifications increase compound resilience to phosphodiesterases.

[0144] In some embodiments, a linker comprises a modified phosphodiester group. Such modifications are known in the art and include, without limitation, substituting phosphodiesters with phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkyl-phosphotriesters, methyl and other alkyl phosphonates including 3′-alkylene phosphonates, 5′-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3′-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates, boranophosphates, and combinations thereof. In some embodiments, a phosphodiester is modified to a phosphoramidates. Suitable phosphoramidates include, without limitation, those listed available at www.glenresearch.com/Reference/StructureListing.php, the entire contents of which are hereby incorporated by reference.

[0145] In some embodiments, a linker does not include phosphorus. In some embodiments, a linker comprises a short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane moieties; sulfide, sulfoxide and sulfone moieties; amides, carboxylates, formacetyl and thioformacetyl moieties; methylene formacetyl and thioformacetyl moieties; riboacetyl moieties; alkene containing moieties; sulfamate moieties; methyleneimino and methylenehydrazino moieties; sulfonate and

sulfonamide moieties; amide moieties; and others having mixed N, O, S, and CH_2 component parts.

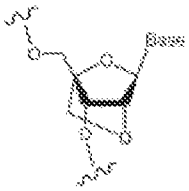
[0146] In addition, a phosphodiester linker may be modified to improve the stability of the compound. For example, in certain instances the P=O linkage is changed to a P=S linkage which is not as susceptible to degradation by nucleases in vivo. In certain instances, the C-2 hydroxyl group of the sugar moiety of a nucleotide is converted to an alkyl or heteroalky ether. This modification renders the oligonucleotide less prone to nucleolytic degradation.

[0147] Additional phosphodiester modification are described by Dellinger et al. Curr Protoc Nucleic Acid Chem. 2004 Oct; Chapter 4:Unit 4; Marshall et al. Science. 1993 Mar 12;259(5101):1564-70, the entire contents of which are hereby incorporated by reference.

[0148] A linker moiety may also comprise one or more modified ribose moieties. In some embodiments, a linker comprises a ribose modified at one of the following at the 2′ or 3’ position: OH; F; O-, S-, or N-alkyl; O-, S-, or N-alkenyl; O-, S- or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl, alkenyl and alkynyl may be substituted or unsubstituted C_1 to C_{10} alkyl or C_2 to C_{10} alkenyl and alkynyl. In some embodiments, 2’ or 3’ modifications include: 2’-O-Me, 2’-O-MOE, 2’-O-allyl, 2’-O-dinitrophosphate, 2’-fluoro, 2’-thio, 2’-aminoethyl, 2’-guanidinopropyl. In some embodiments, 2’ or 3’ modifications include: $O[(CH_2)_nO]_mCH_3$, $O(CH_2)_nOCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nONH_2$, and $O(CH_2)_nON[(CH_2)CH_3]_2$, where n and m are from 1 to about 10. In some embodiments, a linker comprises a ribose modified at the 2′ or 3’ position with: C_1 to C_{10} lower alkyl, substituted lower alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH_3, OCN, Cl, Br, CN, CF_3, OCF_3, $SOCH_3$, SO_2CH_3, ONO_2, NO_2, N_3, NH_2, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA-cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. In some embodiments, a modification includes 2′-O-methoxyethyl (2′-O—$CH_2CH_2OCH_3$, also known as 2′-O-(2-methoxyethyl) or 2′-methoxyethoxy or 2′-MOE) (Martin et al., *Helv. Chim. Acta,* 1995, 78, 486-504) i.e., an alkoxyalkoxy group. A further preferred modification includes 2′-dimethylaminooxyethoxy, i.e., a $O(CH_2)_2ON(CH_3)_2$ group, also known as 2′-DMAOE, and 2′-dimethylamino-ethoxyethoxy (also known in the art as 2′-O-dimethyl-

amino-ethoxy-ethyl or 2′-DMAEOE), i.e., 2′-O—CH_2—O—CH_2—$N(CH_3)_2$. Other modifications to a linker ribose include 2′-methoxy (2′-O—CH_3), 2′-aminopropoxy (2′-$OCH_2CH_2CH_2NH_2$), 2′-allyl (2′-CH_2—CH=CH_2), 2′-O-allyl (2′-O—CH_2—CH=CH_2) and 2′-fluoro (2′-F). The 2′-modification may be in the arabino (up) position or ribo (down) position. In some embodiments, a 2′-arabino modification is 2′-F.

[0149] Similar modifications may also be made at other positions on the oligonucleotide, particularly the 3′ position of the sugar on the 3′ terminal nucleotide and the 5′ position. Oligonucleotides may also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar.

[0150] In some embodiments, a linker ribose is modified to a Locked Nucleic Acids (LNAs) in which the 2′-hydroxyl group is linked to the 3′ or 4′ carbon atom of the sugar ring, thereby forming a bicyclic sugar moiety. The linkage is preferably a methylene (—CH_2—) n group bridging the 2′ oxygen atom and the 4′ carbon atom wherein n is 1 or 2. LNAs and preparation thereof are described in International Patent Publication Nos. WO 98/39352 and WO 99/14226, the entire contents of which are hereby incorporated by reference. In some embodiments, the LNA forms a moiety:  .

[0151] In certain embodiments, a linker comprises a hexose moiety. In some embodiments, the hexose is glucose or mannose. In certain instances, the ribose sugar moiety is replaced with a cyclohexenyl group or polycyclic heteroalkyl ring. In some embodiments, the ribose sugar moiety is replaced with morpholino group. Additional ribose modification are discussed by Engels, New Biotechnology, Vol. 30, 3, p. 302 (2013), the entire contents of which are hereby incorporated by reference.

[0152] In some embodiments A or B is an unmodified or natural nucleobase selected from adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). In some embodiments, A or B is a modified nucleobase. Modified nucleobases are known in the art and include, without limitation, synthetic and natural nucleobases such as 5-

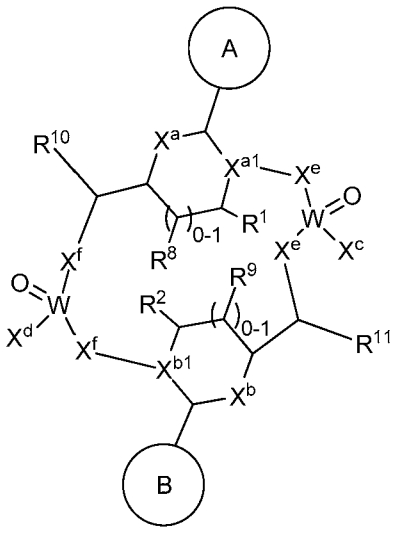
methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl (—C≡C—CH3) uracil and cytosine and other alkynyl derivatives of pyrimidine bases, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 2-F-adenine, 2-amino-adenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Further modified nucleobases include tricyclic pyrimidines such as phenoxazine cytidine(1H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), phenothiazine cytidine (1H-pyrimido[5,4-b][1,4]benzothiazin-2(3H)-one), G-clamps such as a substituted phenoxazine cytidine (e.g. 9-(2-aminoethoxy)-H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), carbazole cytidine (2H-pyrimido[4,5-b]indol-2-one), pyridoindole cytidine (H-pyrido[3′,2′:4,5]pyrrolo[2,3-d]pyrimidin-2-one). Modified nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deazaadenine, 7-deazaguanosine, 2-aminopyridine and 2-pyridone. Further nucleobases include those disclosed in U.S. Pat. No. 3,687,808, those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990, those disclosed by Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613, and those disclosed by Sanghvi, Y. S., Chapter 15, Antisense Research and Applications, pages 289-302, Crooke, S. T. and Lebleu, B., ed., CRC Press, 1993. In some embodiments, a non-natural nucleobase is difluorotolyl, nitropyrrolyl, or nitroimidazolyl. In certain embodiments, a non-natural nucleobase is 7-deazaadenine, 3-deazaadenine, N1-methyl-guanosine, 6-thioguanosine, 2-pyrimidinone, 4-thiouridine, 2-pyridinone, 5-propynyl-uridine, imidazole-4-carboxamide, 5-nitroindol, 3-nitropyrrole, 2-aminopurine, 5-methyl-2-pyrimidinone, N3-thioethylthymidine, 6-thiopurine, 5-iodouridine, 8-azidoadenosine, 5-mercaptouridine, or those derived from 5-bromouracil, diaminopurine, 2-thiouracil, 4-thiouracil, pseudouracil, difluorotoluene, and dihydrouracil. Additional modified nucleobases include those found in www.glenresearch.com/Reference/StructureListing.php and www.thermoscientificbio.com/rna-

pricing-and-modifications/, the entire contents of each of which are hereby incorporated by reference. Additional modifications are discussed by Verma et al. Annu. Rev. Biochem. 1998. 67:99–134.

[0153] In some embodiments, a linker moiety comprises a group that replaces both a phosphodiester and ribose groups of a ribonucleotide. One such linker is referred to as a peptide nucleic acid (PNA). In PNA compounds, the usual sugar-backbone of an oligonucleotide is replaced with an amide containing backbone, for example an aminoethylglycine backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone. Representative United States patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Pat. Nos. 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference. Further teaching of PNA compounds can be found in Nielsen et al., *Science,* 1991, 254, 1497-1500; Nielsen et al., Chem. Soc. Rev., 1997,26, 73-78; Shakeel et al., Journal of Chemical Technology & Biotechnology, Volume 81, Number 6, June 2006 , pp. 892-899(8); Nielsen, CHEMISTRY & BIODIVERSITY – Vol. 7 (2010), p. 786.

[0154] These and other suitable linkers are discussed in U.S. Pat. Nos. 7, 365,058, 8,101,348, 8,088,902, 7,579,451,7,582,744, 8,334,373, 8,017,762, 7,919,612, 7,812,149, and 7,723,508, the entire contents of each of which are hereby incorporated by reference herein.

[0155] In some embodiments, the present invention provides a compound of formula I:

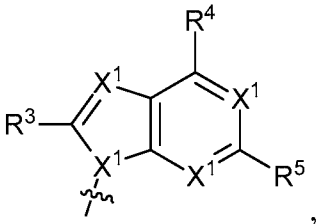
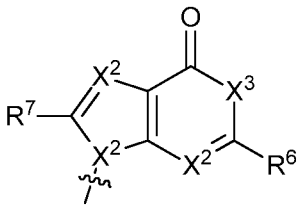
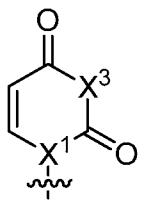
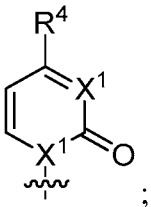


I

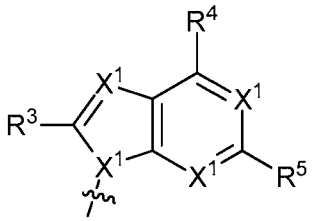
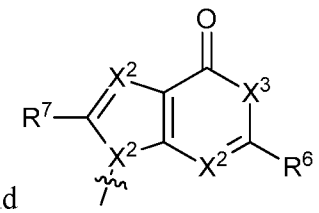
or a pharmaceutically acceptable salt thereof,

wherein:

Ring A is selected from the group consisting of:


,

,

, and

;

Ring B is selected from the group consisting of:


and

;

each X^1 and X^2 is independently –CR- or –N-;

X^3 is $–C(R)_2$-, -O-, or –NR-;

X^a and X^b are independently $–C(R)_2$-, -C(R)=C(R)-, -O-, -S-, -S(O)-, $-S(O)_2$-, or -N(R)-;

X^{a1} and X^{b1} are independently –C(R)- or –N-;

X^c and X^d, when present, are independently optionally substituted oxygen, optionally substituted sulfur, a substituted nitrogen atom, BH_3, or optionally substituted C_{1-12} aliphatic;

each X^e and X^f is independently –O-, -S-, or -N(R)-;

each W is independently P or S;

each R^1 and R^2 is independently selected from the group consisting of hydrogen, halogen, $–NO_2$, –CN, $–OR^a$, –SR, $–N(R)_2$, -C(O)R, $-CO_2R$, -C(O)C(O)R, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, -OC(O)R, -N(R)C(O)R, $–N(R)N(R)_2$, -

$N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, and optionally substituted C_{1-12} aliphatic or C_{1-4} alkoxy-C_{1-4} alkyl;

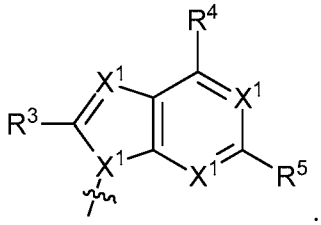
each R^3, R^4, R^5, R^6, R^7, R^{10}, and R^{11} is independently selected from the group consisting of hydrogen, halogen, $–NO_2$, –CN, –OR, –SR, $–N(R)_2$, $-C(O)R$, $-CO_2R$, $-C(O)C(O)R$, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, $-OC(O)R$, $-N(R)C(O)R$, $–N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, or an optionally substituted group selected from C_{1-12} aliphatic, phenyl, a 3-7 membered saturated or partially unsaturated monocyclic carbocyclic ring, a 7-10 membered saturated or partially unsaturated bicyclic carbocyclic ring, a 3-7 membered saturated or partially unsaturated heterocyclic ring having 1-2 heteroatoms independently selected from nitrogen, oxygen, or sulfur, a 7-10 membered saturated or partially unsaturated bicyclic heterocyclic ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur, and a 5-6 membered heteroaryl ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur;

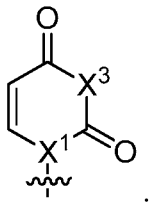
each R^8 and R^9, when present, is independently selected from the group consisting of hydrogen, halogen, $–NO_2$, –CN, $–OR^a$, –SR, $–N(R)_2$, $-C(O)R$, $-CO_2R$, $-C(O)C(O)R$, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, $-OC(O)R$, $-N(R)C(O)R$, $–N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, and an optionally substituted C_{1-12} aliphatic;

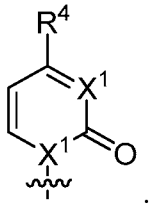
each R is independently selected from the group consisting of hydrogen or an optionally substituted group selected from C_{1-6} aliphatic, phenyl, a 3-7 membered saturated or partially unsaturated carbocyclic ring, a 3-7 membered saturated or partially unsaturated monocyclic heterocyclic ring having 1-2 heteroatoms independently selected from nitrogen, oxygen, or sulfur, and a 5-6 membered heteroaryl ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur; or:

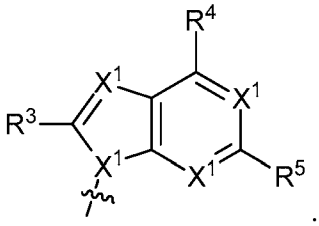
two R groups on the same nitrogen are taken together with their intervening atoms to form an optionally substituted 3-7 membered saturated, partially unsaturated, or heteroaryl ring having 1-4 heteroatoms independently selected from nitrogen, oxygen, or sulfur; and

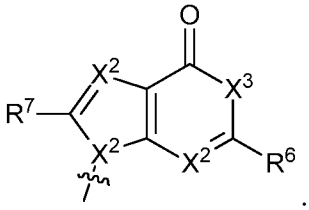
R^a is an oxygen protecting group or R.

[0156] In some embodiments, Ring A is .

[0157] In some embodiments, Ring A is .

[0158] In some embodiments, Ring A is .

[0159] In some embodiments, Ring B is .

[0160] In some embodiments, Ring B is .

[0161] In some embodiments, X^1 is –CR-. In some embodiments, X^1 is –N-. In some embodiments, X^2 is –CR-. In some embodiments, X^2 is –N-. In some embodiments, X^3 is –$C(R)_2$-. In some embodiments, X^3 is –O-. In some embodiments, X^3 is –NR-.

[0162] In certain embodiments, X^a is –$C(R)_2$-. In certain embodiments, X^a is -C(R)=C(R)-. In certain embodiments, X^a is –O-. In certain embodiments, X^a is –S-. In certain embodiments, X^a is –S(O)-.

[0163] In certain embodiments, X^a is –$S(O)_2$-. In certain embodiments, X^a is –NR-.

[0164] In certain embodiments, X^b is $-C(R)_2-$. In certain embodiments, X^b is -C(R)=C(R)-. In certain embodiments, X^b is –O-. In certain embodiments, X^b is –S-. In certain embodiments, X^b is –S(O)-. In certain embodiments, X^b is $-S(O)_2-$. In certain embodiments, X^b is –NR-.

[0165] In certain embodiments, X^{a1} is –C(R)-. In certain embodiments, X^{a1} is –N-. In certain embodiments, X^{b1} is –C(R)-. In certain embodiments, X^{b1} is –N-.

[0166] In some embodiments, X^c is oxygen. In some embodiments, X^c is sulfur. It will be appreciated that in certain embodiments where X^c is oxygen or sulfur, the oxygen or sulfur atom may possess a formal negative charge. In some embodiments, X^c is a substituted nitrogen atom. In some embodiments, the nitrogen is independently substituted with hydrogen or optionally substituted C_{1-12} aliphatic groups. In some embodiments, X^c is optionally substituted C_{1-12} aliphatic.

[0167] In some embodiments, X^d is oxygen. In some embodiments, X^d is sulfur. It will be appreciated that in certain embodiments where X^d is oxygen or sulfur, the oxygen or sulfur atom may possess a formal negative charge. In some embodiments, X^d is a substituted nitrogen atom. In some embodiments, the nitrogen is independently substituted with hydrogen or optionally substituted C_{1-12} aliphatic groups. In some embodiments, X^d is optionally substituted C_{1-12} aliphatic.

[0168] In some embodiments, X^e is –O-. In some embodiments, X^e is –S-. In some embodiments, X^e is –N(R)-.

[0169] In some embodiments, X^f is –O-. In some embodiments, X^f is –S-. In some embodiments, X^f is –N(R)-.

[0170] In some embodiments, W is P. In other embodiments, W is S.

[0171] In some embodiments, R^1 is hydrogen, halogen, $-OR^a$, -SR, $-N(R)_2$, and optionally substituted C_{1-12} aliphatic or C_{1-4} alkoxy-C_{1-4} alkyl. In some embodiments, R^1 is hydrogen. In some embodiments, R^1 is halogen. In some embodiments, R^1 is $-OR^a$. In some embodiments, R^1 is -OH. In some embodiments, R^1 is fluro. In some embodiments, R^1 is C_{1-12} aliphatic. In some embodiments, R^1 is C_{1-6} aliphatic. In some embodiments, R^1 is C_{1-3} aliphatic.

In some embodiments, R^1 is methyl. In some embodiments, R^1 is C_{1-4} alkoxy-C_{1-4} alkyl. In some embodiments, R^1 is methoxy-ethyl.

[0172] In some embodiments, R^2 is hydrogen, halogen, $-OR^a$, -SR, $-N(R)_2$, and optionally substituted C_{1-12} aliphatic or C_{1-4} alkoxy-C_{1-4} alkyl. In some embodiments, R^2 is hydrogen. In some embodiments, R^2 is halogen. In some embodiments, R^2 is $-OR^a$. In some embodiments, R^2 is -OH. In some embodiments, R^2 is fluro. In some embodiments, R^2 is C_{1-12} aliphatic. In some embodiments, R^2 is C_{1-6} aliphatic. In some embodiments, R^2 is C_{1-3} aliphatic. In some embodiments, R^2 is methyl. In some embodiments, R^2 is C_{1-4} alkoxy-C_{1-4} alkyl. In some embodiments, R^2 is methoxy-ethyl.

[0173] In some embodiments, R^3 is hydrogen, halogen, $-NO_2$, –CN, –OR, –SR, $-N(R)_2$, -C(O)R, $-CO_2R$, -C(O)C(O)R, $-C(O)CH_2C(O)R$, –S(O)R, $-S(O)_2R$, $-C(O)N(R)_2$, $-SO_2N(R)_2$, -OC(O)R, -N(R)C(O)R, $-N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, -C(=NR)$N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $-N(R)SO_2N(R)_2$, $-N(R)SO_2R$, $-OC(O)N(R)_2$, or optionally substituted C_{1-12} aliphatic. In certain embodiments, R^3 is hydrogen. In some embodiments, R^3 is halogen. In certain embodiments, R^3 is $-NO_2$. In some embodiments, R^3 is –CN. In certain embodiments, R^3 is –OR. In some embodiments, R^3 is C_{1-12} aliphatic. In some embodiments, R^3 is C_{1-6} aliphatic.

[0174] In some embodiments, R^4 is hydrogen, halogen, $-NO_2$, –CN, –OR, –SR, $-N(R)_2$, -C(O)R, $-CO_2R$, -C(O)C(O)R, $-C(O)CH_2C(O)R$, –S(O)R, $-S(O)_2R$, $-C(O)N(R)_2$, $-SO_2N(R)_2$, -OC(O)R, -N(R)C(O)R, $-N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, -C(=NR)$N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $-N(R)SO_2N(R)_2$, $-N(R)SO_2R$, $-OC(O)N(R)_2$, or optionally substituted C_{1-12} aliphatic. In certain embodiments, R^4 is hydrogen. In some embodiments, R^4 is halogen. In certain embodiments, R^4 is $-NO_2$. In some embodiments, R^4 is –CN. In certain embodiments, R^4 is –OR. In some embodiments, R^4 is C_{1-12} aliphatic. In some embodiments, R^4 is C_{1-6} aliphatic.

[0175] In some embodiments, R^5 is hydrogen, halogen, $-NO_2$, –CN, –OR, –SR, $-N(R)_2$, -C(O)R, $-CO_2R$, -C(O)C(O)R, $-C(O)CH_2C(O)R$, –S(O)R, $-S(O)_2R$, $-C(O)N(R)_2$, $-SO_2N(R)_2$, -OC(O)R, -N(R)C(O)R, $-N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, -C(=NR)$N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $-N(R)SO_2N(R)_2$, –

$N(R)SO_2R$, $-OC(O)N(R)_2$, or optionally substituted C_{1-12} aliphatic. In certain embodiments, R^5 is hydrogen. In some embodiments, R^5 is halogen. In certain embodiments, R^5 is $-NO_2$. In some embodiments, R^5 is –CN. In certain embodiments, R^5 is –OR. In some embodiments, R^5 is C_{1-12} aliphatic. In some embodiments, R^5 is C_{1-6} aliphatic.

[0176] In some embodiments, R^6 is hydrogen, halogen, $-NO_2$, –CN, –OR, –SR, $-N(R)_2$, $-C(O)R$, $-CO_2R$, $-C(O)C(O)R$, $-C(O)CH_2C(O)R$, $-S(O)R$, $-S(O)_2R$, $-C(O)N(R)_2$, $-SO_2N(R)_2$, $-OC(O)R$, $-N(R)C(O)R$, $-N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $-N(R)SO_2N(R)_2$, $-N(R)SO_2R$, $-OC(O)N(R)_2$, or optionally substituted C_{1-12} aliphatic. In certain embodiments, R^6 is hydrogen. In some embodiments, R^6 is halogen. In certain embodiments, R^6 is $-NO_2$. In some embodiments, R^6 is –CN. In certain embodiments, R^6 is –OR. In some embodiments, R^6 is C_{1-12} aliphatic. In some embodiments, R^6 is C_{1-6} aliphatic.

[0177] In some embodiments, R^7 is hydrogen, halogen, $-NO_2$, –CN, –OR, –SR, $-N(R)_2$, $-C(O)R$, $-CO_2R$, $-C(O)C(O)R$, $-C(O)CH_2C(O)R$, $-S(O)R$, $-S(O)_2R$, $-C(O)N(R)_2$, $-SO_2N(R)_2$, $-OC(O)R$, $-N(R)C(O)R$, $-N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $-N(R)SO_2N(R)_2$, $-N(R)SO_2R$, $-OC(O)N(R)_2$, or optionally substituted C_{1-12} aliphatic. In certain embodiments, R^7 is hydrogen. In some embodiments, R^7 is halogen. In certain embodiments, R^7 is $-NO_2$. In some embodiments, R^7 is –CN. In certain embodiments, R^7 is –OR. In some embodiments, R^7 is C_{1-12} aliphatic. In some embodiments, R^7 is C_{1-6} aliphatic.

[0178] In some embodiments, R^8 is present. In other embodiments, R^8 is absent. In some embodiments, R^8 is hydrogen. In some embodiments, R^8 is halogen. In some embodiments, R^8 is $-OR^a$. In some embodiments, R^8 is optionally substituted C_{1-12} aliphatic. In some embodiments, R^8 is C_{1-6} aliphatic. In some embodiments, R^8 is C_{1-3} aliphatic.

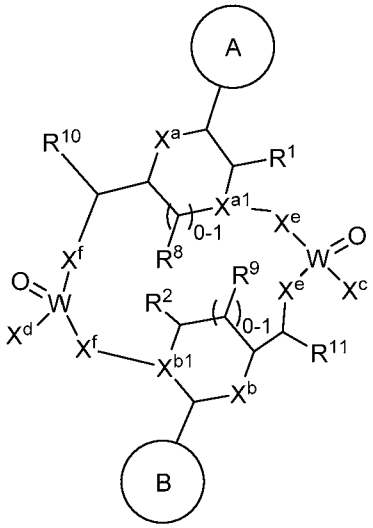
[0179] In some embodiments, R^9 is present. In other embodiments, R^9 is absent. In some embodiments, R^9 is hydrogen. In some embodiments, R^9 is halogen. In some embodiments, R^9 is $-OR^a$. In some embodiments, R^9 is optionally substituted C_{1-12} aliphatic. In some embodiments, R^9 is C_{1-6} aliphatic. In some embodiments, R^9 is C_{1-3} aliphatic.

[0180] In some embodiments, R^{10} is hydrogen, halogen, $–NO_2$, –CN, –OR, –SR, $–N(R)_2$, $-C(O)R$, $-CO_2R$, $-C(O)C(O)R$, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, $-OC(O)R$, $-N(R)C(O)R$, $–N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, or optionally substituted C_{1-12} aliphatic. In some embodiments, R^{10} is optionally substituted phenyl, a 3-7 membered saturated or partially unsaturated monocyclic carbocyclic ring, a 3-7 membered saturated or partially unsaturated heterocyclic ring having 1-2 heteroatoms independently selected from nitrogen, oxygen, or sulfur, or a 5-6 membered heteroaryl ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur In certain embodiments, R^{10} is hydrogen. In some embodiments, R^{10} is halogen. In some embodiments, R^{10} is C_{1-12} aliphatic. In some embodiments, R^{10} is C_{1-6} aliphatic.

[0181] In some embodiments, R^{11} is hydrogen, halogen, $–NO_2$, –CN, –OR, –SR, $–N(R)_2$, $-C(O)R$, $-CO_2R$, $-C(O)C(O)R$, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, $-OC(O)R$, $-N(R)C(O)R$, $–N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, or optionally substituted C_{1-12} aliphatic. In some embodiments, R^{11} is optionally substituted phenyl, a 3-7 membered saturated or partially unsaturated monocyclic carbocyclic ring, a 3-7 membered saturated or partially unsaturated heterocyclic ring having 1-2 heteroatoms independently selected from nitrogen, oxygen, or sulfur, or a 5-6 membered heteroaryl ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur In certain embodiments, R^{11} is hydrogen. In some embodiments, R^{11} is halogen. In some embodiments, R^{11} is C_{1-12} aliphatic. In some embodiments, R^{11} is C_{1-6} aliphatic.

[0182] It will be appreciated that for compounds depicted herein, where negatively charges phosphates are shown, the disclosure contemplates both free and salt forms of such compounds, and tautomers thereof. In some embodiments, a provided compound may have one or more protonated nitrogens that balance the charge of a free phosphate.

[0183] In some embodiments, the present invention provides a compound of formula II:

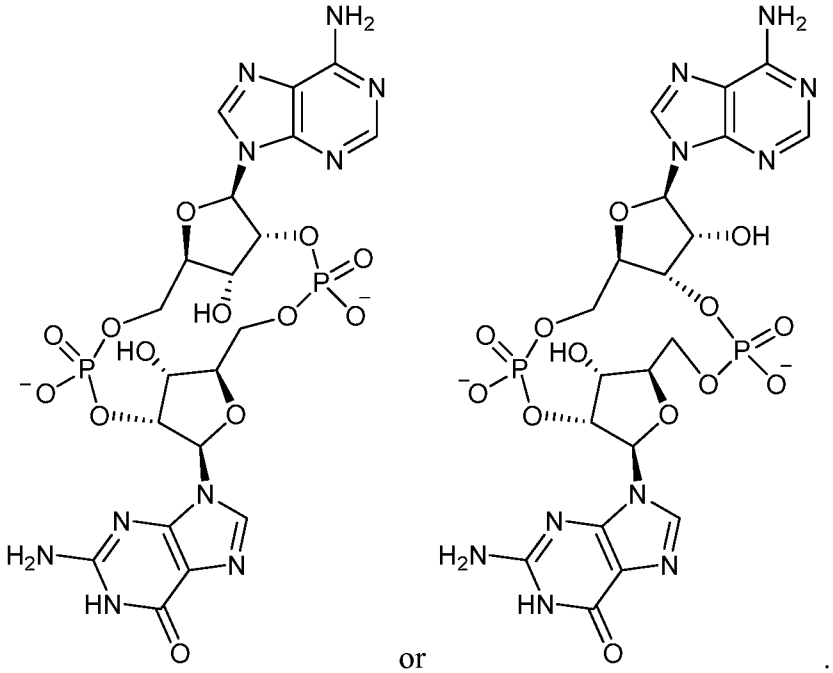


II

or a pharmaceutically acceptable salt thereof,

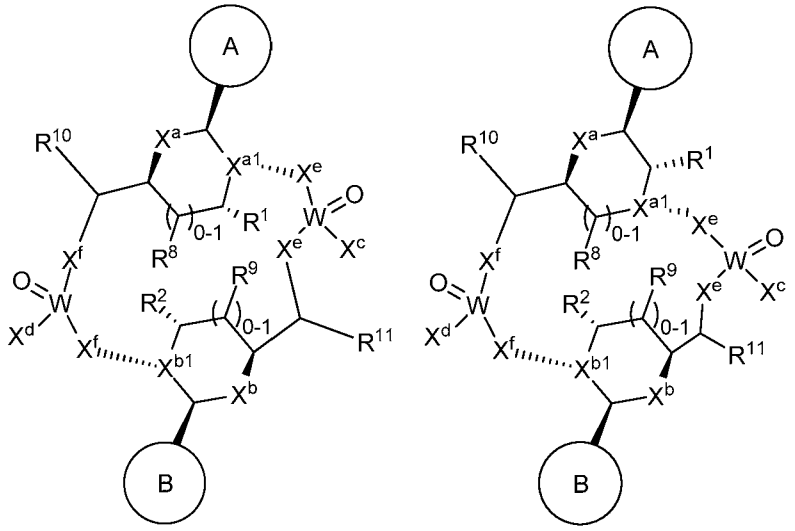
wherein each of Ring A, Ring B, X^a, X^b, X^c, X^d, X^e, X^f, X^{a1}, X^{b1}, X^2, X^3, W, R^1, R^2, R^6, R^7, R^8, R^9, R^{10}, and R^{11} is as defined above and described in classes and subclasses herein, both singly and in combination.

[0184] In some embodiments, a provided compound is other than:



or .

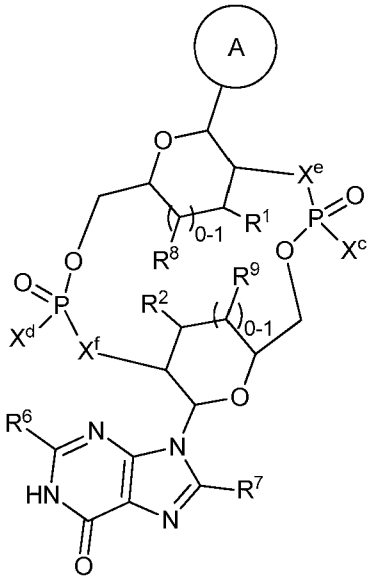
[0185] In some embodiments, a provided compound is of formula I-a or II-a:



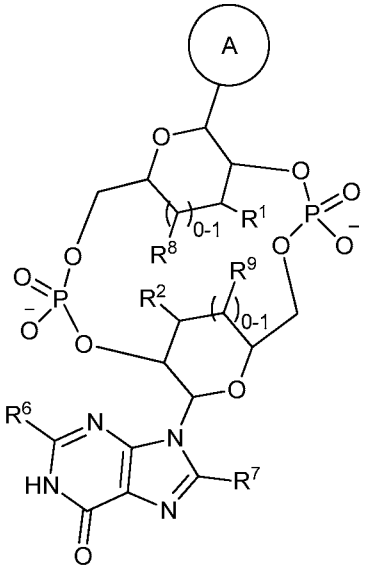
I-a **II-a**

[0186] or a pharmaceutically acceptable salt thereof.

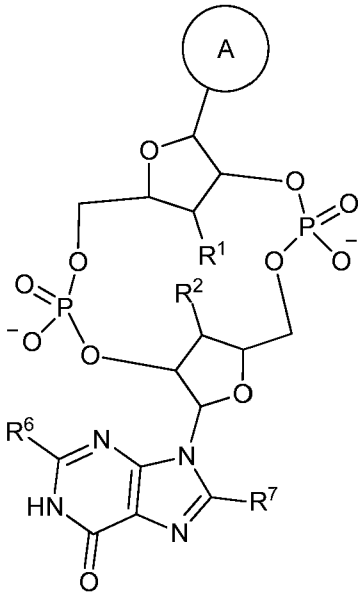
[0187] In some embodiments, a provided compound is of formula III, IV, V, VI, VII, VIII, or IX:



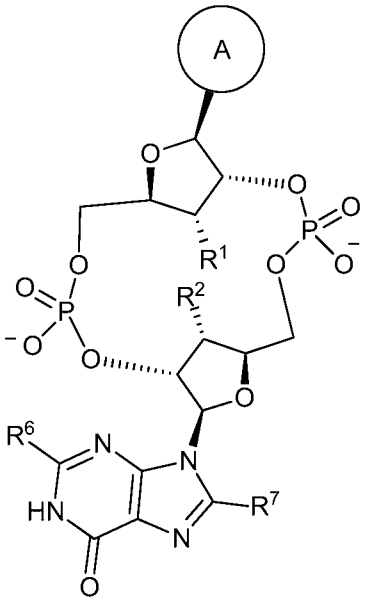
III



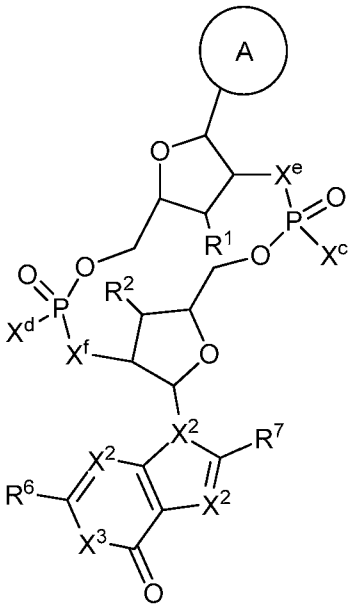
IV



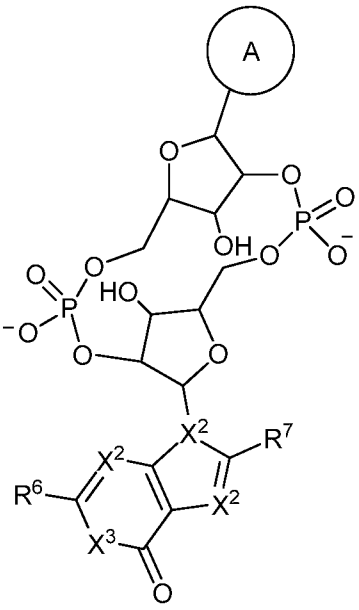
V



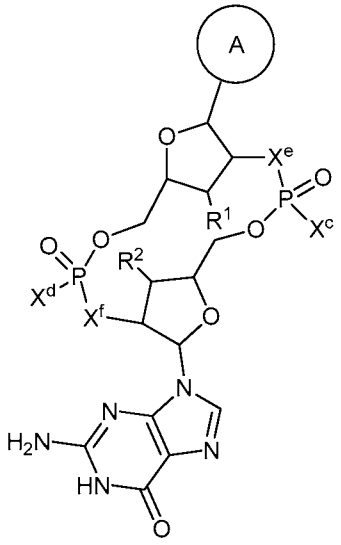
VI



VII



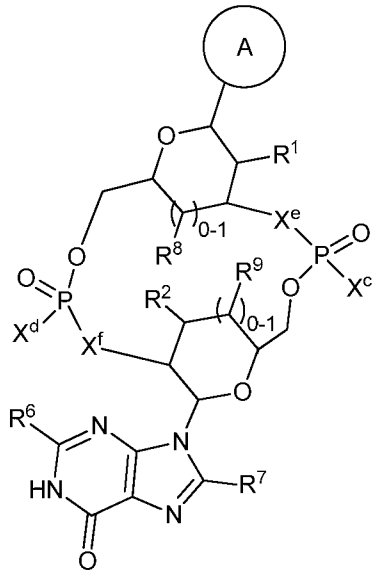
VIII



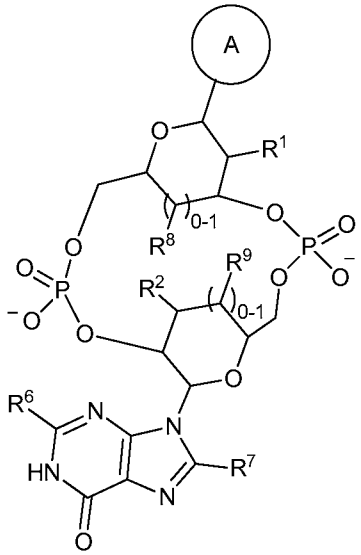
IX

or a pharmaceutically acceptable salt thereof, wherein each of Ring A, X^c, X^d, X^e, X^f, R^1, R^2, R^6, R^7, R^8, R^9, X^2, and X^3 is as defined above and described in classes and subclasses herein, both singly and in combination.

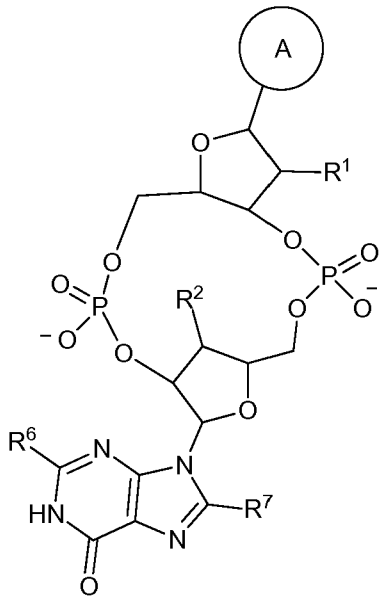
In some embodiments, a provided compound is of formula X, XI, XII, XIII, XIV, XV, or XVI:



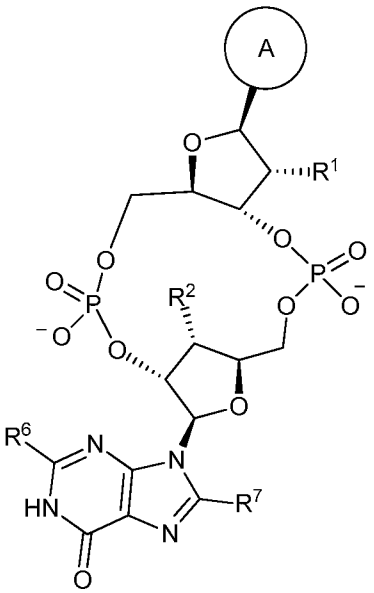
X



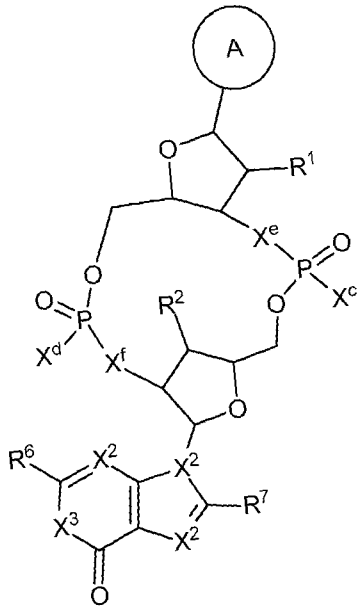
XI



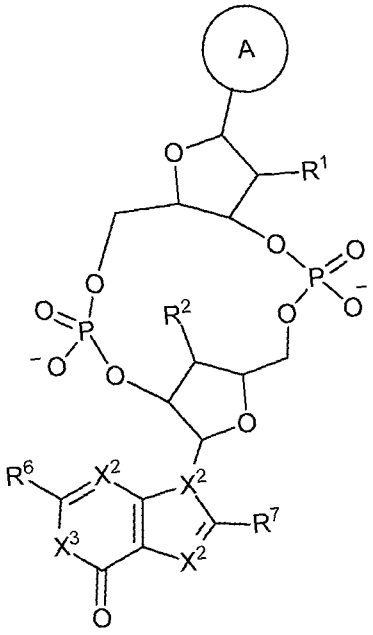
XII



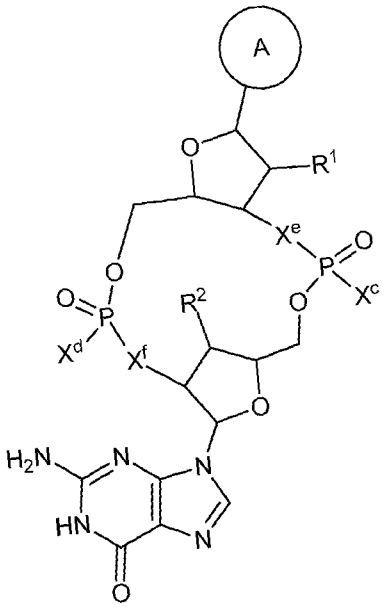
XIII



XIV



XV

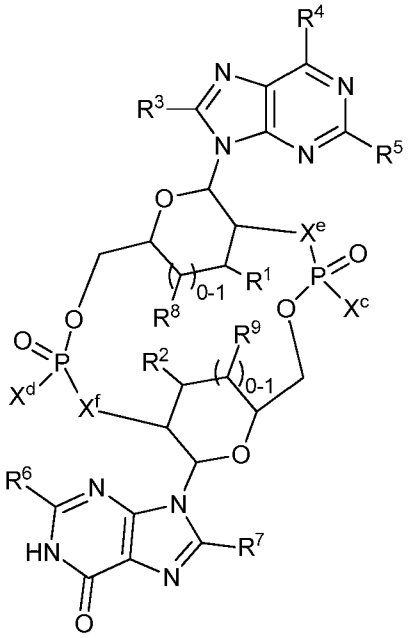


XVI

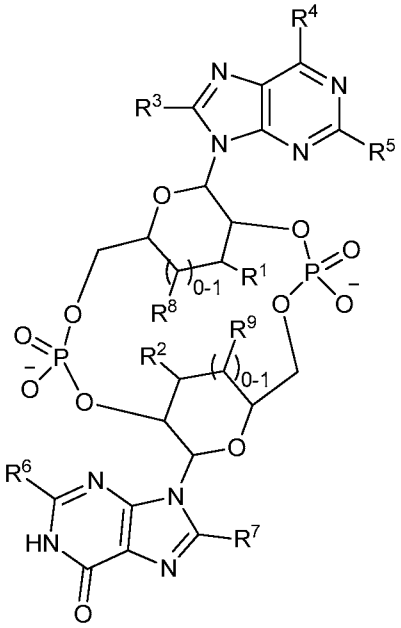
or a pharmaceutically acceptable salt thereof,

wherein each of Ring A, X^c, X^d, X^e, X^f, R^1, R^2, R^6, R^7, R^8, R^9, X^2, and X^3 is as defined above and described in classes and subclasses herein, both singly and in combination.

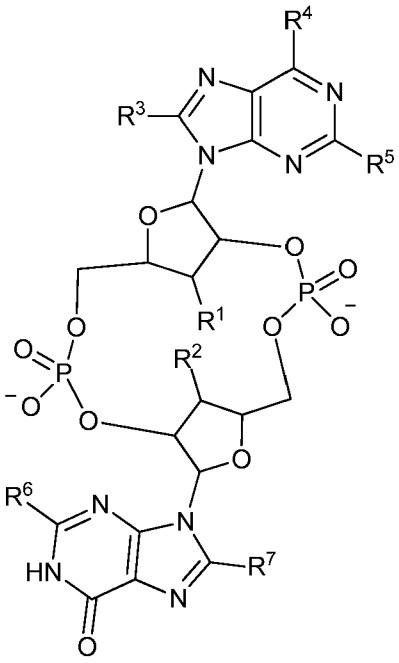
[0188] In some embodiments, a provided compound is selected from:



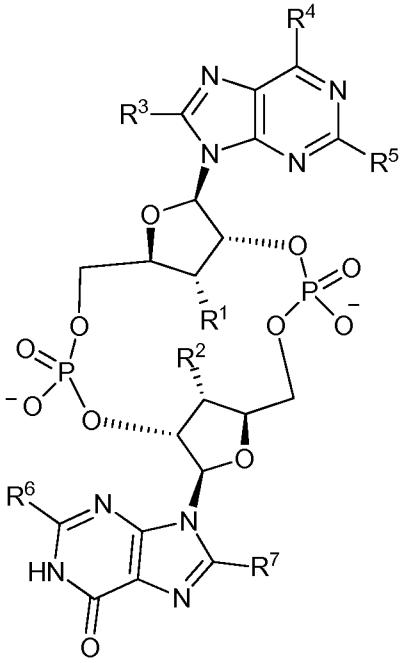
III-a



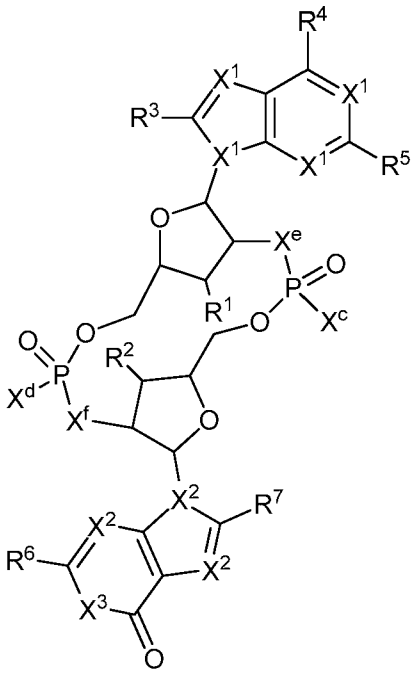
IV-a



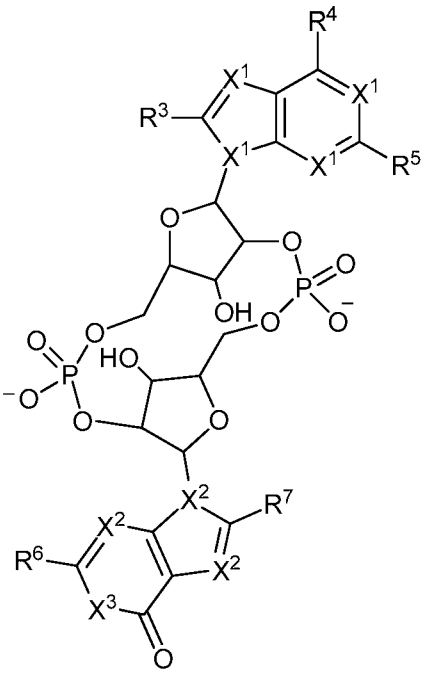
V-a



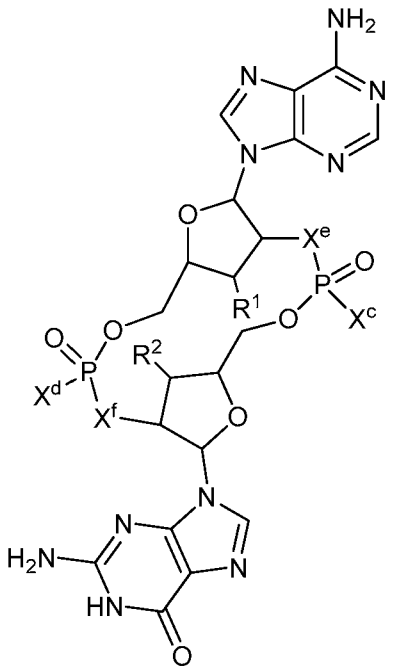
VI-a



VII-a



VIII-a

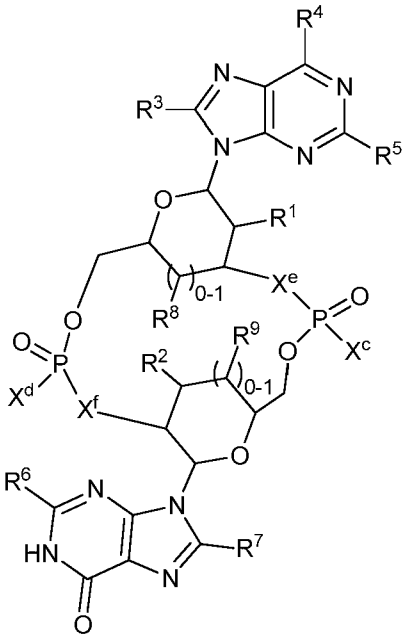


IX-a

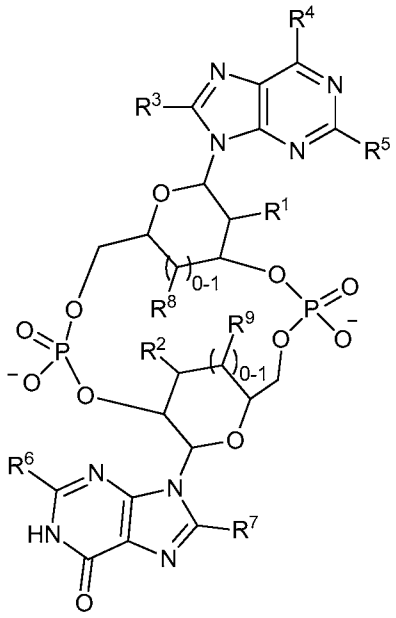
or a pharmaceutically acceptable salt thereof,

wherein each of Ring A, X^c, X^d, X^e, X^f, R^1, R^2, R^6, R^7, R^8, R^9, X^2, and X^3 is as defined above and described in classes and subclasses herein, both singly and in combination.

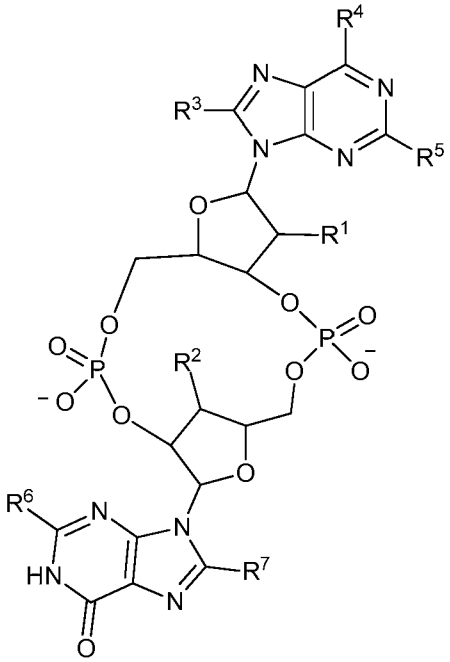
[0189] In some embodiments, a provided compound is selected from:



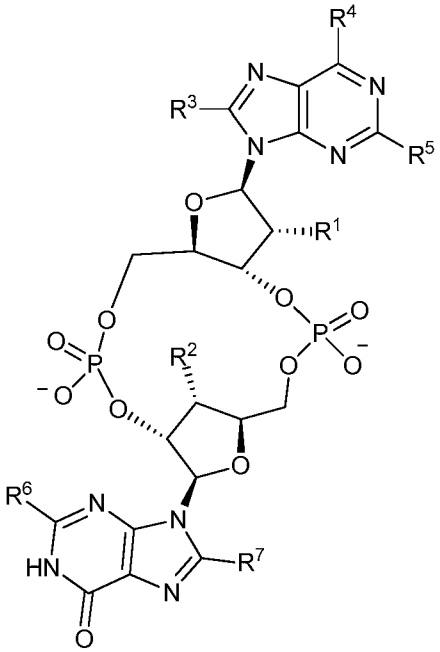
X-a



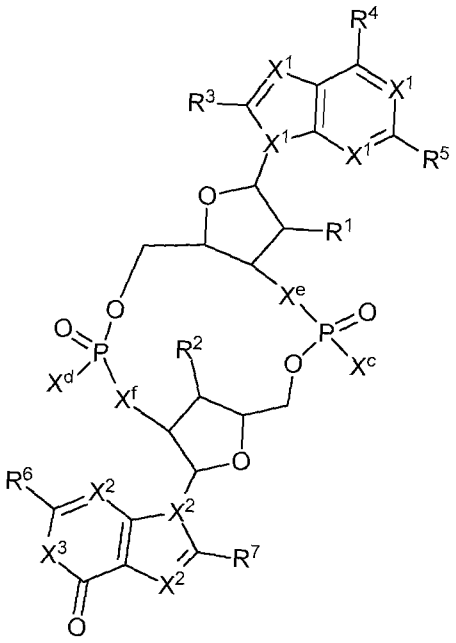
XI-a



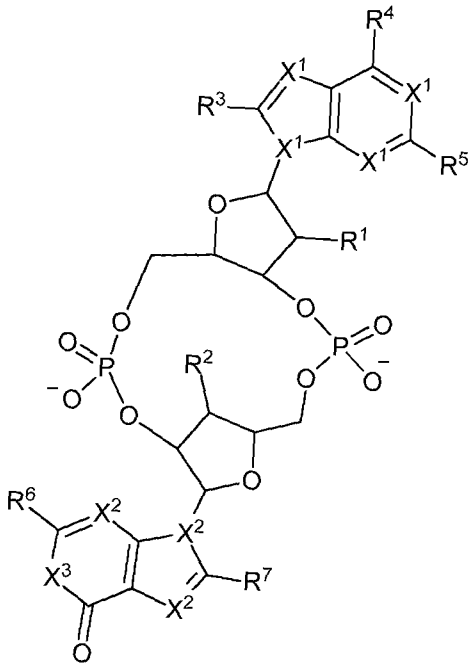
XII-a



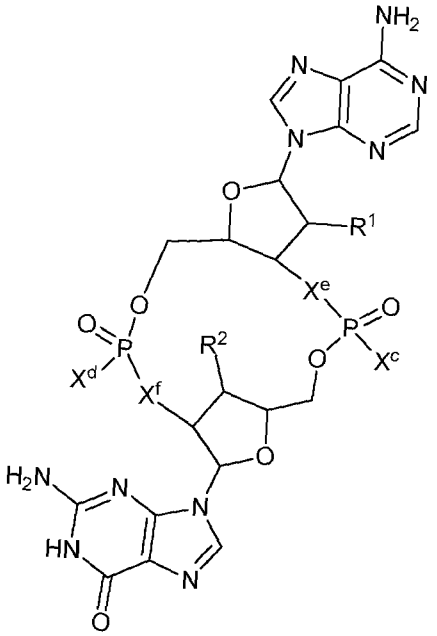
XIII-a



XIV-a



XV-a

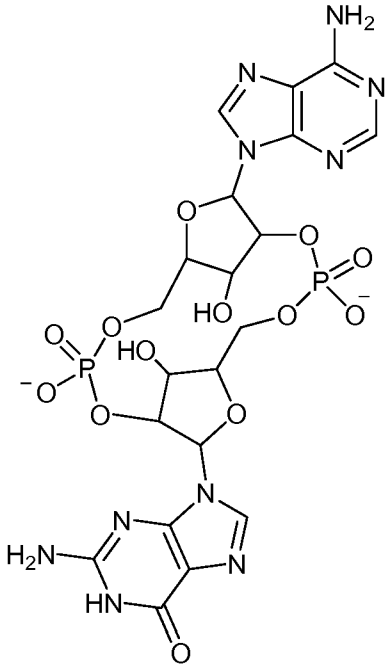
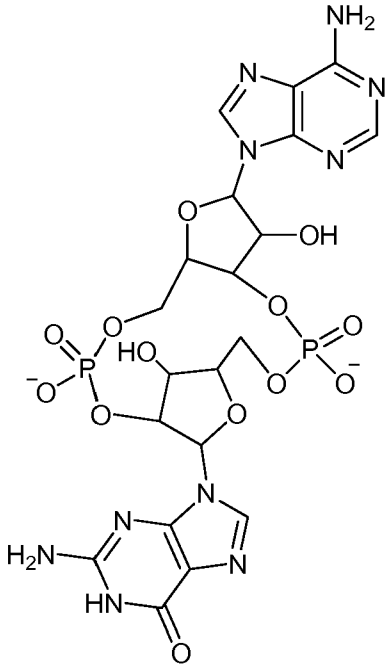
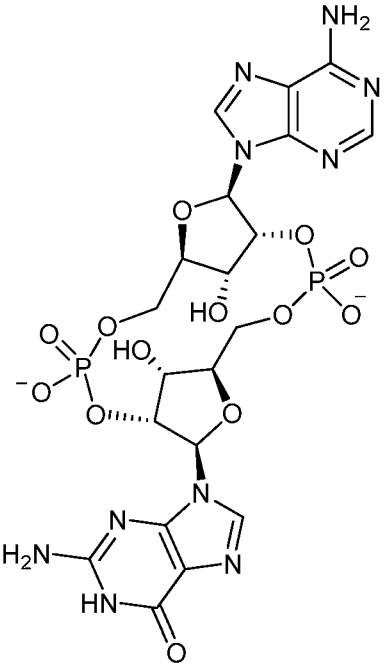
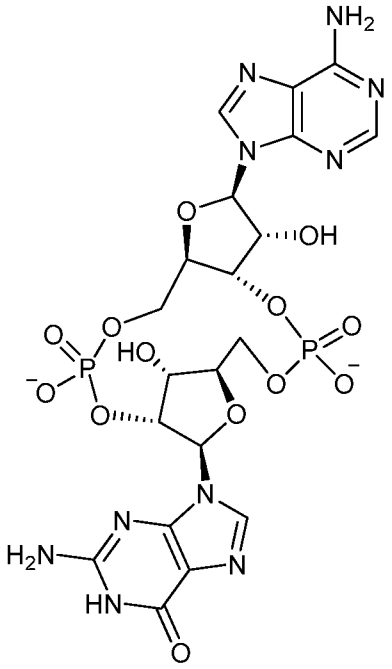


XVI-a

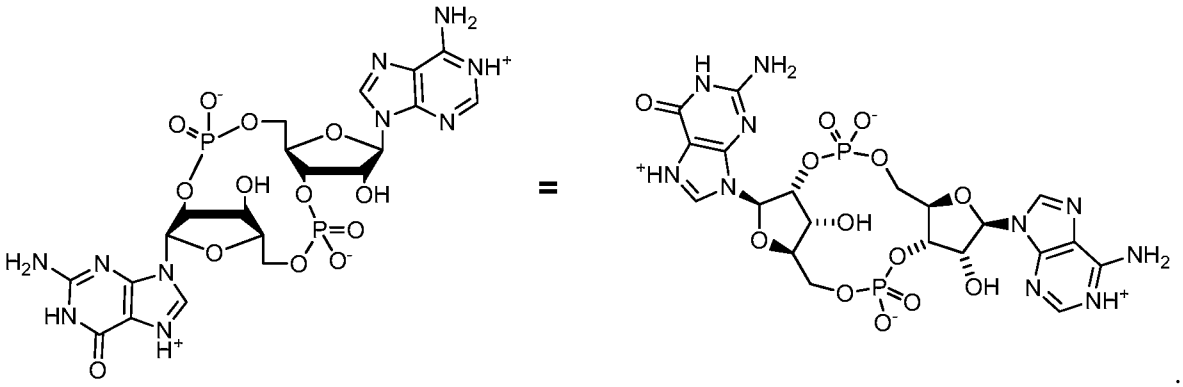
or a pharmaceutically acceptable salt thereof,

wherein each of Ring A, X^c, X^d, X^e, X^f, R^1, R^2, R^6, R^7, R^8, R^9, X^2, and X^3 is as defined above and described in classes and subclasses herein, both singly and in combination.

[0190] In some embodiments, a provided compound is selected from:


,

,

, or

, or a pharmaceutically acceptable salt thereof.

[0191] It will be appreciated that the compounds depicted in the immediately preceding paragraph may be drawn using other conventions. For example, the following two compounds are considered equivalent in chemical structure and stereochemistry:



[0192] In some embodiments, provided compounds are in isolated form. In some embodiments, provided compounds are pure.

Pharmaceutical Compositions

[0193] Provided pharmaceutical compositions can be in a variety of forms including oral dosage forms, topic creams, topical patches, iontophoresis forms, suppository, nasal spray and inhaler, eye drops, intraocular injection forms, depot forms, as well as injectable and infusible solutions. Methods for preparing pharmaceutical composition are well known in the art.

[0194] Pharmaceutical compositions typically contain the active agent described herein in an amount effective to achieve the desired therapeutic effect while avoiding or minimizing adverse side effects. Pharmaceutically acceptable preparations and salts of the active agent are provided herein and are well known in the art. For the administration of cGAS modulators and the like, the amount administered desirably is chosen that is therapeutically effective with few to no adverse side effects. The amount of the therapeutic or pharmaceutical composition which is effective in the treatment of a particular disease, disorder or condition depends on the nature and severity of the disease, the target site of action, the subject's weight, special diets being followed by the subject, concurrent medications being used, the administration route and other factors that are recognized by those skilled in the art. The dosage can be adapted by the clinician in

accordance with conventional factors such as the extent of the disease and different parameters from the subject. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems (e.g., as described by the U.S. Department of Health and Human Services, Food and Drug Administration, and Center for Drug Evaluation and Research in "Guidance for Industry: Estimating Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers", *Pharmacology and Toxicology*, July 2005, the entire contents of which are incorporated herein by reference).

[0195] Various delivery systems are known and can be used to administer active agent described herein or a pharmaceutical composition comprising the same.

[0196] The pharmaceutical compositions described herein can be administered by any suitable route including, but are not limited to enteral, gastroenteral, epidural, oral, transdermal, epidural (peridural), intracerebral (into the cerebrum), intracerebroventricular (into the cerebral ventricles), epicutaneous (application onto the skin), intradermal, (into the skin itself), subcutaneous (under the skin), nasal administration (through the nose), intravenous (into a vein), intraarterial (into an artery), intramuscular (into a muscle), intracardiac (into the heart), intraosseous infusion (into the bone marrow), intrathecal (into the spinal canal), intraperitoneal, (infusion or injection into the peritoneum), intravesical infusion, intravitreal, (through the eye), intracavernous injection, (into the base of the penis), intravaginal administration, intrauterine, extra-amniotic administration, transdermal (diffusion through the intact skin for systemic distribution), transmucosal (diffusion through a mucous membrane), insufflation (snorting), sublingual, sublabial, enema, eye drops (onto the conjunctiva), or in ear drops. In specific embodiments, compositions may be administered in a way which allows them cross the blood-brain barrier, vascular barrier, or other epithelial barrier. Other delivery systems well known in the art can be used for delivery of the pharmaceutical compositions described herein, for example via aqueous solutions, encapsulation in microparticules, or microcapsules. The pharmaceutical compositions of the present invention can also be delivered in a controlled release system. For example, a polymeric material can be used (see, e.g., Smolen and Ball, Controlled Drug Bioavailability, Drug product design and performance, 1984, John Wiley & Sons; Ranade and Hollinger, Drug Delivery Systems, pharmacology and toxicology series, 2003, 2nd edition, CRRC Press). Alternatively, a pump may be used (Saudek et al., *N. Engl. J. Med.*

321:574 (1989)). The compositions described herein may also be coupled to a class of biodegradable polymers useful in achieving controlled release of the drug, for example, polylactic acid, polyorthoesters, cross-linked amphipathic block copolymers and hydrogels, polyhydroxy butyric acid, and polydihydropyrans.

[0197] As described above, pharmaceutical compositions desirably include a pharmaceutically acceptable carrier. The term carrier refers to diluents, adjuvants, excipients or vehicles with which modulators are administered. Such pharmaceutical carriers include sterile liquids such as water and oils including mineral oil, vegetable oil (e.g., soybean oil or corn oil), animal oil or oil of synthetic origin. Aqueous glycerol and dextrose solutions as well as saline solutions may also be employed as liquid carriers of the pharmaceutical compositions of the present invention. The choice of the carrier depends on factors well recognized in the art, such as the nature of the peptide, peptide derivative or peptidomimetic, its solubility and other physiological properties as well as the target site of delivery and application. Examples of suitable pharmaceutical carriers are described in Remington: The Science and Practice of Pharmacy by Alfonso R. Gennaro, 2003, 21th edition, Mack Publishing Company. Moreover, suitable carriers for oral administration are known in the art and are described, for example, in U.S. Patent Nos. 6,086,918, 6,673,574, 6,960,355, and 7,351,741 and in WO2007/131286, the disclosures of which are hereby incorporated by reference.

[0198] Further pharmaceutically suitable materials that may be incorporated in pharmaceutical preparations include absorption enhancers including those intended to increase paracellular absorption, pH regulators and buffers, osmolarity adjusters, preservatives, stabilizers, antioxidants, surfactants, thickeners, emollient, dispersing agents, flavoring agents, coloring agents, and wetting agents.

[0199] Examples of suitable pharmaceutical excipients include, water, glucose, sucrose, lactose, glycol, ethanol, glycerol monostearate, gelatin, starch flour (e.g., rice flour), chalk, sodium stearate, malt, sodium chloride, and the like. The pharmaceutical compositions comprising modulators can take the form of solutions, capsules, tablets, creams, gels, powders sustained release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides (see Remington: The Science and Practice of Pharmacy by Alfonso R. Gennaro, 2003, 21th edition, Mack Publishing Company).

Such compositions contain a therapeutically effective amount of the therapeutic composition, together with a suitable amount of carrier so as to provide the form for proper administration to the subject. The formulations are designed to suit the mode of administration and the target site of action (e.g., a particular organ or cell type).

[0200] The pharmaceutical compositions comprising the active agent described herein also include compositions formulated as neutral or salt forms. Pharmaceutically acceptable salts include those that form with free amino groups and those that react with free carboxyl groups. Non-toxic alkali metal, alkaline earth metal, and ammonium salts commonly used in the pharmaceutical industry include sodium, potassium, lithium, calcium, magnesium, barium, ammonium, and protamine zinc salts, which are prepared by methods well known in the art. Also included are non-toxic acid addition salts, which are generally prepared by reacting the compounds of the present invention with suitable organic or inorganic acid. Representative salts include the hydrobromide, hydrochloride, valerate, oxalate, oleate, laureate, borate, benzoate, sulfate, bisulfate, acetate, phosphate, tysolate, citrate, maleate, fumarate, tartrate, succinate, napsylate salts, and the like.

[0201] Examples of fillers or binders that may be used in accordance with the present invention include acacia, alginic acid, calcium phosphate (dibasic), carboxymethylcellulose, carboxymethylcellulose sodium, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, dextrin, dextrates, sucrose, tylose, pregelatinized starch, calcium sulfate, amylose, glycine, bentonite, maltose, sorbitol, ethylcellulose, disodium hydrogen phosphate, disodium phosphate, disodium pyrosulfite, polyvinyl alcohol, gelatin, glucose, guar gum, liquid glucose, compressible sugar, magnesium aluminum silicate, maltodextrin, polyethylene oxide, polymethacrylates, povidone, sodium alginate, tragacanth microcrystalline cellulose, starch, and zein. In certain embodiments, a filler or binder is microcrystalline cellulose.

[0202] Examples of disintegrating agents that may be used include alginic acid, carboxymethylcellulose, carboxymethylcellulose sodium, hydroxypropylcellulose (low substituted), microcrystalline cellulose, powdered cellulose, colloidal silicon dioxide, sodium croscarmellose, crospovidone, methylcellulose, polacrilin potassium, povidone, sodium alginate, sodium starch glycolate, starch, disodium disulfite, disodium edathamil, disodium edetate,

disodiumethylenediaminetetraacetate (EDTA) crosslinked polyvinylpyrrolidones, pregelatinized starch, carboxymethyl starch, sodium carboxymethyl starch, microcrystalline cellulose.

[0203] Examples of lubricants include calcium stearate, canola oil, glyceryl palmitostearate, hydrogenated vegetable oil (type I), magnesium oxide, magnesium stearate, mineral oil, poloxamer, polyethylene glycol, sodium lauryl sulfate, sodium stearate fumarate, stearic acid, talc and, zinc stearate, glyceryl behapate, magnesium lauryl sulfate, boric acid, sodium benzoate, sodium acetate, sodium benzoate/sodium acetate (in combination), DL-leucine.

[0204] Examples of silica flow conditioners include colloidal silicon dioxide, magnesium aluminum silicate and guar gum. Another most preferred silica flow conditioner consists of silicon dioxide.

[0205] Examples of stabilizing agents include acacia, albumin, polyvinyl alcohol, alginic acid, bentonite, dicalcium phosphate, carboxymethylcellulose, hydroxypropylcellulose, colloidal silicon dioxide, cyclodextrins, glyceryl monostearate, hydroxypropyl methylcellulose, magnesium trisilicate, magnesium aluminum silicate, propylene glycol, propylene glycol alginate, sodium alginate, carnauba wax, xanthan gum, starch, stearate(s), stearic acid, stearic monoglyceride and stearyl alcohol.

[0206] In some embodiments, the present invention contemplates oral formulations containing the active agent described herein. For example, pharmaceutical compositions described herein may include a cyclodextrin or cyclodextrin derivative. Cyclodextrins are generally made up of five or more α-D-glycopyranoside unites linked 1->4. Typically, cyclodextrins contain a number of glucose monomers ranging from six to eight units in a ring, creating a cone shape (α-cyclodextrin: six membered sugar ring molecule, β-cyclodextrin: seven sugar ring molecule, γ-cyclodextrin: eight sugar ring molecule). Exemplary cyclodextrins and cyclodextrin derivatives are disclosed in U.S. Patent No. 7,723,304, U.S. Publication No. 2010/0196452, and U.S. Publication No. 2010/0144624, the entire contents of each of which are incorporated herein by reference. For example, in some embodiments, a cyclodextrin in accordance with the present invention is an alkylated cyclodextrin, hydroxyalkylated cyclodextrin, or acylated cyclodextrin. In some embodiments, a cyclodextrin is a hydroxypropyl β-cyclodextrin. Exemplary cyclodextrin derivatives are disclosed in Szejtli, J. Chem Rev,

(1998), 98, 1743-1753; and Szente, L and Szejtli, J., Advance Drug Delivery Reviews, 36 (1999) 17-28, the entire contents of each of which are hereby incorporated by reference. Examples of cyclodextin derivatives include methylated cyclodextrins (e.g., RAMEB; randomly methylated β-cyclodextrin); hydroxyalkylated cyclodextrins (hydroxypropyl-β-cyclodextrin and hydroxypropyl γ-cyclodextrin); acetylated cyclodextrins (acetyl-γ-cyclodextrin); reactive cyclodextrins (chlorotriazinyl β- cyclodextrin); and branched cyclodextrins (glucosyl- and maltosyl β-cyclodextrin); acetyl-γ-cyclodextrin; acetyl-β-cyclodextrin, sulfobutyl-β cyclodextrin, sulfated α-, β- and γ-cyclodextrins; sulfoalkylated cyclodextrins; and hydroxypropyl β-cyclodextrin.

Dosing

[0207] Typically, active agent described herein in an amount ranging from 0.001 to 100 mg/kg/day is administered to the subject. For example, in some embodiments, about 0.01 mg/kg/day to about 25 mg/kg/day, about 1 mg/kg/day to about 20 mg/kg/day, 0.2 mg/kg/day to about 10 mg/kg/day, about 0.02 mg/kg/day to about 0.1 mg/kg/day, or about 1 mg/kg/day to about 100 mg/kg/day is administered to the subject. In some embodiments, active agent described herein in an amount of about 10 μg/kg/day, 50 μg/kg/day, 100 μg/kg/day, 200 μg/kg/day, 300 μg/kg/day, 400 μg/kg/day, 500 μg/kg/day, 600 μg/kg/day, 700 μg/kg/day, 800 μg/kg/day, 900 μg/kg/day, or 1000 μg/kg/day is administered to the subject.

[0208] In some embodiments, the compound is administered at an effective dose ranging from about 1-1,000 μg/kg/day (e.g., ranging from about 1-900 μg/kg/day, 1-800 μg/kg/day, 1-700 μg/kg/day, 1-600 μg/kg/day, 1-500 μg/kg/day, 1-400 μg/kg/day, 1-300 μg/kg/day, 1-200 μg/kg/day, 1-100 μg/kg/day, 1-90 μg/kg/day, 1-80 μg/kg/day, 1-70 μg/kg/day, 1-60 μg/kg/day, 1-50 μg/kg/day, 1-40 μg/kg/day, 1-30 μg/kg/day, 1-20 μg/kg/day, 1-10 μg/kg/day). In some embodiments, the compound is administered at an effective dose ranging from about 1-500 μg/kg/day. In some embodiments, the compound is administered at an effective dose ranging from about 1-100 μg/kg/day. In some embodiments, the compound is administered at an effective dose ranging from about 1-60 μg/kg/day. In some embodiments, the compound is administered at an effective dose selected from about 1, 2, 4, 6, 8, 10, 15, 20, 25, 30, 35, 40, 45,

50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, or 1,000 ug/kg/day.

[0209] In some embodiments, a therapeutically effective amount of a compound may be an amount ranging from about 10-1,000 mg (e.g., about 20 mg – 1,000 mg, 30 mg – 1,000 mg, 40 mg – 1,000 mg, 50 mg – 1,000 mg, 60 mg – 1,000 mg, 70 mg – 1,000 mg, 80 mg – 1,000 mg, 90 mg – 1,000 mg, about 10-900 mg, 10-800 mg, 10-700 mg, 10-600 mg, 10-500 mg, 100-1000 mg, 100-900 mg, 100-800 mg, 100-700 mg, 100-600 mg, 100-500 mg, 100-400 mg, 100-300 mg, 200-1000 mg, 200-900 mg, 200-800 mg, 200-700 mg, 200-600 mg, 200-500 mg, 200-400 mg, 300-1000 mg, 300-900 mg, 300-800 mg, 300-700 mg, 300-600 mg, 300-500 mg, 400 mg – 1,000 mg, 500 mg – 1,000 mg, 100 mg - 900 mg, 200 mg – 800 mg, 300 mg – 700 mg, 400 mg – 700 mg, and 500 mg – 600 mg). In some embodiments, a compound is present in an amount of or greater than about 10 mg, 50 mg, 100 mg, 150 mg, 200 mg, 250 mg, 300 mg, 350 mg, 400 mg, 450 mg, 500 mg, 550 mg, 600 mg, 650 mg, 700 mg, 750 mg, 800 mg. In some embodiments, a compound is present in an amount of or less than about 1000 mg, 950 mg, 900 mg, 850 mg, 800 mg, 750 mg, 700 mg, 650 mg, 600 mg, 550 mg, 500 mg, 450 mg, 400 mg, 350 mg, 300 mg, 250 mg, 200 mg, 150 mg, or 100 mg. In some embodiments, the therapeutically effective amount described herein is provided in one dose. In some embodiments, the therapeutically effective amount described herein is provided in one day.

[0210] In other embodiments, a therapeutically effective amount may be, for example, about 0.001 mg/kg weight to 500 mg/kg weight, e.g., from about 0.001 mg/kg weight to 400 mg/kg weight, from about 0.001 mg/kg weight to 300 mg/kg weight, from about 0.001 mg/kg weight to 200 mg/kg weight, from about 0.001 mg/kg weight to 100 mg/kg weight, from about 0.001 mg/kg weight to 90 mg/kg weight, from about 0.001 mg/kg weight to 80 mg/kg weight, from about 0.001 mg/kg weight to 70 mg/kg weight, from about 0.001 mg/kg weight to 60 mg/kg weight, from about 0.001 mg/kg weight to 50 mg/kg weight, from about 0.001 mg/kg weight to 40 mg/kg weight, from about 0.001 mg/kg weight to 30 mg/kg weight, from about 0.001 mg/kg weight to 25 mg/kg weight, from about 0.001 mg/kg weight to 20 mg/kg weight, from about 0.001 mg/kg weight to 15 mg/kg weight, from about 0.001 mg/kg weight to 10 mg/kg weight. In some embodiments, the therapeutically effective amount described herein is

provided in one dose. In some embodiments, the therapeutically effective amount described herein is provided in one day.

[0211] In still other embodiments, a therapeutically effective amount may be, for example, about 0.0001 mg/kg weight to 0.1 mg/kg weight, e.g. from about 0.0001 mg/kg weight to 0.09 mg/kg weight, from about 0.0001 mg/kg weight to 0.08 mg/kg weight, from about 0.0001 mg/kg weight to 0.07 mg/kg weight, from about 0.0001 mg/kg weight to 0.06 mg/kg weight, from about 0.0001 mg/kg weight to 0.05 mg/kg weight, from about 0.0001 mg/kg weight to about 0.04 mg/kg weight, from about 0.0001 mg/kg weight to 0.03 mg/kg weight, from about 0.0001 mg/kg weight to 0.02 mg/kg weight, from about 0.0001 mg/kg weight to 0.019 mg/kg weight, from about 0.0001 mg/kg weight to 0.018 mg/kg weight, from about 0.0001 mg/kg weight to 0.017 mg/kg weight, from about 0.0001 mg/kg weight to 0.016 mg/kg weight, from about 0.0001 mg/kg weight to 0.015 mg/kg weight, from about 0.0001 mg/kg weight to 0.014 mg/kg weight, from about 0.0001 mg/kg weight to 0.013 mg/kg weight, from about 0.0001 mg/kg weight to 0.012 mg/kg weight, from about 0.0001 mg/kg weight to 0.011 mg/kg weight, from about 0.0001 mg/kg weight to 0.01 mg/kg weight, from about 0.0001 mg/kg weight to 0.009 mg/kg weight, from about 0.0001 mg/kg weight to 0.008 mg/kg weight, from about 0.0001 mg/kg weight to 0.007 mg/kg weight, from about 0.0001 mg/kg weight to 0.006 mg/kg weight, from about 0.0001 mg/kg weight to 0.005 mg/kg weight, from about 0.0001 mg/kg weight to 0.004 mg/kg weight, from about 0.0001 mg/kg weight to 0.003 mg/kg weight, from about 0.0001 mg/kg weight to 0.002 mg/kg weight. In some embodiments, the therapeutically effective dose may be 0.0001 mg/kg weight, 0.0002 mg/kg weight, 0.0003 mg/kg weight, 0.0004 mg/kg weight, 0.0005 mg/kg weight, 0.0006 mg/kg weight, 0.0007 mg/kg weight, 0.0008 mg/kg weight, 0.0009 mg/kg weight, 0.001 mg/kg weight, 0.002 mg/kg weight, 0.003 mg/kg weight, 0.004 mg/kg weight, 0.005 mg/kg weight, 0.006 mg/kg weight, 0.007 mg/kg weight, 0.008 mg/kg weight, 0.009 mg/kg weight, 0.01 mg/kg weight, 0.02 mg/kg weight, 0.03 mg/kg weight, 0.04 mg/kg weight, 0.05 mg/kg weight, 0.06 mg/kg weight, 0.07 mg/kg weight, 0.08 mg/kg weight, 0.09 mg/kg weight, or 0.1 mg/kg weight. The effective dose for a particular individual can be varied (e.g., increased or decreased) over time, depending on the needs of the individual. In some embodiments, the therapeutically effective amount described herein is provided in one

dose. In some embodiments, the therapeutically effective amount described herein is provided in one day.

[0212] In some embodiments, a formulation comprising a compound as described herein administered as a single dose. In some embodiments, a formulation comprising a compound as described herein is administered at regular intervals. Administration at an "interval," as used herein, indicates that the therapeutically effective amount is administered periodically (as distinguished from a one-time dose). The interval can be determined by standard clinical techniques. In some embodiments, a formulation comprising a compound as described herein is administered bimonthly, monthly, twice monthly, triweekly, biweekly, weekly, twice weekly, thrice weekly, daily, twice daily, or every six hours. The administration interval for a single individual need not be a fixed interval, but can be varied over time, depending on the needs of the individual.

[0213] As used herein, the term "bimonthly" means administration once per two months (*i.e.,* once every two months); the term "monthly" means administration once per month; the term "triweekly" means administration once per three weeks (*i.e.*, once every three weeks); the term "biweekly" means administration once per two weeks (*i.e.*, once every two weeks); the term "weekly" means administration once per week; and the term "daily" means administration once per day.

[0214] In some embodiments, a formulation comprising a compound as described herein is administered at regular intervals indefinitely. In some embodiments, a formulation comprising a compound as described herein is administered at regular intervals for a defined period. In some embodiments, a formulation comprising a compound as described herein is administered at regular intervals for 5 years, 4, years, 3, years, 2, years, 1 year, 11 months, 10 months, 9 months, 8 months, 7 months, 6 months, 5 months, 4 months, 3 months, 2 months, a month, 3 weeks, 2, weeks, a week, 6 days, 5 days, 4 days, 3 days, 2 days, or a day.

Methods of Use

[0215] In certain embodiments provided compounds are useful in medicine. In some embodiments, provided compounds are useful in treating immune disease, disorders, or conditions. In some embodiments, the present invention provides a method for the treatment or prevention of an immune disease, disorder, or condition comprising administering to a subject in need thereof a provided compound or a pharmaceutical composition thereof.

[0216] In some embodiments, the immune disease, disorder, or condition is an autoimmune disease, disorder, or condition. In certain embodiments, the immune disease, disorder, or condition is selected from the group consisting of any of a variety of diseases, disorders, and/or conditions, including but not limited to one or more of the following: autoimmune disorders (*e.g.* diabetes, lupus, multiple sclerosis, psoriasis, rheumatoid arthritis); inflammatory disorders (*e.g.* arthritis, pelvic inflammatory disease); infectious diseases (*e.g.* viral infections (*e.g.*, HIV, HCV, RSV), bacterial infections, fungal infections, sepsis); neurological disorders (*e.g.* Alzheimer's disease, Huntington's disease; autism; Duchenne muscular dystrophy); cardiovascular disorders (*e.g.* atherosclerosis, hypercholesterolemia, thrombosis, clotting disorders, angiogenic disorders such as macular degeneration); proliferative disorders (*e.g.* cancer, benign neoplasms); respiratory disorders (*e.g.* chronic obstructive pulmonary disease); digestive disorders (*e.g.* inflammatory bowel disease, ulcers); musculoskeletal disorders (*e.g.* fibromyalgia, arthritis); endocrine, metabolic, and nutritional disorders (*e.g.* diabetes, osteoporosis); urological disorders (*e.g.* renal disease); psychological disorders (*e.g.* depression, schizophrenia); skin disorders (*e.g.* wounds, eczema); blood and lymphatic disorders (*e.g.* anemia, hemophilia); *etc.* In some embodiments, the immune disease, disorder, or condition is characterized by inflammation. In some embodiments, the immune disease, disorder, or condition is caused by, sustained by, or related to cGAS activation. In some embodiments, the immune disease, disorder, or condition is caused by, sustained by, or related to STING activation.

[0217] In some embodiments the autoimmune disorder or disease is selected from Acute disseminated encephalomyelitis (ADEM), Addison's disease, Agammaglobulinemia, Alopecia areata, Amyotrophic lateral sclerosis (Also Lou Gehrig's disease; Motor Neuron Disease),

Ankylosing Spondylitis, Antiphospholipid syndrome, Antisynthetase syndrome, Atopic allergy, Atopic dermatitis, Autoimmune aplastic anemia, Autoimmune cardiomyopathy, Autoimmune enteropathy, Autoimmune hemolytic anemia, Autoimmune hepatitis, Autoimmune inner ear disease, Autoimmune lymphoproliferative syndrome, Autoimmune peripheral neuropathy, Autoimmune pancreatitis, Autoimmune polyendocrine syndrome, Autoimmune progesterone dermatitis, Autoimmune thrombocytopenic purpura, Autoimmune urticarial, Autoimmune uveitis, Balo disease/Balo concentric sclerosis, Behçet's disease, Berger's disease, Bickerstaff's encephalitis, Blau syndrome, Bullous pemphigoid Cancer, Castleman's disease, Celiac disease, Chagas disease, Chronic inflammatory demyelinating polyneuropathy, Chronic recurrent multifocal osteomyelitis, Chronic obstructive pulmonary disease, Churg-Strauss syndrome, Cicatricial pemphigoid Cogan syndrome, Cold agglutinin disease, Complement component 2 deficiency, Contact dermatitis, Cranial arteritis, CREST syndrome, Crohn's disease (idiopathic inflammatory bowel disease "IBD"), Cushing's Syndrome, Cutaneous leukocytoclastic angiitis, Dego's disease, Dermatitis herpetiformis, Dermatomyositis, Diabetes mellitus type 1, Diffuse cutaneous systemic sclerosis, Dressler's syndrome, Drug-induced lupus, Discoid lupus erythematosus, Eczema, Endometriosis, Enthesitis-related arthritis, Eosinophilic fasciitis, Eosinophilic gastroenteritis, Epidermolysis bullosa acquisita, Erythema nodosum, Erythroblastosis fetalis, Essential mixed cryoglobulinemia, Evan's syndrome, Fibrodysplasia ossificans progressiva, Fibrosing alveolitis (or Idiopathic pulmonary fibrosis), Gastritis, Gastrointestinal pemphigoid, Glomerulonephritis, Goodpasture's syndrome, Graves' disease, Guillain-Barré syndrome (GBS), Hashimoto's encephalopathy, Hashimoto's thyroiditis, Henoch-Schonlein purpura, Herpes gestationis aka Gestational Pemphigoid, Hidradenitis suppurativa, Hughes-Stovin syndrome, Hypogammaglobulinemia, Idiopathic inflammatory demyelinating diseases, Idiopathic pulmonary fibrosis, Idiopathic thrombocytopenic purpura, IgA nephropathy, Inclusion body myositis, Chronic inflammatory demyelinating polyneuropathy, Interstitial cystitis, Juvenile idiopathic arthritis aka Juvenile rheumatoid arthritis, Kawasaki's disease, Lambert-Eaton myasthenic syndrome, Leukocytoclastic vasculitis, Lichen planus, Lichen sclerosus, Linear IgA disease (LAD), Lupoid hepatitis aka Autoimmune hepatitis, Lupus erythematosus, Majeed syndrome, Ménière's disease, Microscopic polyangiitis, Miller-Fisher syndrome see Guillain-Barre Syndrome, Mixed connective tissue disease, Morphea, Mucha-

Habermann disease aka Pityriasis lichenoides et varioliformis acuta, Multiple sclerosis, Myasthenia gravis, Myositis, Narcolepsy, Neuromyelitis optica (also Devic's disease), Neuromyotonia, Occular cicatricial pemphigoid, Opsoclonus myoclonus syndrome, Ord's thyroiditis, Palindromic rheumatism, PANDAS (pediatric autoimmune neuropsychiatric disorders associated with streptococcus), Paraneoplastic cerebellar degeneration, Paroxysmal nocturnal hemoglobinuria (PNH), Parry Romberg syndrome, Parsonage-Turner syndrome, Pars planitis, Pemphigus vulgaris, Pernicious anaemia, Perivenous encephalomyelitis, POEMS syndrome, Polyarteritis nodosa, Polymyalgia rheumatica, Polymyositis, Primary biliary cirrhosis, Primary sclerosing cholangitis, Progressive inflammatory neuropathy, Psoriasis, Psoriatic arthritis, Pyoderma gangrenosum, Pure red cell aplasia, Rasmussen's encephalitis, Raynaud phenomenon, Relapsing polychondritis, Reiter's syndrome, Restless leg syndrome, Retroperitoneal fibrosis, Rheumatoid arthritis, Rheumatic fever, Sarcoidosis, Schizophrenia, Schmidt syndrome, Schnitzler syndrome, Scleritis, Scleroderma, Serum Sickness, Sjögren's syndrome, Spondyloarthropathy, Still's disease, Subacute bacterial endocarditis (SBE), Susac's syndrome, Sweet's syndrome, Sydenham chorea see PANDAS, Sympathetic ophthalmia, Systemic lupus erythematosis, Takayasu's arteritis, Temporal arteritis (also known as "giant cell arteritis"), Thrombocytopenia, Tolosa-Hunt syndrome, Transverse myelitis, Ulcerative colitis (one of two types of idiopathic inflammatory bowel disease "IBD"), Undifferentiated connective tissue disease different from Mixed connective tissue disease, Undifferentiated spondyloarthropathy, Urticarial vasculitis, Vasculitis, Vitiligo, and Wegener's granulomatosis.

[0218] In certain embodiments, administration of a compound to a patient in need thereof results in a decrease of cGAS activity. In some embodiments, administration of a compound to a patient in need thereof results in a decrease of STING activity.

[0219] In some embodiments, compounds used in the provided methods are prepared by chemical synthesis.

[0220] In certain embodiments, the present invention provides a method of inhibiting cGAS comprising contacting cGAS with a provided compound. In some embodiments, the present invention provides a method of inhibiting cGAS in a patient comprising administering to a patient a provided compound. In certain embodiments, the present invention provides a method of inhibiting STING comprising contacting STING with a provided compound. In some

embodiments, the present invention provides a method of inhibiting STING in a patient comprising administering to a patient a provided compound.

[0221] In certain embodiments, the present invention provides a method of modulating activity of an cGAS polypeptide, the method comprising contacting the cGAS polypeptide with a cGAS modulator designed by the methods disclosed herein, which modulating agent is not a known modulator, substrate, or product of cGAS. In some embodiments, the modulating agent is a provided compound.

Kits

[0222] The invention provides a variety of kits for conveniently and/or effectively carrying out methods of the present invention. Typically kits will comprise sufficient amounts and/or numbers of components to allow a user to perform multiple treatments of a subject(s) and/or to perform multiple experiments.

[0223] In one aspect, the present invention provides kits comprising the molecules (compounds and compositions as described above) of the invention. In one embodiment, the kit comprises one or more functional antibodies or function fragments thereof.

[0224] Kits of the invention may comprise one or more cGAMP parent molecules, or any mimic, analog or variant thereof. Kits may also comprise any of the cGAS variants, derivatives or mutants described herein. The kit may further comprise packaging and instructions and/or a delivery agent to form a formulation composition. The delivery agent may comprise a saline, a buffered solution, a lipid or any delivery agent disclosed herein.

[0225] In one embodiment, the buffer solution may include sodium chloride, calcium chloride, phosphate and/or EDTA. In another embodiment, the buffer solution may include, but is not limited to, saline, saline with 2mM calcium, 5% sucrose, 5% sucrose with 2mM calcium, 5% Mannitol, 5% Mannitol with 2mM calcium, Ringer's lactate, sodium chloride, sodium chloride with 2mM calcium and mannose (See e.g., U.S. Pub. No. 20120258046; herein incorporated by reference in its entirety). In a futher embodiment, the buffer solutions may be precipitated or it may be lyophilized. The amount of each component may be varied to enable consistent, reproducible higher concentration saline or simple buffer formulations. The

components may also be varied in order to increase the stability of the compound or composition in the buffer solution over a period of time and/or under a variety of conditions. In one aspect, the present invention provides kits for research applications related to cGAS activity or cGAMP signaling, provided in an amount effective to study the concomitant signaling pathways when introduced into a target cell. The kits may further comprise a second or further compound or composition described herein. Such second or further molecules may modulate the immune response or an inflammatory process or comprise one or more therapeutic molecules. In one embodiment, a kit comprises at least one cGAS polypeptide and at least one cGAMP molecule. In one embodiment, the kits of the present invention comprise packaging and instructions.

cGAS Crystal Structures

[0226] Among other things, the present invention provides a crystalline (i.e., containing at least one crystal) or crystallizable composition comprising an cGAS polypeptide as described herein (see also Gao et al. *Cell* 153, 1094-1107 (2013), including supplementary materials). In some embodiments, such a provided composition consists of or consists essentially of the cGAS polypeptide. In some embodiments, a composition is considered to "consist of" cGAS polypeptide if it includes only the polypeptide, one or more solvents, and optionally salts and/or metals. In some embodiments, such a provided composition includes one or more other agents such as one or more other polypeptides (e.g., one or more potential or actual cGAS binding partner polypeptides or nucleic acids) and/or one or more interacting agents (e.g., small molecules).

[0227] The present invention also provides structural information and/or analyses of cGAS polypeptide crystals and/or sets thereof. In some embodiments, such structural information includes, but is not limited to, diffraction patterns, and/or coordinates, as well as any data sets, images, models, and/or graphical representations thereof or generated therefrom. In some embodiments, such graphical representations may include, for example, space-filling models, molecular surface representations, shell or boundary models, ribbon models, stick models; and/or combinations thereof.

[0228] In some embodiments, provided information is or comprises differences observed between or among structures that differ from one another in the presence or absence of one or more binding partners and/or interacting agents. In some embodiments, provided information is or comprises differences observed between or among structures that differ from one another in the presence or absence of one or more binding partners and/or one or more modulators.

[0229] In some embodiments, such structural information and/or analyses may be embodied in a tangible medium (e.g., a computer-readable medium) or a storage environment. Thus, the present invention provides tangible embodiments of cGAS polypeptide crystal structure information, as well as its use, for example, by or with a computer system, in any of a variety of applications. For example, in some embodiments, such structural information and/or analyses may be accessed by, transported to or from, and/or otherwise utilized by a computer system or program running thereon.

Structure-based Drug Design

[0230] In some embodiments, the present disclosure provides systems for identifying and/or characterizing cGAS modulators. In some embodiments, the present disclosure provides a method of designing or characterizing a cGAS modulator comprising the steps of:

[0231] a) providing an image of a cGAS crystal that includes at least one potential interaction site;

[0232] b) docking in the image at least one moiety that is a potential cGAS modulator structural element; and

[0233] c) assessing one or more features of a potential moiety-interaction site interaction.

[0234] In some embodiments, the at least one potential interaction site includes a site selected from the group consisting of Ser199, Ser420, Lys402, Glu211, Asp213, Asp307, Tyr421, Arg364, and combinations thereof. In certain embodiments, the at least one potential interaction site includes a site selected from the group consisting of Tyr421, Thr197, Ser366, Ser368, Arg364, and combinations thereof. In certain embodiments, the at least one potential interaction site includes a site selected from the group consisting of Tyr421, Asp213, Asp307,

Arg364, and combinations thereof. In some embodiments, the at least one potential interaction site includes Arg161. In some embodiments, the modulator is a compound disclosed herein.

[0235] In some embodiments, the one or more features include at least one feature selected from the group consisting of: spatial separation between the moiety and the potential interaction site; energy of the potential moiety-interaction site interaction, and/or combinations thereof.

[0236] In some embodiments, a method further comprises a step of providing an image of a potential cGAS modulator comprising the moiety docked with the image of the cGAS crystal. In some embodiments, a method further comprises a step of comparing the image with that of a cGAS crystal including a bound known modulator, substrate, or product.

Computer Systems

[0237] As will be appreciated by those skilled in the art, reading the present disclosure, in some aspects, the present invention is ideally suited for use in computer-implemented inventions. As shown in Figure S10, an implementation of an exemplary cloud computing environment 2400 is shown and described. The cloud computing environment 2400 may include one or more resource providers 2402a, 2402b, 2402c (collectively, 2402). Each resource provider 2402 may include computing resources. In some implementations, computing resources may include any hardware and/or software used to process data. For example, computing resources may include hardware and/or software capable of executing algorithms, computer programs, and/or computer applications. In some implementations, exemplary computing resources may include application servers and/or databases with storage and retrieval capabilities. Each resource provider 2402 may be connected to any other resource provider 2402 in the cloud computing environment 2400. In some implementations, the resource providers 2402 may be connected over a computer network 2408. Each resource provider 2402 may be connected to one or more computing device 2404a, 2404b, 2404c (collectively, 2404), over the computer network 2408.

[0238] The cloud computing environment 2400 may include a resource manager 2406. The resource manager 2406 may be connected to the resource providers 2402 and the computing

devices 2404 over the computer network 2408. In some implementations, the resource manager 2406 may facilitate the provision of computing resources by one or more resource providers 2402 to one or more computing devices 2404. The resource manager 2406 may receive a request for a computing resource from a particular computing device 2404. The resource manager 2406 may identify one or more resource providers 2402 capable of providing the computing resource requested by the computing device 2404. The resource manager 2406 may select a resource provider 2402 to provide the computing resource. The resource manager 2406 may facilitate a connection between the resource provider 2402 and a particular computing device 2404. In some implementations, the resource manager 2406 may establish a connection between a particular resource provider 2402 and a particular computing device 2404. In some implementations, the resource manager 2406 may redirect a particular computing device 2404 to a particular resource provider 2402 with the requested computing resource.

[0239] Figure S11 shows an example of a computing device 2500 and a mobile computing device 2550 that can be used to implement the techniques described in this disclosure. The computing device 2500 is intended to represent various forms of digital computers, such as laptops, desktops, workstations, personal digital assistants, servers, blade servers, mainframes, and other appropriate computers. The mobile computing device 2550 is intended to represent various forms of mobile devices, such as personal digital assistants, cellular telephones, smart-phones, tablet computers, and other similar computing devices. The components shown here, their connections and relationships, and their functions, are meant to be examples only, and are not meant to be limiting.

[0240] The computing device 2500 includes a processor 2502, a memory 2504, a storage device 2506, a high-speed interface 2508 connecting to the memory 2504 and multiple high-speed expansion ports 2510, and a low-speed interface 2512 connecting to a low-speed expansion port 2514 and the storage device 2506. Each of the processor 2502, the memory 2504, the storage device 2506, the high-speed interface 2508, the high-speed expansion ports 2510, and the low-speed interface 2512, are interconnected using various busses, and may be mounted on a common motherboard or in other manners as appropriate. The processor 2502 can process instructions for execution within the computing device 2500, including instructions stored in the memory 2504 or on the storage device 2506 to display graphical information for a

GUI on an external input/output device, such as a display 2516 coupled to the high-speed interface 2508. In other implementations, multiple processors and/or multiple buses may be used, as appropriate, along with multiple memories and types of memory. Also, multiple computing devices may be connected, with each device providing portions of the necessary operations (e.g., as a server bank, a group of blade servers, or a multi-processor system).

[0241] The memory 2504 stores information within the computing device 2500. In some implementations, the memory 2504 is a volatile memory unit or units. In some implementations, the memory 2504 is a non-volatile memory unit or units. The memory 2504 may also be another form of computer-readable medium, such as a magnetic or optical disk.

[0242] The storage device 2506 is capable of providing mass storage for the computing device 2500. In some implementations, the storage device 2506 may be or contain a computer-readable medium, such as a floppy disk device, a hard disk device, an optical disk device, or a tape device, a flash memory or other similar solid state memory device, or an array of devices, including devices in a storage area network or other configurations. Instructions can be stored in an information carrier. The instructions, when executed by one or more processing devices (for example, processor 2502), perform one or more methods, such as those described above. The instructions can also be stored by one or more storage devices such as computer- or machine-readable mediums (for example, the memory 2504, the storage device 2506, or memory on the processor 2502).

[0243] The high-speed interface 2508 manages bandwidth-intensive operations for the computing device 2500, while the low-speed interface 2512 manages lower bandwidth-intensive operations. Such allocation of functions is an example only. In some implementations, the high-speed interface 2508 is coupled to the memory 2504, the display 2516 (e.g., through a graphics processor or accelerator), and to the high-speed expansion ports 2510, which may accept various expansion cards (not shown). In the implementation, the low-speed interface 2512 is coupled to the storage device 2506 and the low-speed expansion port 2514. The low-speed expansion port 2514, which may include various communication ports (e.g., USB, Bluetooth®, Ethernet, wireless Ethernet) may be coupled to one or more input/output devices, such as a keyboard, a pointing device, a scanner, or a networking device such as a switch or router, e.g., through a network adapter.

[0244] The computing device 2500 may be implemented in a number of different forms, as shown in the figure. For example, it may be implemented as a standard server 2520, or multiple times in a group of such servers. In addition, it may be implemented in a personal computer such as a laptop computer 2522. It may also be implemented as part of a rack server system 2524. Alternatively, components from the computing device 2500 may be combined with other components in a mobile device (not shown), such as a mobile computing device 2550. Each of such devices may contain one or more of the computing device 2500 and the mobile computing device 2550, and an entire system may be made up of multiple computing devices communicating with each other.

[0245] The mobile computing device 2550 includes a processor 2552, a memory 2564, an input/output device such as a display 2554, a communication interface 2566, and a transceiver 2568, among other components. The mobile computing device 2550 may also be provided with a storage device, such as a micro-drive or other device, to provide additional storage. Each of the processor 2552, the memory 2564, the display 2554, the communication interface 2566, and the transceiver 2568, are interconnected using various buses, and several of the components may be mounted on a common motherboard or in other manners as appropriate.

[0246] The processor 2552 can execute instructions within the mobile computing device 2550, including instructions stored in the memory 2564. The processor 2552 may be implemented as a chipset of chips that include separate and multiple analog and digital processors. The processor 2552 may provide, for example, for coordination of the other components of the mobile computing device 2550, such as control of user interfaces, applications run by the mobile computing device 2550, and wireless communication by the mobile computing device 2550.

[0247] The processor 2552 may communicate with a user through a control interface 2558 and a display interface 2556 coupled to the display 2554. The display 2554 may be, for example, a TFT (Thin-Film-Transistor Liquid Crystal Display) display or an OLED (Organic Light Emitting Diode) display, or other appropriate display technology. The display interface 2556 may comprise appropriate circuitry for driving the display 2554 to present graphical and other information to a user. The control interface 2558 may receive commands from a user and convert them for submission to the processor 2552. In addition, an external interface 2562 may

provide communication with the processor 2552, so as to enable near area communication of the mobile computing device 2550 with other devices. The external interface 2562 may provide, for example, for wired communication in some implementations, or for wireless communication in other implementations, and multiple interfaces may also be used.

[0248] The memory 2564 stores information within the mobile computing device 2550. The memory 2564 can be implemented as one or more of a computer-readable medium or media, a volatile memory unit or units, or a non-volatile memory unit or units. An expansion memory 2574 may also be provided and connected to the mobile computing device 2550 through an expansion interface 2572, which may include, for example, a SIMM (Single In Line Memory Module) card interface. The expansion memory 2574 may provide extra storage space for the mobile computing device 2550, or may also store applications or other information for the mobile computing device 2550. Specifically, the expansion memory 2574 may include instructions to carry out or supplement the processes described above, and may include secure information also. Thus, for example, the expansion memory 2574 may be provide as a security module for the mobile computing device 2550, and may be programmed with instructions that permit secure use of the mobile computing device 2550. In addition, secure applications may be provided via the SIMM cards, along with additional information, such as placing identifying information on the SIMM card in a non-hackable manner.

[0249] The memory may include, for example, flash memory and/or NVRAM memory (non-volatile random access memory), as discussed below. In some implementations, instructions are stored in an information carrier. that the instructions, when executed by one or more processing devices (for example, processor 2552), perform one or more methods, such as those described above. The instructions can also be stored by one or more storage devices, such as one or more computer- or machine-readable mediums (for example, the memory 2564, the expansion memory 2574, or memory on the processor 2552). In some implementations, the instructions can be received in a propagated signal, for example, over the transceiver 2568 or the external interface 2562.

[0250] The mobile computing device 2550 may communicate wirelessly through the communication interface 2566, which may include digital signal processing circuitry where necessary. The communication interface 2566 may provide for communications under various

modes or protocols, such as GSM voice calls (Global System for Mobile communications), SMS (Short Message Service), EMS (Enhanced Messaging Service), or MMS messaging (Multimedia Messaging Service), CDMA (code division multiple access), TDMA (time division multiple access), PDC (Personal Digital Cellular), WCDMA (Wideband Code Division Multiple Access), CDMA2000, or GPRS (General Packet Radio Service), among others. Such communication may occur, for example, through the transceiver 2568 using a radio-frequency. In addition, short-range communication may occur, such as using a Bluetooth®, Wi-Fi™, or other such transceiver (not shown). In addition, a GPS (Global Positioning System) receiver module 2570 may provide additional navigation- and location-related wireless data to the mobile computing device 2550, which may be used as appropriate by applications running on the mobile computing device 2550.

[0251] The mobile computing device 2550 may also communicate audibly using an audio codec 2560, which may receive spoken information from a user and convert it to usable digital information. The audio codec 2560 may likewise generate audible sound for a user, such as through a speaker, e.g., in a handset of the mobile computing device 2550. Such sound may include sound from voice telephone calls, may include recorded sound (e.g., voice messages, music files, etc.) and may also include sound generated by applications operating on the mobile computing device 2550.

[0252] The mobile computing device 2550 may be implemented in a number of different forms, as shown in the figure. For example, it may be implemented as a cellular telephone 2580. It may also be implemented as part of a smart-phone 2582, personal digital assistant, or other similar mobile device.

[0253] Various implementations of the systems and techniques described here can be realized in digital electronic circuitry, integrated circuitry, specially designed ASICs (application specific integrated circuits), computer hardware, firmware, software, and/or combinations thereof. These various implementations can include implementation in one or more computer programs that are executable and/or interpretable on a programmable system including at least one programmable processor, which may be special or general purpose, coupled to receive data and instructions from, and to transmit data and instructions to, a storage system, at least one input device, and at least one output device.

[0254] These computer programs (also known as programs, software, software applications or code) include machine instructions for a programmable processor, and can be implemented in a high-level procedural and/or object-oriented programming language, and/or in assembly/machine language. As used herein, the terms machine-readable medium and computer-readable medium refer to any computer program product, apparatus and/or device (e.g., magnetic discs, optical disks, memory, Programmable Logic Devices (PLDs)) used to provide machine instructions and/or data to a programmable processor, including a machine-readable medium that receives machine instructions as a machine-readable signal. The term machine-readable signal refers to any signal used to provide machine instructions and/or data to a programmable processor.

[0255] To provide for interaction with a user, the systems and techniques described here can be implemented on a computer having a display device (e.g., a CRT (cathode ray tube) or LCD (liquid crystal display) monitor) for displaying information to the user and a keyboard and a pointing device (e.g., a mouse or a trackball) by which the user can provide input to the computer. Other kinds of devices can be used to provide for interaction with a user as well; for example, feedback provided to the user can be any form of sensory feedback (e.g., visual feedback, auditory feedback, or tactile feedback); and input from the user can be received in any form, including acoustic, speech, or tactile input.

[0256] The systems and techniques described here can be implemented in a computing system that includes a back end component (e.g., as a data server), or that includes a middleware component (e.g., an application server), or that includes a front end component (e.g., a client computer having a graphical user interface or a Web browser through which a user can interact with an implementation of the systems and techniques described here), or any combination of such back end, middleware, or front end components. The components of the system can be interconnected by any form or medium of digital data communication (e.g., a communication network). Examples of communication networks include a local area network (LAN), a wide area network (WAN), and the Internet.

[0257] The computing system can include clients and servers. A client and server are generally remote from each other and typically interact through a communication network. The

relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other.

[0258] In certain embodiments, the present invention provides a system comprising a computer or computer readable medium in which a cGAS crystal structure, or coordinates thereof, is embedded and/or displayed.

[0259] In some embodiments, the present invention provides a method of designing and/or characterizing an cGAS modulator, which method comprises steps of:

(i) using a provided system to assess one or more structural features of the cGAS modulator; and

(ii) performing one or more *in vitro, in vivo* or cell-based assays to characterize the cGAS modulator.

[0260] In some embodiments, the method further comprises the step of performing a competition experiment between the cGAS modulator and a known cGAS modulator, substrate, or product. In some embodiments, the method further comprises the step of defining the three-dimensional shape of the inhibitor.

[0261] In some embodiments, the present invention provides a computer system containing a set of information to perform a design or characterization of an cGAS inhibitor having a user interface comprising a display unit, the set of information comprising:

(i) logic for inputting an information regarding a binding of a cGAS protein to a moiety known to bind cGAS protein;

(ii) logic for design a candidate cGAS inhibitor based on the binding of the cGAS protein to the moiety known to bind cGAS protein;

(iii) logic for determining an information regarding a binding of the cGAS protein to the candidate cGAS inhibitor; and

(iv) logic for making a conclusion regarding a cGAS inhibitory properties of the candidate cGAS inhibitor based on the determination of step (iii).

[0262] In some embodiments, the present invention provides a computer-readable storage medium containing a set of information for a general purpose computer having a user interface comprising, a display unit, the set of information comprising:

(i) logic for inputting an information regarding a binding of a cGAS protein to a chemical known to binding cGAS protein;

(ii) logic for design a candidate cGAS inhibitor based on the binding of the cGAS protein to the chemical known to bind cGAS protein;

(iii) logic for determining an information regarding a binding of the cGAS protein to the candidate cGAS inhibitor; and

(iv) logic for making a conclusion regarding a cGAS inhibitory properties of the candidate cGAS inhibitor based on the determination step of step (iii).

[0263] In some embodiments, the present invention provides an electronic signal or carrier wave that is propagated over the internet between computers comprising a set of information for a general purpose computer having a user interface comprising a display unit, the set of information comprising a computer-readable storage medium containing a set of information for a general purpose computer having a user interface comprising a display unit, the set of information comprising:

(i) logic for inputting an information regarding a binding of a cGAS protein to a chemical known to bind cGAS protein;

(ii) logic for designing a candidate cGAS inhibitor based on the binding of the cGAS protein to the chemical known to bind cGAS protein;

(iii) logic for determining an information regarding a binding of the cGAS protein to the candidate cGAS inhibitor; and

(iv) logic for making a conclusion regarding a cGAS inhibitory properties of the candidate cGAS inhibitor based on the determination of step (iii).

EXAMPLES

[0264] The following coordinates have been deposited in the RCSB Protein Data Bank, with which the skilled artisan will be familiar, and correspond to Tables 1-7 referenced herein. See also Gao et al. *Cell* 153, 1094-1107 (2013), including supplementary materials, the entire contents of which are hereby incorporated by reference herein. Furthermore, in the context of the ensuing Figures 1-4, 7, and S1-S4, the data presented in Tables 1-7 of United States provisional patent application no. 61/819,369, filed May 3, 2013, is hereby incorporated by reference.

Table E1

Sample	PDB code	rcsb code [1]	Table
cGAS	4K8V	RCSB079037	1
cGAS + DNA	4K96	RCSB079048	2
cGAS + DNA + ATP	4K97	RCSB079049	3
cGAS + DNA + 5'-pppG(2',5')pG	4K98	RCSB079050	4
cGAS + DNA + 5'-pppdG(2',5')pdG	4K99	RCSB079051	5
cGAS + DNA + 5'-pG(2',5')pA	4K9A	RCSB079052	6
cGAS + DNA + c[G(2',5')pA(3',5')p]	4K9B	RCSB079053	7

[1] One method of accessing the RCSB Protein Data Bank is online at www.rcsb.org.

Example 1

Crystal Structures

Protein Expression and Purification

[0265] The gene encoding mouse cGAS was purchased from Open Biosystems Inc. The sequences corresponding to full-length and residues 147-507 of cGAS were inserted into a modified pRSFDuet-1 vector (Novagen), in which cGAS was separated from the preceding His_6-SUMO tag by an ubiquitin-like protease (ULP1) cleavage site. The gene sequences were

subsequently confirmed by sequencing. The fusion proteins were expressed in BL21 (DE3) RIL cell strain. The cells were grown at 37 °C until OD600 reached approx. 0.6. The temperature was then shifted to 18 °C and the cells were induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to the culture medium at a final concentration of 0.3 mM. After induction, the cells were grown overnight. The fusion protein was purified over a Ni-NTA affinity column. The His_6-SUMO tag was removed by ULP1 cleavage during dialysis against buffer containing 40 mM Tris-HCl, 0.3 M NaCl, 1 mM DTT, pH 7.5. After dialysis, the protein sample was further fractionated over a Heparin column, followed by gel filtration on a 16/60 G200 Superdex column. The final sample of cGAS (full-length) and cGAS (147-507) contain about 30mg/ml protein, 20 mM Tris, 300 mM NaCl, 1 mM DTT, pH 7.5. The Se-methionine substituted protein was expressed in Se-methionine (Sigma) containing M9 medium and purified using the same protocol used for the wild-type protein. All the mutants were cloned and purified using the same protocol as used for preparation of the wild-type protein.

Crystallization

[0266] For crystallization of cGAS (147-507) in the free state, the protein was first diluted into about 15 mg/ml and then mixed with equal volume reservoir solution (0.1 M HEPES, 0.1 M $MgAc_2$, 20% PEG3350, pH 7.6) at 4 °C by using hanging drop vapor diffusion method.

[0267] For cGAS (147-507)-dsDNA binary complex, the sample was prepared by direct mixing protein with a 16-bp DNA (1-nt 5'-overhang at either end: upper strand 5'-AAATTGCCGAAGACGAA-3'; lower strand 5'- TTTCGTCTTCGGCAATT-3') in a 1:1.2 molar ratio. The crystals were generated by hanging drop vapor diffusion method at 20 °C, from drops mixed from 1 μl of cGAS-dsDNA solution and 1 μl of reservoir solution (0.1 M MES, 8% MPD, pH 6.6). The crystals of Se-methionine substituted cGAS (147-507) in complex with dsDNA were grown under the same condition.

[0268] The cGAS (147-507)-dsDNA-ATP, cGAS (147-507)-dsDNA-GTP, and cGAS (147-507)-dsDNA-3'-dGTP ternary complexes were prepared by mixing protein with dsDNA in a 1:1.2 molar ratio, and then incubated in the presence of ATP/GTP/3'-dGTP (5 mM) and $MgCl_2$ (10 mM) for 0.5 h at room temperature. The crystals for cGAS (147-507)-dsDNA-ATP complex

were generated by hanging drop vapor diffusion method at 20 °C, from drops mixed from 1 μl of cGAS-dsDNA-ATP solution and 1 μl of reservoir solution (0.1 M HEPES, 0.2 M $CaAc_2$, 20% PEG300, pH 7.7). For cGAS (147-507)-dsDNA-GTP and cGAS (147-507)-dsDNA-3'-dGTP complexes, the crystals were generated by sitting drop vapor diffusion method at 20 °C, by mixing equal volume reservoir solution (for GTP: 0.1 M NaAc, 10% MPD, pH 5.0; for 3'-dGTP: 0.1 M NaAc, 12% MPD, pH 5.2) with the samples.

[0269] The cGAS (147-507)-dsDNA-GMP+ATP and cGAS (147-507)-dsDNA-GTP+ATP ternary complexes were prepared by mixing protein with dsDNA in a 1:1.2 molar ratio, and then incubated with GMP/GTP (5 mM), ATP (5 mM) and $MgCl_2$ (10 mM) for 0.5 h at room temperature. The crystals for cGAS (147-507)-dsDNA-GMP+ATP complex were generated by sitting drop vapor diffusion method at 20 °C, from drops mixed cGAS-dsDNA-GMP+ATP solution with equal volume reservoir solution (0.1 M MES, 40% MPD, pH 6.0). The crystals for cGAS (147-507)-DNA-GTP+ATP complex were generated over two weeks by sitting drop vapor diffusion method at 20 °C, by mixing equal volume reservoir solution (0.1 M HEPES, 0.2 M $MgCl_2$, 30% PEG300, pH 7.5) with the sample.

Structure Determination

[0270] The heavy atom derivative crystal of the free state was generated by soaking in a reservoir solution with 5 mM thimerosal for 24h. The diffraction data sets for cGAS (147-507) in free state (both native and Hg-derivative) and DNA-bound state (both native and Se-derivative) were collected at the Brookhaven National Laboratory. The data sets for all the ternary complexes were collected at the Advanced Photo Source (APS) at the Argonne National Laboratory. The diffraction data were indexed, integrated and scaled using the HKL2000 program (Otwinowski and Minor, 1997). The structure of Hg-substituted cGAS (147-507) in free state and Se-substituted cGAS (147-507) in DNA bound state were both solved using single-wavelength anomalous dispersion method as implemented in the program PHENIX (Adams et al., 2010). The model building was carried out using the program COOT (Emsley et al., 2010) and structural refinement was carried out using the program PHENIX (Adams et al., 2010). The statistics of the data collection and refinement for free and binary structures are shown in Table S1. The structures of all the ternary complexes were solved using molecular replacement method

in PHASER (McCoy et al., 2007) using the binary structure as the search model. The model building was conducted using the program COOT (Emsley et al., 2010) and structural refinement was conducted using the program PHENIX (Adams et al., 2010). The statistics of the data collection and refinement are shown in Table S2 and S3.

Table S1

Data collection and refinement statistics for structures of cGAS in free and DNA bound state

Crystal	cGAS	cGAS + DNA
Beam line	NSLS-29X	NSLS-29X
Wavelength	0.9790	0.9790
Space group	$P2_1$	$C2$
Unit cell		
a, b, c (Å)	86.6, 84.1, 124.7	181.9, 93.8, 75.5
α, β, γ (°)	90.0, 92.7, 90.0	90.0, 97.7, 90.0
Resolution (Å)	50-2.0 (2.07-2.00)[a]	50-2.1 (2.18-2.10)[a]
R_{merge}	0.179 (0.493)	0.089 (0.511)
I/ σ (I)	15.7 (3.6)	15.8 (3.1)
Completeness (%)	99.3 (98.6)	99.8 (100)
Redundancy	7.6 (7.7)	5.5 (5.3)
Number of unique reflections	118611	74352
R_{work}/R_{free} (%)	17.5/20.8	20.2/22.6
Number of non-H atoms		
Protein/DNA	11957	7257
Water	1357	722
Ion	4	2
Average B factors (Å^2)		
Protein	34.01	40.90
DNA		68.30
Water	35.56	43.80
Ion	13.80	56.96
R.m.s. deviations		
Bond lengths (Å)	0.010	0.003
Bond angles (°)	1.207	0.947

[a] Highest resolution shell (in Å) shown in parentheses.

Table S2

Data collection and refinement statistics of ternary complexes of cGAS and dsDNA with ATP, GTP, and 3'-dGTP

Crystal	**cGAS + DNA + ATP**	**cGAS + DNA + GTP [5'-pppG(2',5')pG]**	**cGAS + DNA + 3'-dGTP [5'-pppdG(2',5')pdG]**
Beam line	APS-24ID-C	APS-24ID-E	APS-24ID-E
Wavelength	0.9823	0.9792	0.9792
Space group	*I*222	*I*222	*I*222
Unit cell			
a, b, c (Å)	86.2, 99.4, 131.5	85.4, 97.9, 133.5	85.1, 97.6, 131.4
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	50-2.4 (2.54-2.41)[a]	50-1.9 (2.04-1.94)[a]	50-2.0 (2.05-1.95)[a]
R_{merge}	0.079 (0.577)	0.067 (0.643)	0.079 (0.700)
I/ σ (I)	14.2 (2.4)	20.5 (2.2)	17.5 (2.3)
Completeness (%)	99.6 (97.8)	98.9 (93.5)	100 (100)
Redundancy	6.3 (6.3)	9.1 (5.6)	10.7 (10.7)
Number of unique reflections	22099	41225	40199
R_{work}/R_{free} (%)	19.3/23.6	15.7/19.9	17.6/21.2
Number of non-H atoms			
Protein/DNA	3492	3521	3521
Water	147	363	318
Ion	3	3	3
Other ligands	31	55	53
Average B factors (Å^2)			
Protein	51.83	35.77	38.17
DNA	83.79	64.04	64.83
Water	44.70	40.68	40.78
Ion	35.56	21.95	35.32
Other ligands	58.44	32.66	43.53
R.m.s. deviations			
Bond lengths (Å)	0.007	0.018	0.008
Bond angles (°)	1.190	2.183	1.718

[a] Highest resolution shell (in Å) shown in parentheses.

Table S3.

Data collection and refinement statistics of ternary complexes of cGAS and dsDNA with GMP + ATP and GTP + ATP

Crystal	cGAS + DNA + GMP + ATP [5'-pG(2',5')pA]	cGAS + DNA + GTP + ATP c[G(2',5')pA(3',5')p]
Beam line	APS-24ID-C	APS-24ID-E
Wavelength	0.9795	0.9792
Space group	*I*222	*I*222
Unit cell		
a, b, c (Å)	85.4, 98.0, 131.3	85.3, 98.3, 130.0
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	50-2.3 (2.39-2.26)[a]	50-2.3 (2.38-2.26)[a]
R_{merge}	0.059 (0.303)	0.084 (0.904)
I/ σ (I)	31.5 (4.8)	17.6 (2.4)
Completeness (%)	98.3 (89.2)	100 (100)
Redundancy	12.1 (5.5)	9.7 (9.9)
Number of unique reflections	25590	25950
R_{work}/R_{free} (%)	16.8/21.0	17.6/21.8
Number of non-H atoms		
Protein/DNA	3521	3518
Water	209	189
Ion	1	1
Other ligands	46	45
Average B factors (Å^2)		
Protein	43.30	48.35
DNA	74.18	71.50
Water	43.86	43.94
Ion	30.20	25.89
Other ligands	54.77	89.75
R.m.s. deviations		
Bond lengths (Å)	0.005	0.010
Bond angles (°)	1.433	1.933

[a] Highest resolution shell (in Å) shown in parentheses.

Structure of cyclic GMP-AMP Synthase (cGAS)

[0271] We have solved the 2.0 Å crystal structure of cGAS (construct 147-507) (Figure S1A) in the free state (Figure 1A and S1B). The protein adopts a bilobal scaffold with mixed α/β topology (Figure 1A; x-ray statistics in Table S1) characteristic of members of the nucleotidyltransferase superfamily. A DALI search identified ILF2/NF45, which contains a nucleotidyltransferase fold (PDB: 4AT8) (Wolkowicz and Cook, 2012), as most closely resembling the fold of cGAS, with a Z score of 15.1 and r.m.s.d of 3.8 Å. In addition, the free state of human oligoadenylate synthetase 1 (OAS1) (PDB: 1PX5) (Hartmann et al. 2003) exhibited a Z score of 13.3 and a r.m.s.d. of 4.1 Å (comparison of cGAS and OAS1 in the free state in stereo in Figure S1C).

Structure of Binary Complex of cGAS with Bound dsDNA

[0272] We have cocrystallized cGAS bound to a 16-bp complementary dsDNA (plus 1-nt 5'-overhang at either end) and solved the structure of the binary complex at 2.1 Å resolution (x-ray statistics in Table S1). The structure of the binary complex is shown in Figure 1B. The majority of the intermolecular contacts in the binary cGAS-dsDNA complex (summarized in Figure 1C) are between cGAS and the sugar-phosphate backbone of the DNA (Figure S2A, B), with only one base-specific contact (Figure S2B). The superposed structures of cGAS in the free (light gray) and dsDNA-bound (dark gray) states are shown in Figure 1D. There are large conformational changes on formation of the binary dsDNA complex as can be seen within a β-sheet segment containing catalytic Glu211, Asp213 and Asp307 residues (Figure 1E), as well for loop and helical segments within the catalytic pocket containing Ser199 (Figure 1F). Thus, a β-sheet segment shifts by 5.1 Å on complex formation (Figure S2C), as does Arg161 involved in base-specific recognition by 9.2 Å (Figure S2D), as do Tyr and Lys residues within loop segments by up to 17.6 Å (Figure S2E). Equally important, a very narrow entrance leads to the catalytic pocket for cGAS in the free state (Figure 1G), while this entrance widened significantly in the binary complex with DNA (Figure 1H). The cGAS fold in the dsDNA bound state is similar to that reported recently for the OAS1 in the dsRNA bound state plus 2'-dATP (PDB:

4IG8) (Donovan et al. 2013) (comparison of proteins in complexes in stereo in Figure S2D; Z score of 18.2 and r.m.s.d. of 3.2 Å between the two protein folds).

Structure of Ternary Complex of cGAS with dsDNA and Bound ATP

[0273] We cocrystallized cGAS bound to dsDNA and ATP and solved the structure of the ternary complex at 2.4 Å resolution. The ATP is bound in the catalytic pocket positioned within the interior of the cGAS in the ternary complex (Figure 2A). There is close superposition of the binary complex of cGAS and dsDNA with the ternary complex containing bound ATP as shown in Figure 2B, with essentially minimal conformational changes in either the β-sheet segment carrying the catalytic acidic residues (Figure 2C) or the loop and helix segments forming the catalytic pocket (Figure 2D) on ternary complex formation. The only notable change is the movement of the side chain of Glu211 towards the other two acidic residues in the ternary complex (Figure 2C). The triphosphate group of ATP is hydrogen bonded to polar side chains (Ser199, Ser420 and Lys402), while two bound cations (tentatively assigned to Mg^{2+}) serve a bridging role for interactions between the triphosphate and the side chains of catalytic acidic residues (Glu211, Asp213 and Asp307) (Figure 2E). In addition, the adenine ring of bound ATP stacks over Tyr421 in one direction and partially over the guanidinium group of Arg364 in the other direction (Figure 2F).

[0274] It should be noted that we observe additional weak electron density (dark gray contours in Figure 2G) that is unaccounted for at this time in the 2.4 Å structure of the ternary complex. The additional density could be either water molecules or an AMP molecule with modest (30 %) occupancy. A view of the bound ATP looking into the catalytic pocket of the ternary complex is shown in Figure 2H.

Structure of Ternary Complex of cGAS with dsDNA and Bound 5'-pppG(2',5')pG

[0275] We cocrystallized cGAS bound to dsDNA and GTP and solved the structure of the ternary complex at 1.9 Å resolution. The structure of the ternary complex is shown in Figure 3A (x-ray statistics in Table S2). Notably, phosphodiester bond formation has occurred in the catalytic pocket yielding the bound ligand 5'-pppGpG (shown positioned in the catalytic pocket in space-filling representation in Figure 3A). Importantly, minimal conformational changes

occurred on proceeding from the binary complex to the ternary complex with bound 5'-pppGpG (Figure S3A-C).

[0276] Strikingly, the GpG linkage of 5'-pppGpG is 2',5' rather than the anticipated 3',5', with the first and second G residues in addition adopting *syn* and *anti* glycosidic torsion orientations, respectively (Figure S3D). The triphosphate group of 5'-pppG(2',5')pG is coordinated to two cations (Figure 3B), with the first G stacked on Tyr421, while the second G uses its Watson-Crick edge to hydrogen bond with polar side chains (Thr197, Ser366 and Ser368), and its Hoogsteen edge to hydrogen bond with Arg364 (Figure 3C). The observed 5'-pppG(*syn*)(2',5')pG(*anti*) topology can be traced with a high degree of confidence because of the clear density observed for this intermediate of the reaction in the 1.9 Å structure of the ternary complex (2Fo-Fc map in Figure 3D, with two views of the Fo-Fc omit map shown in Figure S3E). A view of the bound 5'-pppG(2',5')pG looking into the catalytic pocket of the ternary complex shown is in Figure 3E. We have superposed the structures of bound 5'-pppG(2',5')pG (gray) and ATP (dark gray) in their respective ternary complexes with cGAS and dsDNA, and observe that the first G of the bound 5'-pppG(2',5')pG is positioned in the plane of the bound ATP (Figure S3F). The two bound cations have been tentatively assigned to Mg^{2+} based on omit maps (Figure S3G) and the octahedral coordination around each cation (Figure S3H, I).

[0277] We also grew crystals of the ternary complex with 3'-dGTP, and observed formation of the related 5'-pppdG(2',5')pdG intermediate (cannot form a 3',5' linkage) in the 2.1 Å structure of this complex (x-ray statistics in Table S2).

Structure of Ternary Complex of cGAS with dsDNA and Bound 5'-pG(2',5')pA

[0278] We have also cocrystallized cGAS in the presence of dsDNA, GMP, and ATP and solved the structure of a complex at 2.3 Å resolution (structural statistics in Table S3). By using GMP rather than GTP, we hoped to trap the intermediate following formation of the first phosphodiester bond, and observed indeed the bound linear product of 5'-pG(*syn*)(2',5')pA(*anti*) (Figure 3F, G). No Mg^{2+} cations were observed in the absence of a triphosphate moiety in the product. Notably, attempts at cocrystallization of cGAS with dsDNA, GTP, and AMP only yielded crystals that diffracted very poorly (12 Å resolution). We observed good superposition of

the intermediates 5'-pppG(*syn*)(2',5')pG(*anti*) (in dark gray) and 5'-pG(*syn*)(2',5')pA(*anti*) (in light gray) as shown in Figure 3H.

Structure of Ternary Complex of cGAS with dsDNA and Bound c[G(2',5')pA(3',5')p]

[0279] We cocrystallized cGAS with dsDNA, GTP, and ATP and solved the structure of the complex at 2.3 Å resolution. These crystals took two weeks to grow, unlike other crystals mentioned above, that grew within a few days. The structure of the ternary complex is shown in Figure 4A (x-ray statistics in Table S3). Most unexpectedly, the bound small ligand shown in a space-filling representation in Figure 4A, is a cyclic dinucleotide. Notably, no conformational changes occurred on proceeding from the binary complex to the ternary complex with bound cyclic dinucleotide, with even the side chain of Glu211 adopting identical orientations (Figure S4A-C).

[0280] Importantly, we can trace the GpA step in the bound cyclic dinucleotide without ambiguity (the 3'-OH of G can be traced) and establish that this linkage is 2',5' (Figure. S4D). On the other hand, the linkage at the ApG step in the bound cyclic dinucleotide could be either 2',5' or 3',5' based on the observed density, and cannot be assigned with certainty solely based on structure. We have undertaken the refinement with a 3',5' linkage at the ApG step based on evidence outlined later and prepared the drawings in Figure 4 and S4 with 2',5' linkage at the GpA step and 3',5' linkage at the ApG step. We can distinguish G from A based on the observed density for the 2-amino group of G, and note that both adopt *anti* alignments in the bound cyclic dinucleotide {c[G(2',5')pA(3',5')p]} (Figure S4D). The A residue of the bound c[G(2',5')pA(3',5')p] is stacked on Tyr421 (Figure 4B, C), while the G residue of the bound c[G(2',5')pA(3',5')p] is anchored in place through hydrogen bonding to the side chains of Asp213, Asp307 and Arg364 (Figure 4B,C). Further, the A and G residues partially stack on each other. The 2Fo-Fc electron density for the bound c[G(2',5')pA(3',5')p] is shown in Figure 4D, with omit maps shown in Figure S4E. A view of the bound c[G(2',5')pA(3',5')p] looking into the catalytic pocket of the ternary complex is shown in Figure 4E, with the c[G(2',5')pA(3',5')p] bound towards one end of the opening. We also do not observe bound cations, given that c[G(2',5')pA(3',5')p] does not contain triphosphates, and the G base directly coordinates with Asp213 and Asp307 (Figure 4C). We have superposed the structures of bound

c[G(2',5')pA(3',5')p] and ATP in their respective ternary complexes with cGAS and dsDNA, and observe that the A of the bound c[G(2',5')pA(3',5')p] is positioned in the plane of the bound ATP (Figure S4F).

[0281] A view of c[G(2',5')pA(3',5')p] highlighting the 2',5' linkage at the GpA step and the 3',5' linkage at the ApG step is shown in Figure 4F. We note that in the ternary complex with 5'-pG(2',5')pA linear product, it is the G base that stacks on Tyr421 (Figure 3G and 4G), while in the ternary complex with c[G(2',5')pA(3',5')p] product, it is the A base that stacks on Tyr421 (Figure 4C, H). Thus, the linear product and cyclic final product adopt different alignments within the catalytic pocket.

Example 2

Biochemical Characterization of cGAS Activity

[0282] To validate the structural results, we established an activity assay using thin-layer chromatography (Figure S5) and monitored cyclic dinucleotide c[G(2',5')pA(3',5')p] formation from ATP and GTP using purified recombinant full-length and truncated cGAS proteins. cGAS required the presence of dsDNA and Mg^{2+} or Mn^{2+} for activity (Figure 5A-B). We tested c[G(2',5')pA(3',5')p] formation as a function of dsDNA length and found that dsDNA of 36 bp or longer were optimal, yet the 16 bp dsDNA used for crystallography elicited some activity (Figure S6A). Double-stranded RNA, a DNA/RNA duplex, or single-stranded DNA or RNA did not stimulate cyclic dinucleotide formation (Figure 5B). The trace amount of c[G(2',5')pA(3',5')p] detected for the specific ssDNA used in this experiment was attributable to a stretch of sequence complementarity, and the substitution of G by 8-oxoguanine (8-oxoG) was sufficient to destabilize its predicted interaction and eliminate the residual cGAS activity. Replacement of guanine by 8-oxoG within the dsDNA did not alter cGAS activity.

[0283] We quantified the activity of cGAS to yield c[G(2',5')pA(3',5')p] under multiple turnover conditions. Over 78% (s.d. +/- 2.6%, n=5) of the original ATP and GTP provided was converted to c[G(2',5')pA(3',5')p] within 40 min leading to an estimated observed rate constant of 0.19 min^{-1} and involving over 750 turnovers per enzyme molecule.

[0284] To determine the order of intermediate formation, we first substituted GTP by GMP or GDP (Figure 5C). Both compounds led to the formation of the respective 5'-pG(2',5')pA product and 5'-ppG(2',5')pA intermediates, and only 5'-ppG(2',5')pA could react

further to yield a reduced amount of the cylic-dinucleotide. Substitution of ATP by ADP or AMP resulted in no product or intermediate formation, with only ADP (with GTP) leading to the generation of reduced levels of c[G(2',5')pA(3',5')p].

[0285] In order to determine the involvement of 2' or 3'-hydroxyl (OH) groups of GTP and ATP for formation of c[G(2',5')pA(3',5')p], we tested 2'- and 3'-deoxyguanosine triphosphate and 2' and 3' and deoxyadenosine triphosphate as substrates for cGAS (Figure 5D and S6B). 2'-dGTP was unable to form a cyclic dinucleotide, unlike 3'-dGTP, indicating that the 2'-OH of guanosine was required for the formation of the linkage with the α-phosphate at the 5' position of adenosine. In contrast, both 2'- and 3'-dATP led to markedly reduced formation of cyclic GA-dinucleotides, although 2'-dATP + GTP yielded noticeably more product. In both cases, we observed accumulation of 5'-pppG(2',5')pdA reaction intermediates (Figure 5D, lanes 2 and 3), which migrated slightly faster than the all ribose intermediate (5'-pppG(2',5')pA, lane 1).

Example 3

Identification of c[G(2',5')pA(3',5')p] as the Product Formed by dsDNA-dependent cGAS Activity

[0286] The syntheses and purification of the three isomeric cGAMP molecules, c[G(2',5')pA(2',5')p] (2',5' linkages at GpA and ApG steps) 6, c[G(2',5')pA(3',5')p] (2',5' at GpA step and 3',5' at ApG step) 11, and c[G(3',5')pA(3',5')p] (3',5' at GpA and ApG steps) 15 shown in Figure S7 were carried out using procedures previously reported by the Jones laboratory (Gaffney et al. 2010; Gaffney and Jones, 2012). The identity of the three isomeric cGAMP molecules 6, 11 and 15 (Figure S7) was validated from heteronuclear NMR analysis using through-bond connectivities. The experimental NMR data for c[G(2',5')pA(3',5')p] is outlined in Figure S8, with the proton and carbon chemical shifts for all three isomeric GMP molecules listed in Table S4.

[0287] We analyzed the product generated by dsDNA-dependent cGAS activity using reverse-phase high-performance liquid chromatography (HPLC) and compared its elution profile to chemically synthesized c[G(3',5')pA(3',5')p], c[G(2',5')pA(2',5')p], and c[G(2',5')pA(3',5')p] compounds (Figure 6A and S6D). A prominent peak consistently eluted

from the HPLC system at precisely 23.5 min, which corresponded to the elution profile of c[G(2',5')pA(3',5')p]. Co-injection with c[G(3',5')pA(3',5')p] or c[G(2',5')pA(2',5')p] demonstrated that the cGAS reaction product does not co-elute, unlike co-injection with chemically synthesized c[G(2',5')pA(3',5')p].

[0288] To demonstrate that the *in vitro* produced cyclic dinucleotide matched the molecule determined crystallographically, we analyzed by HPLC the dissolved cGAS crystals that had been co-incubated with DNA, ATP, and GTP (Figure 6B). A peak corresponding to c[G(2',5')pA(3',5')p] was observed, distinct from the c[G(2',5')pA(2',5')p] co-injected reference molecule as before. Additional unidentified peaks of longer retention times were also seen. Presumably, these unidentified compounds originating from the crystallization buffer and/or additives were not completely removed despite washing the crystals prior to HPLC analysis.

[0289] In addition, cGAS-generated c[G(2',5')pA(3',5')p] was purified by HPLC and subjected to one-dimensional NMR analysis. Its NMR spectrum in the sugar H1' region is identical to that of chemically synthesized standard c[G(2',5')pA(3',5')p] and distinct from chemically synthesized c[G(2',5')pA(2',5')p] and c[G(3',5')pA(3',5')p] (Figure 6C). Thus, both HPLC (Figure 6A) and NMR (Figure 6C) independently validate that the product generated by cGAS is c[G(2',5')pA(3',5')p].

Example 4

Functional Analysis of cGAS Mutant Proteins

[0290] We next assayed the biochemical and functional consequences of mutations on cGAS in its capacity to form c[G(2',5')pA(3',5')p] *in vitro* and to stimulate the type I interferon pathway in cells. We generated alanine-substitution mutants corresponding to amino acid residues that the co-crystal structures revealed to be involved in dsDNA binding or cGAS activity. For *in vitro* cGAS activity assays, we generated and purified six recombinant mutant cGAS forms; four were predicted to eliminate dsDNA-binding and two point mutant proteins were substituted with alanine at potentially key catalytic residues. Incubation of DNA with mutant cGAS proteins led to little or no c[G(2',5')pA(3',5')p] formation for all but two mutants (R161A, S199A, Figure 7A).

[0291] To assess the impact on cGAS function in cells, we generated additional alanine mutants of cGAS for expression in mammalian cells. The full-length cGAS mutants together with STING and an IFN-β luciferase reporter were transiently expressed in HEK 293 cells. In this assay cGAS is engaged by the co-transfected DNA plasmids, and WT cGAS expression resulted in close to 15-fold enhanced luciferase activity compared to a control plasmid (Figure 7B). Single mutations of DNA binding residues, including Arg161 responsible for the only direct interaction with a DNA base, were not sufficient to impair cGAS activity. However, ablation of interactions with two or three consecutive phosphodiesters in either DNA strand (Figure 1C, S2B-C) resulted in diminished, or entirely abrogated cGAS function (Figure 7B). At the catalytic site, single mutants Glu211, Asp213 or Asp307 affecting the binding of divalent cations (Figure 2E, 3B) all resulted in non-functional cGAS (Figure 7C). Furthermore, abrogation of cGAS activity required mutation of both amino acid residues involved in (i) the binding of ATP (or GTP) gamma phosphate (Lys402, Ser420; Figure 2E, 3B), (ii) the binding of ATP adenosine (Glu371, Lys424; Figure 2F), or (iii) the base stacking of ATP and c[G(2',5')pA(3',5')p] (Arg364, Tyr 421; Fig 2F, 4B), while single mutants of these residues only slightly impaired cGAS function (Figure 7C). Gly198 and Ser199 are highly conserved residues that were found to undergo significant conformational changes upon ligand binding (Fig. 1F, 2D). Nevertheless, single mutations G198A and S199A did not severely impair cGAS function, but the double mutant of these positions was not functional (Figure S6E). Similarly, conversion of the highly mobile Gly198 to sterically restricted proline abrogated cGAS activity (Figure S6E).

Example 5

Studies on Conformational Transitions, Bond Formation, and Intermediates

Conformational Transitions in cGAS on Complex Formation

[0292] Our structural studies highlight the fact that cGAS undergoes a pronounced conformational change upon binding of dsDNA (Figure 1D), whereby it repositions catalytic residues Glu211, Asp213, Asp307, as well as Ser199 (Figure 1E, F), while at the same time opening access to the catalytic pocket (Figure 1G, H). In essence, cGAS adopts a catalytically competent conformation only when engaging dsDNA, thereby accounting for its role as a cytosolic dsDNA sensor. By contrast, only minimal conformational changes that are restricted to

the side chain of Glu211 are observed when proceeding from the binary complex of cGAS and dsDNA to the ternary complex with ATP (Figure 2B-D) and GTP (where the pocket contains off-pathway intermediate 5'-pppG(2',5')G; Figure S3A-C), with no change observed even for the side chain of Glu211 on formation of the ternary complex with ATP + GTP (where the pocket contains product c[G(2',5')pA(3',5')p]; Figure S4A-C).

Phosphodiester Bond Formation in Catalytic Pocket

[0293] The structural and functional experiments both established phosphodiester bond formation in the catalytic pocket after binding of dsDNA to cGAS, in the absence of any additional components. The structural studies on cGAS in the presence of dsDNA and GTP identified accumulation of linear reaction intermediate 5'-pppG(2',5')pG (Figure 3B,C), while in the presence GMP and ATP identified accumulation of linear reaction product 5'-pG(2',5')pA. While not wishing to be bound by any particular theory, it is believed that the former intermediate is off-pathway and therefore impaired for the formation of the second phosphodiester bond to form a cyclic product, given that the first G is *syn*, and the distance is long between the 2'-OH (or 3'-OH) of the second G and the α-phosphate of the triphosphate moiety. Nevertheless, these results suggest that formation of c[G(2',5')pA(3',5')p] is likely to occur in a stepwise manner, involving formation of sequential phosphodiester bonds to yield the cyclic dinucleotide product. By contrast, structural studies on cGAS in the presence of dsDNA, GTP, and ATP resulted in formation of c[G(2',5')pA(3',5')p] (Figure 4B, C), without accumulation of an intermediate and consistent with an on-pathway reaction involving formation of a pair of sequential phosphodiester linkages.

Positioning of G and A Residues of Bound c[G(2',5')pA(3',5')p]

[0294] The G and A residues of c[G(2',5')pA(3',5')p] adopt distinct positions in the structure of the ternary complex with cGAS and dsDNA. The A residue of the bound c[G(2',5')pA(3',5')p] is stacked on Tyr421 (Figure 4B) and occupies the position of the adenine ring in the ATP complex (Figure 2F) and the first base in the 5'-pppG(2',5')pG (Figure 3C) and 5'-pG(2',5')pA (Figure 3G) complexes. The A residue of bound c[G(2',5')pA(3',5')p] is not involved in any intermolecular hydrogen bonds and hence could potentially be replaced by even

a pyrimidine (C or U) residue. By contrast, the G residue of bound c[G(2',5')pA(3',5')p], which is partially stacked over the A residue, forms a network of intermolecular hydrogen bonds involving its Watson-Crick and Hoogsteen edges (Figure 4B, C) and cannot be replaced by any of the other three bases (C, A and U). Thus, the cGAS-binding pocket has distinct recognition elements that distinguish between G and A and hence can bind c[G(2',5')pA(3',5')p] in a unique orientation.

[0295] In agreement with the crystallographic data, the biochemical results indicate a strong preference for GTP, consistent with the elaborate amino acid interactions observed in the structure between cGAS and this base. Incubation of cGAS with GTP alone, as well as GTP plus either CTP or UTP, can lead to cyclical dinucleotide formation (Figure S6C). While ATP alone does not yield any cyclic or intermediate products, incubation with UTP results in some cyclic product formation, suggesting that UTP can also substitute for GTP albeit at a much reduced reaction rate. Together, these findings indicate that cGAS has more relaxed requirements for the second nucleotide compared to the first guanosine.

Structural Comparison of Linear 5'-pG(2',5')pA and Cyclic c[G(2',5')pA(3',5')p] Product

[0296] We observe a striking difference in alignment within the catalytic pocket between the off-pathway linear 5'-pG(*syn*)(2',5')pA(*anti*) product (Figure 3F, G) and cyclic 5'-pG(*anti*)(2',5')A(*anti*) product (Figure 4B, C). In the former case, it is the G base that stacks over Tyr421 (Figure 3G), while in the latter it is the A base that stacks over Tyr421 (Figure 4C). The two alignments are compared in stereo where the linear 5'-pG(2',5')pA is shown in Figure 4G and the cyclic c[G(2',5')pA(3',5')p] in Figure 4H. This implies that the intermediate may have to rearrange its orientation by a complete flip-over within the catalytic pocket prior to the cyclization reaction. This may not be too surprising since judging from the three basic (Glu211, Asp213 and Asp307) and one polar (Ser199) amino acid lining the catalytic pocket, there is only a single set of catalytic residues and hence following the first phosphodiester bond formation, the intermediate may have to realign so as to facilitate the second phosphodiester bond formation to complete cyclization.

Phosphodiester Linkages

[0297] A clear assumption in the earlier studies leading to the identification of cyclic GAMP as a second messenger generated by the cytoplasmic dsDNA sensor cGAS (Sun et al.

2013; Wu et al. 2013) was that both phosphodiester linkages were of a 3',5' nature. Such 3',5' linkages have been observed previously in structures of bacterial second messenger c-di-GMP bound to both STING (Yin et al 2010; Ouygang et al. 2012; Huang et al. 2012; Shu et al. 2012) and riboswitches (Smith et al. 2009; Kulshina et al. 2009). Nevertheless, the mass spectroscopic approach utilized (Wu et al. 2013) cannot distinguish between 3',5' and 2',5' linkages for one or both phosphodiester bonds of cyclic GAMP.

[0298] The first indication of a 2',5' linkage emerged from the structure of cGAS, dsDNA and GTP, where an off-pathway product formed in the catalytic pocket, exhibited a 5'-pppG(2',5')pG linkage (Figure 3B, C). In addition, pG(2',5')pA was observed in the structure of cGAS, dsDNA and GMP + ATP (Figure 3F, G). More importantly, a 2',5' linkage was also observed for the GpA step of the bound c[G(2',5')pA(3',5')p] product in the catalytic pocket in the structure of cGAS, dsDNA and GTP + ATP (Figure 4B, C).

[0299] Our initial biochemical analyses indicated that the 2',5' linkage between GTP and ATP occurs first, prior to the cyclization of the adenosine back to guanosine. This evidence was further supported by the observation that incubation of 2'-dATP with 2'-dGTP could not react to form any cyclic reaction products (Figure S6). In the second step, formation of a cyclical dinucleotide via the 2' or 3' OH of adenosine can proceed even when the other position is blocked through removal of oxygen, although there was an observable preference for utilization of 2'-dATP. Cyclic dinucleotide production in either case was very inefficient, suggesting that both positions may participate in the formation of a transition state for efficient phosphate hydrolysis and cyclization. This perplexing result, combined with ternary structural data concerning the connection from adenosine to guanosine, prompted us to further examine whether cGAS would ultimately have a preference for generating a 2',5' or 3',5' link for cyclization.

[0300] We observed a single HPLC peak, distinct from two cyclic-GA dinucleotide reference molecules (either both 2',5' or both 3',5' linkages) and coincident with c[G(2',5')pA(3',5')p], as the product of dsDNA-dependent cGAS activity (Figure 6A). This conclusion was validated from an independent NMR study (Figure 6C). While biological production of cyclic dinucleotides appears evolutionarily conserved from prokaryotes to eukaryotes, their formation based on the chemical linkages is distinct. In the case of cGAS, formation of c[G(2',5')pA(3',5')p] appears similar to the 2',5' oligoadenylates generated by

OAS but also to the 3',5' dinucleotide linkages created by bacterial cyclases (Sadler and Williams, 2008; Kodym et al. 2009; Donovan et al. 2013).

Potential benefits of 2',5' (GA step) and 3',5' (AG step) linkages in c[G(2',5')pA(3',5')p]

[0301] It is not clear why cGAS prefers to generate both a 2',5' (GpA step) and 3',5' (ApG step) cyclic GA-dinucleotide. A 2',5' phosphodiester bond is uncommon and few nucleases are known to be able to hydrolyze such a linkage (Kubota et al. 2004). Without wishing to be bound by any particular theory, 2',5' linkages might promote greater stability in cells to enable effective transduction of the second messenger, but the 3',5' linkage may facilitate its breakdown by numerous conventional endonucleases to prevent prolonged interferon response. Taken together, our structure and functional studies have identified the chemical nature of metazoan cGAMP, highlighting the role of 2',5' linkages in second messengers that activate the type I interferon pathway.

Implications of cGAS dsDNA-binding mutants

[0302] Structural studies have identified intermolecular protein-DNA contacts on formation of the cGAS-dsDNA complex (Figure 1C). Since these are primarily of an electrostatic nature and involve non-specific recognition of the DNA phosphodiester backbone, they have been classified into three sets of triple mutants, with the *in vitro* (Figure 7A) and cellular assays (Figure 7B) establishing complete loss in activity and ability to stimulate interferon production for the S165A, N172A, K372A triple mutant, the N196A, Y200A, K372A triple mutant and the R158A, R161A, K395A triple mutant, and partial loss in activity for the S165A, N172A, Y200A triple mutant. This reinforces the importance of complex formation between cGAS and dsDNA for the catalytic activity of cGAS.

Implications of cGAS Catalytic Pocket Mutants

[0303] Structural studies of ternary complexes of cGAS and dsDNA with bound NTPs have identified Glu211, Asp213, Asp307 as important catalytic residues for phophodiester bond formation. All three catalytic acidic residues are functionally dead on replacement by Ala as observed in either *in vitro* (Figure 7A; Glu211A) or cellular (Figure 7C, all three catalytic

residues) assays, while the S199A mutation retained substantial activity. Tyr421 is involved in stacking interactions with A, while Arg364 is hydrogen bonded with G in the cGAMP ternary complex (Figure 4B, C). Dual mutation of Y421A, R364A result in loss in the majority of activity in cellular assays (Figure 7C).

Role of Divalent Cations

[0304] Structural studies of ternary complexes of dsDNA-bound cGAS with ligands have shown that the triphosphate moieties of ATP (Figure 2E, F) and 5'-pppG(2',5')pG (Figure 3B, C) are coordinated to a pair of cations (tentatively assigned to Mg^{2+}). Indeed, functional studies have highlighted the importance of divalent cations to phosphodiester bond formation. Omission of divalent cations or use of EDTA prevented c[G(2',5')pA(3',5')p] formation, whereas Mg^{2+} and Mn^{2+} promoted cGAS activity (Figure 5A).

Comparison with Cytoplasmic dsRNA Sensor OAS1

[0305] In a parallel study to our contribution, structural studies and biochemical assays have been recently reported on the characterization of the dsRNA sensor human oligoadenylate synthetase 1 (OAS1) which polymerizes ATP into linear 2',5'-linked oligoadenylate (Donovan et al. 2013). The crystallographic studies unequivocally demonstrated conformational transitions in OAS1 on proceeding from the free state (Hartmann et al. 2003) to the ternary complex with bound dsRNA and 2'-dATP (Donovan et al. 2013), which follow a similar pattern to those observed by us in this study for complex formation of cGAS with dsDNA and bound ligands. Thus, three catalytic Glu residues of OAS1 are brought into close proximity on formation of the ternary complex with dsRNA and 2'-dATP, thereby creating the coordination geometry for binding of two Mg^{2+} ions and 2'-dATP (Donovan et al. 2013), similar to what we observe for the cGAS system. Given that the only available structures were for free OAS1 (Hartmann et al. 2003) and its ternary complex with dsRNA and bound ligand (Donovan et al. 2013), these authors were not in a position to determine how the conformational transition was partitioned between steps reflecting conversion from free OAS1 to the binary complex with dsRNA and conversion from the binary complex to the ternary complex in the presence of 2'-dATP. Our results on the cytosolic dsDNA sensor cGAS suggest that the major conformational transition

will most likely be restricted for the step involving conversion of OAS1 from the free state to the dsRNA-binding complex, with minimal changes on addition of 2'-dATP to form the ternary complex.

[0306] In addition to similarities mentioned above, there are also differences in protein-nucleic acid recognition principles between the cGAS dsDNA sensor (our study) and the OAS1 dsRNA sensor (Donovan et al. 2013), in that cGAS targets the sugar-phosphate backbone of dsDNA within a central segment of the dsDNA duplex (Figure 1C), while OAS1 targets the sugar-phosphate backbone of dsRNA by contacting two minor groove segments separated by 30 Å (Donovan et al. 2013). The helical parameters of dsRNA and dsDNA are very distinct, and different recognition principles are used in protein-dsRNA (reviewed in Lunde et al. 2007) and protein-dsDNA (reviewed in Huffman and Brennan, 2002) complexes. Nevertheless, common principles are utilized to generate the critical catalytic site architecture, which in turn couples nucleic acid recognition (dsRNA or dsDNA in the cytoplasm) with the cascade of downstream events leading to an antiviral state including type I interferon response (cGAS) and RNase L activation (OAS1).

[0307] Further, the formation of linear 2',5'-linked iso-RNA mediated by OAS1 parallels the formation of c[G(2',5')pA(3',5')p] containing 2',5' linkage at the GpA step by cGAS. Thus, unlike earlier sole emphasis on 3',5' linkages as observed previously for bacterial second messenger c-di-GMP (reviewed in Romling et al. 2013), we highlight that the metazoan second messenger c[G(2',5')pA(3',5')p] utilizes mixed linkages involving 2',5' at the GpA step and 3',5' at the ApG step.

cGAS Contains a Single Active Site for Step-wise Phosphodiester Bond Formation

[0308] Previous studies have established that diguanylate cyclase PleD forms a head-to-tail homodimer to form a reaction center at its interface, so that the intermediate does not have to change its orientation on the pathway to form c-di-GMP (Chan et al. 2004). By contrast, in our current studies of ternary complexes of cGAS, dsDNA and bound ligands, we observe no evidence for dimer or higher order oligomer formation in the crystal. Further, the ligand-binding pocket in our structures is buried within the cGAS topology and is not located on the surface, as it is in PleD (Chan et al. 2004).

[0309] Indeed, cGAS contains a single active site for the sequential phosphodiester bond formation steps, a feature quite remarkable given that the ligands are GTP and ATP and that the GpA linkage which forms first is 2'-5' and the ApG linkage which forms second is 3',5', resulting in generation of c[G(2',5')pA(3',5')p]. Without wishing to be bound by any particular theory, we outline a model for formation of c[G(2',5')pA(3',5')p] from GTP and ATP within the single catalytic pocket of dsDNA-bound cGAS in Figure 7D. In this model, the first step involves formation of a 5'-pppGpA intermediate followed in the second step by formation of c[G(2',5')pA(3',5')p]. Note, also that the bound ligand is believed to undergo two flip-overs on the pathway to c[G(2',5')pA(3',5')p] formation.

Implications for the Di-nucleotide Cyclase DncV

[0310] In an earlier study, the bacterial dinucleotide cyclase DncV was shown to generate cyclic GMP-AMP (cGAMP) (Davies et al. 2012). This first report on formation of cGAMP raises the interesting question as to the nature of the pair of phosphodiester linkages in this bacterial system.

Example 6

NMR Spectral Analysis of Synthesized cGAMP Linkage Isomers

[0311] Lyophilized cGAMP linkage isomers were dissolved in 99.9% D_2O in 10mM K_2HPO_4-KH_2PO_4 (pH 6.6) buffer. All NMR experiments are conducted at 35°C on a Bruker 900 MHz spectrometer at New York Structural Biology Center. Resonance assignments are made based on HMBC (2 s recycling delay, 0.8 s ^1H acquisition time, 20 ms ^{13}C acquisition time, phase-insensitive ^{13}C acquisition, and anti-phase ^1H detection with absolute value mode processing), double-quantum filtered COSY (2 s recycling delay, 0.8 s direct acquisition time, 12 ms indirect acquisition time), and HSQC experiments (1 s recycling delay, 48 ms ^1H acquisition time, 20 ms ^{13}C acquisition time). The 1D proton spectra with water presaturation are accumulated over 8 scans for the synthesized cGAMP linkage isomers standards and 128 scans for the bio-enzymatically produced cGAS reaction.

Example 7

Thin Layer Chromatography (TLC) Analysis

Preparation of Oligonucleotides for TLC assays

[0312] Oligonucleotides used for biochemical assays of cGAS nucleotidyltransferase activity are listed in Table S6. Oligodeoxynucleotides were synthesized in-house using a 3400 DNA synthesizer (Applied Biosystems), oligoribonucleotides were purchased (Dharmacon). Double-stranded DNA, RNA, and DNA/RNA duplexes were annealed in 70 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, at equimolar concentrations by incubation initiated at 95 °C followed by a 0.1 °C decrease per second to 25 °C in a Peltier thermoycler (MJ Research), and verified for annealing by agarose gel electrophoresis prior to use.

TLC Analysis of c[G(2',5')pA(3',5')p] Formation

[0313] Purified recombinant full length (fl, amino acids 1-507) and truncated (tr, amino acids 147-507) murine cGAS, including truncated mutant versions 1-6, were incubated in 20 μL reactions containing: 1 μM cGAS, 3.3 μM dsDNA, 5 mM $MgCl_2$, 150 mM NaCl, 20 mM Tris-HCl, pH 7.5 at 25°C, 1 mM DTT, 10% glycerol, 1 mM each of nucleotides (typically ATP and GTP), and α^{32}P or γ^{32}P radiolabelled NTPs or dNTPs at 37°C for 40 min. Reactions were stopped by addition of 20 μL of 50 mM EDTA. 2 μL of reaction solution was spotted onto high-performance TLC plates (HPTLC silica gel, 60 Å pores, F_{254}, 10x10 cm, cat # 1.05628.0001, EMD Millipore) and products were separated with Solvent 1 (NH_4HCO_3:C_2H_5OH:H_2O [0.2 M:30%:70%], w:v:v) or 2 (NH_4HCO_3:C_2H_5OH:H_2O [0.025 M:30%:70%], w:v:v) at 25 °C for 1 h. Reaction products were visualized by UV (254 nm) and phosphorimaging (Typhoon FLA 9500, GE Healthcare). Images were processed using Adobe Photoshop and Illustrator CS5. The TLC conditions used were largely based on a protocol established to separate 3',5' cAMP (Higashida et al, 2012).

Example 8

Quantitation of cGAS reaction products

[0314] The yield of c[G(2',5')pA(3',5')p] generated was calculated by densitometry analysis of TLC experiments, using FIJI (ImageJ 1.47i) or spectrophotemetrically (absorbance at 260 nm, E_{260}=25.4 x 10^3) after purification from HPLC. For densitometry analyses, the fraction of α^{32}P-labelled c[G(2',5')pA(3',5')p] over total radioactivity per lane (c[G(2',5')pA(3',5')p] plus remaining α^{32}P-labelled ATP or GTP) was calculated.

Example 9

Preparation of cGAS reaction products for high-performance liquid chromatography analysis

[0315] *In vitro* generated c[G(2',5')pA(3',5')p] reaction products, or washed and dissolved cGAS crystals, were treated with 25 units of Benzonase (Novagen, cat. # 70746, Purity > 90%) for 30 min at 37°C, heat inactivated for 10 min at 95°C, then centrifuged at 21,000 g for 15 min (Sorvall Legend Micro 21R, Thermo Scientific); the supernatant was used for HPLC analysis. Reaction products, with or without 3-8 nmoles of chemically synthesized all 3',5' cGAMP, all 2',5' cGAMP, or c[G(2',5')pA(3',5')p] were subjected to reverse-phase HPLC analysis (AKTA Purifier, GE Healthcare) using a C18 column (25 cm x 4.5 mm, 5 μM pore, Supelco Analytical). Analytes were monitored by UV 260 and 280 nm. A 0-10% solvent B (2 column volume), 10-50% solvent B (2 column volume) two-step linear gradient was used; solvent A (triethylammonium acetate:acetonitrile:H_2O [0.1 M:3%:97%], w:v:v) and solvent B (methanol:acetonitrile:H_2O [45%:45%:10:], v:v:v).

Preparation of cGAS reaction product for 1D NMR analysis

[0316] 100 μl of *in vitro* generated c[G(2',5')pA(3',5')p] reaction product was benzonase and heat treated as before, prior to fractionation by HPLC. Three serial HPLC runs were performed (two 40 μl and one 20 μl reaction injection), and the peak corresponding to c[G(2',5')pA(3',5')p] was collected into a 15 ml falcon tube (approx. 4.5 ml total). Solvent removal was accomplished by vacuum centrifuge (Vacufuge, Eppendorf) for 3 days at room temperature until completely dry.

Example 10

Cellular Assays

Generation of cGAS Point Mutants

[0317] The murine cGAS CDS was inserted into a modified pMAX-cloning vector (Amaxa, Cologne, Germany). Site-directed mutagenesis was performed using the Quikchange method (Agilent, Santa Clara, CA) using Pfu Ultra Hot Start DNA Polymerase (Agilent) or KOD Hot Start DNA Polymerase (Merck, Darmstadt, Germany). The murine STING CDS and Firefly Luciferase (Promega, Madison, WI) were cloned into an EF1-promoter- modified pLenti6

(Invitrogen, Carlsbad, CA) expression plasmid. pGL3 IFN-beta Gluc reporter was obtained from Brian Monks (Institute of Innate Immunity, University of Bonn, Germany). All constructs were verified by sequencing of the CDS.

Luciferase assay

[0318] $3x10^4$ HEK293 cells per 96-well were reverse-transfected in triplicates with a mixture of pGL3-IFNbeta-Gluc (50 ng), pLenti-EF1-Fluc (25 ng), pLenti-EF1-mSTING (25 ng) and cGAS-expression plasmid (25ng, pMAX-cGAS WT or mutants) or Control plasmid pMAX-GFP (Amaxa) using Trans-IT LT1 (MirusBio, Madison, WI). After 36 h cells were lysed in passive lysis buffer. Firefly and gaussia Luciferase activities were determined on an EnVision reader (Perkin Elmer, Waltham, MA) using the respective substrates D-luciferin and coelenterazine(PJK GmbH, Kleinblittersdorf, Germany). IFNbeta-Gluc values were normalized to constitutive firefly luciferase values and fold induction was calculated in relation to control-plasmid pMAX-GFP.

Example 11

Synthesis of Cyclic GA-dinucelotides

[0319] Preparation of all 2',5'-cGAMP (6, Figure S7), c[G(2',5')pA(3',5')p] (11, Figure S7), and all 3',5'-cGAMP (15, Figure S7) were carried out using the procedure previously reported by the Jones laboratory (Gaffney et al. 2010; Gaffney and Jones 2012). To adenosine phosphoramidite, 1 or 7, (0.784 g, 0.793 mmol) dissolved in 5 mL of CH_3CN and water (0.028 mL, 1.6 mmol, 2 equiv) was added pyridinium trifluoroacetate (0.184 g, 0.95 mmol, 1.2 equiv). After 1 min, 6 mL of *tert*-$BuNH_2$ was added. After another 10 min, the mixture was concentrated. To the residue dissolved in 10 mL of CH_2Cl_2 was added H_2O (0.14 mL, 7.9 mmol, 10 equiv), followed by 10 mL of 6% dichloroacetic acid (DCA, 7.5 mmol) in CH_2Cl_2. After 10 min, the reaction was quenched by addition of pyridine (1.2 mL, 15 mmol, 2 equiv rel to DCA). The mixture was then concentrated, and the residue was dissolved in 7 mL of CH_3CN and concentrated again.

[0320] This process was repeated two more times, the last time leaving the A H-phosphonate, 2 or 8, in 2 mL. To this solution was added a dried solution of G amidite, 3 or 12 (1.00g, 1.03 mmol, 1.3 equiv) in 3 mL CH_3CN. After 2 min, anhydrous *tert*-butyl hydroperoxide

5.5 M in decane (0.43 mL, 2.4 mmol, 3 equiv) was added. After 30 min, 0.20 g of $NaHSO_3$ dissolved in 0.5 mL H_2O was added. The mixture was stirred for 5 min, and then concentrated. The residual oil was dissolved in 14 mL of CH_2Cl_2, followed by addition of H_2O (0.15 mL, 8.5 mmol, 10 equiv) and then 14 mL of 6% DCA (9.8 mmol) in CH_2Cl_2. After 10 min, the reaction was quenched with 9 mL of pyridine. The mixture was concentrated to a small volume, 25 mL more pyridine was added, and the solution was concentrated again, leaving the linear dimer, 4, 9, or 13, in 17 mL. To this solution was added 5,5-dimethyl-2-oxo-2-chloro-1,3,2-dioxaphosphinane (DMOCP, 0.54 g of 95% reagent, 2.8 mmol, 3.5 equiv). After 10 min, the reaction was quenched by addition of H_2O (0.50 mL, 28 mmol, 10 equiv rel to DMOCP), and I_2 (0.26 g, 1.0 mmol, 1.3 equiv) was added immediately. After 5 min, the mixture was poured into 120 mL of H_2O containing 0.17 g $NaHSO_3$. After 5 min of stirring, 3.4 g of $NaHCO_3$ was slowly added. After 5 min more of stirring, the aqueous solution containing solid was partitioned with 135 mL 1:1 EtOAc:Et_2O. The separated aqueous layer was then partitioned with an additional 35 mL of 1:1 EtOAc:Et_2O.

[0321] The organic layers containing 5, 10, or 14 were combined and concentrated to an oil. For 14, the oil was dissolved in 5 mL CH_3CN and the cyanoethyl group was removed by addition of 5 mL of *tert*-$BuNH_2$ for 10 min. The residue was purified on a 80 g SiO_2 column, using a gradient of 0 to 25% CH_3OH in CH_2Cl_2 over 50 min. 5 and 10 were directly purified on SiO_2 without *tert*-$BuNH_2$ treatment. In each case the residue after purification was treated with 21 mL of CH_3NH_2 in anhydrous EtOH (33% by weight, 168 mmol, 212 equiv rel to the amino protecting groups). After 4 h at room temperature, the mixture was concentrated to a solid, to which 3 mL of pyridine and 1 mL of Et_3N were added. The mixture was concentrated to an oil, and this process was repeated two more times to convert the *tert*-$BuNH_3^+$ to the Et_3NH^+ salt. To the oil was added 1mL of pyridine, and the flask was placed in an oil bath at 55 °C. Et_3N (7.5 mL) and Et_3N•3HF (2.6 mL, 48 mmol F^-, 30 eq rel to each TBS) were added simultaneously. The mixture was stirred at 55 °C. After 3 h, the flask was removed from the oil bath and HPLC grade acetone (70 ml) was slowly added to the stirring mixture. After 10 min, the solid was collected by filtration, washed 5x with 3 mL portions of acetone, and dried in a desiccator over KOH overnight. This process gave pure 15, but 6 and 11 were purified on a 19x300 mm Prep Nova-Pak C18 column using a gradient of 2 to 20% CH_3CN in 0.1 M NH_4HCO_3.

[0322] Analytical reversed phase HPLC was carried out on a Waters 2965 system with a photodiode array detector, using an Atlantis C18 column, 100 Å, 4.6 mm × 50 mm, 3.0 μm. Gradients of CH_3CN and 0.1 M triethylammonium acetate buffer (pH 6.8) were used with a flow rate of 1.0 mL/min. Low resolution ESI-MS was routinely acquired using a Waters Micromass single quadrupole LCZ system. LCMS of 6, 11, and 15 displayed *m/z* (M-H) 673 (calculated for $C_{20}H_{23}N_{10}O_{13}P_2^-$: 673).

Example 12

STING-dependent induction of murine alpha-interferon and human CXCL10 by cGAMP compounds.

THP-1 culture and assay conditions

[0323] THP-1 cells were cultured in RPMI1640 with 10% FBS, sodium pyruvate and penicillin/streptomycin (Gibco, Life Technologies). 8x104 cells were plated per 96-well in 100 μl of Medium and equilibrated for 2 h at 37°C/5% CO2. To generate macrophage-like cells, 8x104 THP-1 cells were differentiated overnight with 10 ng/ml PMA (Sigma), medium was changed and cells were incubated for additional 24 h prior to stimulation.

BMDM culture and assay conditions

[0324] Bone marrow derived macrophage cells (BMDM) were flushed from femurs of C57BL/6 mice. Erythrocytes were lysed (PharmLyse, BD Biosciences) and 1x107 cells per Petri dish were incubated in DMEM 10% FBS, sodium pyruvate and penicillin/streptomycin (Gibco, Life Technologies) with 30% L929-supernatant for 7 days . Cells were harvested with PBS 2 mM EDTA and plated at a density of 1x105 cells per 96-well. BMDMs were digitonin-permeabilized or control-treated in the presence of indicated cGAMP concentrations for 30 min, then supplemented with fresh medium. Supernatants were taken after 18 h and cytokines were determined by ELISA.

Cell permeabilization/stimulation

[0325] Cell permeabilization for delivery of cyclic di-nucleotides was performed as previously described(Woodward et al., 2010). Briefly, supernatant was removed and cells were covered with 50 μl Perm-buffer (50 mM HEPES pH 7.0, 100 mM KCl, 3 mM MgCl2, 1 mM ATP, 0.1 mM GTP, 0.1 mM DTT, 85 mM sucrose, 0.2% BSA) +/- 10 μg/ml digitonin and serial

dilutions of cGAMP isomers, followed by 30 min incubation at 37°C. Perm-buffer was then removed and cells were covered with 100 µl of pre-warmed medium. Viability of permeabilized cells was >50% compared to untreated, as monitored by light microscopy and Cell Titer Blue (Roche).

[0326] Supernatants were collected 16 h after stimulation and cytokines were determined by ELISA and HEK-Blue™ IFN-α/β bioassay, respectively.

ELISA

[0327] Human CXCL-10 was determined by ELISA (BD Opteia human IP-10 ELISA-Set) according to manufacturer's recommendations. Murine Ifna was determined by sandwich-ELISA: Monoclonal rat-anti Ifna (clone RMMA-1) was used as capture antibody, recombinant Ifna was used as standard and polyclonal rabbit serum against Ifna for detection (all from PBL Interferon Source, Piscataway NJ, USA), followed by anti-rabbit HRP (Bio-Rad).

Fitting of dose-response curves

[0328] 4-parametric sigmoidal dose-response curves and EC50 values were analyzed with Graph Pad Prism (Graph Pad Software, La Jolla CA, USA).

Table S4. Proton and carbon chemical shifts for cGAMPs

Proton chemical shifts list

		c[G(3',5')pA(3',5')p]	c[G(2',5')pA(2',5')p]	c[G(2',5')pA(3',5')p]
G	H8	7.96	7.85	7.86
	H1'	5.92	5.99	5.93
	H2'	4.72	5.31	5.62
	H3'	4.91	4.66	4.58
	H4'	4.39	4.45	4.39
	H5'	4.35	4.13	4.15
	H5''	4.07	4.21	4.22
A	H8	8.35	8.21	8.31
	H2	8.13	8.11	8.27

	H1’	6.11	6.29	6.17
	H2’	4.76	5.23	4.77
	H3’	4.92	4.69	5.03
	H4’	4.45	4.51	4.47
	H5’	4.38	4.16	4.45
	H5’’	4.09	4.24	4.13

Carbon chemical shifts list

		c[G(3’,5’)pA(3’,5’)p]	c[G(2’,5’)pA(2’,5’)p]	c[G(2’,5’)pA(3’,5’)p]
G	C8	139.72	142.36	143.29
	C5	118.97	119.88	120.06
	C4	153.09	154.55	154.53
	C1’	92.23	89.13	88.90
	C2’	76.14	78.77	77.11
	C3’	73.18	73.87	74.05
	C4’	82.58	86.18	86.11
	C5’	65.03	67.59	68.38
A	C8	142.13	142.57	141.61
	C6	157.97	157.89	157.97
	C5	121.44	121.10	121.47
	C4	150.63	152.22	150.52
	C2	155.32	155.08	155.05
	C1’	92.56	87.63	92.36
	C2’	76.23	82.04	76.58
	C3’	73.16	74.75	73.17
	C4’	82.75	87.03	82.68
	C5’	65.05	68.01	64.89

Table S5. Primers used in the Quickchange mutagenesis PCRs (Metabion, Martinsried, D):

Primer Name	Sequence (5'-3')
N196A_Y200A	caaaggtgtggagcagctggccactggcagcgcctatgaacatgtgaagatt
N196A_Y200A_antisense	aatcttcacatgttcataggcgctgccagtggccagctgctccacacctttg
K372A	cctctctttctctcacactgaagcgtacattttgaataatcacggg
K372A_antisense	cccgtgattattcaaaatgtacgcttcagtgtgagagaaagagagg
S165A	ttgaaacgcaaagatatcgcggaggcggccg
S165A_antisense	cggccgcctccgcgatatctttgcgtttcaa
N172A	ggcggccgagacggtggctaaagttgtggaacgc
N172A_antisense	gcgttccacaactttagccaccgtctcggccgcc
Y200A	gcagctgaacactggcagcgcctatgaacatgtgaagatt
Y200A_antisense	aatcttcacatgttcataggcgctgccagtgttcagctgc
R158A_R161A	agaaggtgctggacaaattggcattgaaagccaaagatatctcggaggcgg
R158A_R161A_antisense	ccgcctccgagatatctttggctttcaatgccaatttgtccagcaccttct
K395A	aatcttccggagcaaaatgctgcagagcagaatgtttaaaattaatgaaatacc
K395A_antisense	ggtatttcattaattttaaacattctgctctgcagcattttgctccggaagatt
R161A	ggacaaattgagattgaaagccaaagatatctcggaggcg

R161A_antisense	cgcctccgagatatctttggctttcaatctcaatttgtcc
G198P_S199A	gtggagcagctgaacactgccgcctactatgaacatgtgaag
G198P_S199A_antisense	cttcacatgttcatagtaggcggcagtgttcagctgctccac
G198A	ggagcagctgaacactgccagctactatgaacatg
G198A_antisense	catgttcatagtagctggcagtgttcagctgctcc
G198P	gtggagcagctgaacactcccagctactatgaacatgt
G198P_antisense	acatgttcatagtagctgggagtgttcagctgctccac
G198P_G199A	ggtgtggagcagctgaacactcccgcctactatgaacatgtgaagatt
G198P_S199A_antisense	aatcttcacatgttcatagtaggcgggagtgttcagctgctccacacc
S199A	ggagcagctgaacactggcgcctactatgaacatgtgaag
S199A_antisense	cttcacatgttcatagtaggcgccagtgttcagctgctcc
E211A	tgtgaagatttctgctcctaatgcatttgatgttatgtttaaactgg
E211A_antisense	ccagtttaaacataacatcaaatgcattaggagcagaaatcttcaca
K402A	gcaaaatgctgcagaaaagaatgtttaaaattaatggcatacctttggaacagttgaaaaaa
K402A_antisense	ttttttcaactgttccaaaaggtatgccattaattttaaacattcttttctgcagcattttgc
S420A	ttcaagagctggatgcattctgtgcctaccatgtga
S420A_antisense	tcacatggtaggcacagaatgcatccagctcttgaa
E371A	gcctctctttctctcacactgcaaagtacattttgaataatcac
E371A_antisense	gtgattattcaaaatgtactttgcagtgtgagagaaagagaggc

K424A	gcattctgttcctaccatgtggcaactgccatctttcacatgtg
K424A_antisense	cacatgtgaaagatggcagttgccacatggtaggaacagaatgc
R364A	tcaaggagagacctgggccctctctttctctcac
R364A_antisense	gtgagagaaagagagggcccaggtctctccttga
Y421A	gctggatgcattctgttccgcccatgtgaaaactgccatc
Y421A_antisense	gatggcagttttcacatgggcggaacagaatgcatccagc
mcGASfwXhoI	atatatctcgagatggaagatccgcgtagaagga
mcGASrevBglII	atatatagatctctatcaaagcttgtcaaaaattggaaacccat

Table S6.Oligonucleotides utilized for TLC analyses:

Oligonucleotide Name	Sequence (5'-3')
45 mer DNA (top strand)	tacagatctactagtgatctatgactgatctgtacatgatctaca
45 mer DNA (bottom strand)	tgtagatcatgtacagatcagtcatagatcactagtagatctgta
45 mer RNA (top strand)	uacagaucuacuagugaucuaugacugaucuguacaugaucuaca
45 mer RNA (bottom strand)	uguagaucauguacagaucagucauagaucacuaguagaucugua

17 mer DNA (top strand)	aaattgccgaagacgaa
17 mer DNA (bottom strand)	tttcgtcttcggcaatt
36 mer DNA (top strand)	acacacacacacacacacacacacacacacacac
36 mer DNA (bottom strand)	ctctctctctctctctctctctctctctctct

[0329] Bold and underlined nucleotides represent 8-oxoguanosines that were utilized in separately generated modified oligos. Oligonucleotides were synthesized in-house 3400 DNA synthesizer (Applied Biosystems) or purchased (Dharmacon).

REFERENCES

[0330] Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213-221.

[0331] Burckstummer, T., Baumann, C., Blumi, S., Dixit, E., Durnberger, G., Jahn, H., Planyavsky, M., Bilban, M., Colinge, J., Bennet, 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.

[0332] Chan, C., Paul, R., Samoray, D., Amiot, N.C., Giese, B., Jenal, U. and Schirmer, T. (2004). Structural basis of activity and allosteric control of diguanylate cyclase. Proc. Natl. Acad. Scis. USA. *101*, 17084-17089.

[0333] 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.

[0334] Donovan, J., Dufner, M., and Korennykh, A. (2013). Structural basis for cytosolic double-stranded RNA surveillance by human oligoadenylate synthetase 1. Proc. Natl. Acad. Sci. USA *110*, 1652-1657.

[0335] Egli, M., Gessner, R.V., Williams, L.D., Quigley, G.J., van der Marel, G.A., van Boom, J.H., Rich, A., and Frederick, C.A. (1990). Atomic-resolution structure of the cellulose synthase regulator cyclic diguanylic acid. Proc. Natl. Acad. Sci. USA *87*, 3235-3239.

[0336] Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486-501.

[0337] 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.

[0338] Gaffney, B.L., and Jones, R.A. (2012). One-flask syntheses of cyclic diguanosine monophosphate (c-di-GMP). Current Protocols in Nucleic Acid Chemistry *14*, 14.18.11-14.18.17.

[0339] Gaffney, B.L., Veliath, E., Zhao, J., and Jones, R.A. (2010). One-flask syntheses of c-di-GMP and the [*R*p,*R*p] and [*R*p,*S*p] thiophosphate analogues. Org Lett *12*, 3269-3271.

[0340] Hartmann, R., Justesen, J., Sarkar, S.N., Sen, G.C., and Yee, V.C. (2003). Crystal structure of the 2'-specific and double-stranded RNA-activated interferon-induced antiviral protein 2'-5'-oligoadenylate synthetase. Mol. Cell *12*, 1173-1185.

[0341] Higashida, H., Hossain, K.Z., Takahagi, H., and Noda, M. (2002). Measurement of adenylyl cyclase by separating cyclic AMP on silica gel thin-layer chromatography. Anal. Biochem. *308*, 106-111.

[0342] 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.

[0343] Hornung, V., and Latz, E. (2010). Intracellular DNA recognition. Nat. Rev. Immunol. *10*, 123-130.

[0344] Huang, Y.H., Liu, X.Y., Du, X.X., Jiang, Z.F., and Su, X.D. (2012). The structural basis for the sensing and binding of cyclic di-GMP by STING. Nat Struct Mol Biol *19*, 728-730.

[0345] Huffman, J.L. and Brennan, R.G. (2002). Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *12*, 98-106.

[0346] Ishikawa, H., and Barber, G.N. (2008). STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. Nature *455*, 674-678.

[0347] Jin, L., Waterman, P.M., Jonscher, K.R., Short, C.M., Reisdorph, N.A., and Cambier, J.C. (2008). MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex class II and mediates transduction of apoptotic signals. Mol. Cell Biol. *28*, 5014-5026.

[0348] Jin, T., Perry, A., Jiang, J., Smith, P., Curry, J.A., Unterholzner, L., Jiang, Z., Horvath, G., Rathinam, V.A., Johnstone, R.W., *et al.* (2012). Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor. Immunity *36*, 561-571.

[0349] Keating, S.E., Baran, M., and Bowie, A.G. (2011). Cytosolic DNA sensors regulating type I interferon induction. Trends Immunol. *32*, 574-581.

[0350] Kerur, N., Veettil, M.V., Sharma-Walia, N., Bottero, V., Sadagopan, S., Otageri, P., and Chandran, B. (2011). IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection. Cell Host Microbe *9*, 363-375.

[0351] Kodym, R., Kodym, E., and Story, M.D. (2009). 2'-5'-Oligoadenylate synthetase is activated by a specific RNA sequence motif. Biochem. Biophys. Res. Commun. *388*, 317-322.

[0352] Krasteva, P.V., Giglio, K.M., and Sondermann, H. (2012). Sensing the messenger: the diverse ways that bacteria signal through c-di-GMP. Protein Sci. *21*, 929-948.

[0353] Kubota, K., Nakahara, K., Ohtsuka, T., Yoshida, S., Kawaguchi, J., Fujita, Y., Ozeki, Y., Hara, A., Yoshimura, C., Furukawa, H. et al. Identification of 2'-phophodiesterase, which plays a role in 2-5A system reulated by interferon. J. Biol. Chem, *279*, 37832-37841.

[0354] Kulshina, N., Baird, N.J., and Ferre-D'Amare, A.R. (2009). Recognition of the bacterial second messenger cyclic diguanylate by its cognate riboswitch. Nat. Struct. Mol. Biol. *16*, 1212-1217.

[0355] Lee, E.R., Baker, J.L., Weinberg, Z., Sudarsan, N., and Breaker, R.R. (2010). An allosteric self-splicing ribozyme triggered by a bacterial second messenger. Science *329*, 845-848.

[0356] Lunde, B.M., Moore, C. and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nat. Rev. Mol Cell. Biol. 8, 479-490.

[0357] McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658-674.

[0358] O'Neill, L.A. (2013). Immunology. Sensing the dark side of DNA. Science *339*, 763-764.

[0359] Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods in Enzymology *276*, 307-326.

[0360] 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.

[0361] Romling, U., Galperin, M.Y., and Gomelsky, M. (2013). Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. Microbiol. Mol. Biol. Rev. *77*, 1-52.

[0362] Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G.A., van Boom, J.H., *et al.* (1987). Regulation of cellulose synthesis in Acetobacter xylinum by cyclic diguanylic acid. Nature *325*, 279-281.

[0363] Sadler, A.J., and Williams, B.R. (2008). Interferon-inducible antiviral effectors. Nat. Rev. Immunol. *8*, 559-568.

[0364] Schirmer, T., and Jenal, U. (2009). Structural and mechanistic determinants of c-di-GMP signalling. Nat. Rev. Microbiol. *7*, 724-735.

[0365] 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. Nat. Struct. Mol. Biol. *19*, 722-724.

[0366] Smith, K.D., Lipchock, S.V., Ames, T.D., Wang, J., Breaker, R.R., and Strobel, S.A. (2009). Structural basis of ligand binding by a c-di-GMP riboswitch. Nat. Struct. Mol. Biol. *16*, 1218-1223.

[0367] Sudarsan, N., Lee, E.R., Weinberg, Z., Moy, R.H., Kim, J.N., Link, K.H., and Breaker, R.R. (2008). Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science *321*, 411-413.

[0368] Sun, L., Wu, J., Du, F., Chen, X., and Chen, Z.J. (2013). Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. Science *339*, 786-791.

[0369] Sun, W., Li, Y., Chen, L., Chen, H., You, F., Zhou, X., Zhou, Y., Zhai, Z., Chen, D., and Jiang, Z. (2009). ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. Proc. Natl. Acad. Sci. USA *106*, 8653-8658.

[0370] 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.

[0371] 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.

[0372] Wolkowicz, U.M., and Cook, A.G. (2012). NF45 dimerizes with NF90, Zfr and SPNR via a conserved domain that has a nucleotidyltransferase fold. Nucleic Acids Res. *40*, 9356-9368.

[0373] 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 (New York, N.Y 328, 1703–1705.

[0374] Wu, J., Sun, L., Chen, X., Du, F., Shi, H., Chen, C., and Chen, Z.J. (2013). Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. Science *339*, 826-830.

[0375] Yang, P., An, H., Liu, X., Wen, M., Zheng, Y., Rui, Y., and Cao, X. (2010). The cytosolic nucleic acid sensor LRRFIP1 mediates the production of type I interferon via a beta-catenin-dependent pathway. Nat. Immunol. *11*, 487-494.

[0376] 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.

[0377] 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.

[0378] Zhong, B., Yang, Y., Li, S., Wang, Y.Y., Li, Y., Diao, F., Lei, C., He, X., Zhang, L., Tien, P., *et al.* (2008). The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. Immunity *29*, 538-550.

Equivalents and Scope

[0379] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments in accordance with the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the appended claims.

[0380] In the claims, articles such as "a," "an," and "the" may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from

the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process.

[0381] It is also noted that the term "comprising" is intended to be open and permits but does not require the inclusion of additional elements or steps. When the term "comprising" is used herein, the term "consisting of" is thus also encompassed and disclosed.

[0382] Where ranges are given, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different embodiments of the invention, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

[0383] In addition, it is to be understood that any particular embodiment of the present invention that falls within the prior art may be explicitly excluded from any one or more of the claims. Since such embodiments are deemed to be known to one of ordinary skill in the art, they may be excluded even if the exclusion is not set forth explicitly herein. Any particular embodiment of the compositions of the invention (*e.g.*, any nucleic acid or protein encoded thereby; any method of production; any method of use; *etc.*) can be excluded from any one or more claims, for any reason, whether or not related to the existence of prior art.

[0384] All cited sources, for example, references, publications, databases, database entries, and art cited herein, are incorporated into this application by reference, even if not expressly stated in the citation. In case of conflicting statements of a cited source and the instant application, the statement in the instant application shall control.

[0385] Section and table headings are not intended to be limiting.

CLAIMS

We claim:

1. A modulator of cGAS having a structure comprising the following features:

A ----- L ----- B

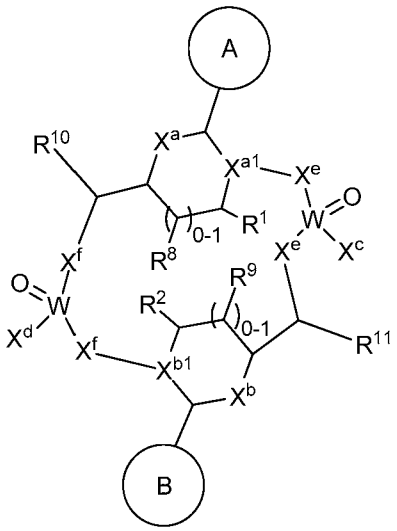
wherein:

A is or comprises a moiety that fits in the cGAS adenosine binding site;

B is or comprises a moiety that fits in the cGAS guanosine binding site; and

L is a linker moiety linking A and B in a manner which allows A and B to adopt appropriate interactions to bind cGAS.

2. A compound of Formula I:

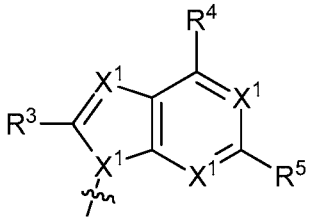
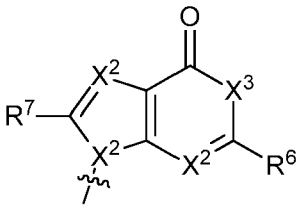
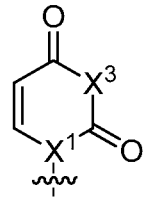
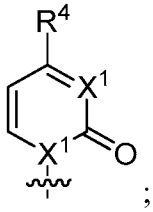


I

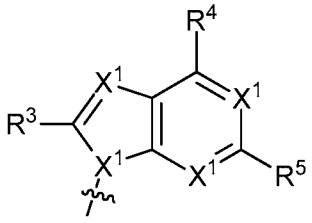
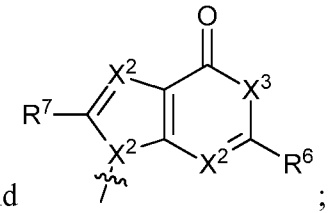
or a pharmaceutically acceptable salt thereof,

wherein:

Ring A is selected from the group consisting of:


,

,

, and

;

Ring B is selected from the group consisting of:


and

;

each X^1 and X^2 is independently –CR- or –N-;

X^3 is $–C(R)_2$-, -O-, or –NR-;

X^a and X^b are independently $–C(R)_2$-, -C(R)=C(R)-, -O-, -S-, -S(O)-, $-S(O)_2$-, or -N(R)-;

X^{a1} and X^{b1} are independently –C(R)- or –N-;

X^c and X^d, when present, are independently optionally substituted oxygen, optionally substituted sulfur, a substituted nitrogen atom, BH_3, or optionally substituted C_{1-12} aliphatic;

each X^e and X^f is independently –O-, -S-, or -N(R)-;

each W is independently P or S;

each R^1 and R^2 is independently selected from the group consisting of hydrogen, halogen, $–NO_2$, –CN, $–OR^a$, –SR, $–N(R)_2$, -C(O)R, $-CO_2R$, -C(O)C(O)R, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, -OC(O)R, -N(R)C(O)R, $–N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, and optionally substituted C_{1-12} aliphatic or C_{1-4} alkoxy-C_{1-4} alkyl;

each R^3, R^4, R^5, R^6, R^7, R^{10}, and R^{11} is independently selected from the group consisting of hydrogen, halogen, $–NO_2$, –CN, –OR, –SR, $–N(R)_2$, -C(O)R, $-CO_2R$, -C(O)C(O)R, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, -OC(O)R, -N(R)C(O)R, $–N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, or an optionally substituted group selected from C_{1-12} aliphatic, phenyl, a 3-7 membered saturated or partially unsaturated monocyclic carbocyclic ring, a 7-10 membered saturated or partially unsaturated bicyclic carbocyclic ring, a 3-7 membered saturated or partially unsaturated heterocyclic ring having 1-2

heteroatoms independently selected from nitrogen, oxygen, or sulfur, a 7-10 membered saturated or partially unsaturated bicyclic heterocyclic ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur, and a 5-6 membered heteroaryl ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur;

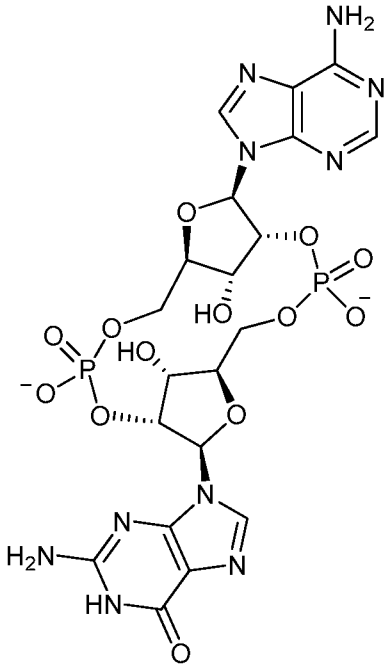
each R^8 and R^9, when present, is independently selected from the group consisting of hydrogen, halogen, $-NO_2$, $-CN$, $-OR^a$, $-SR$, $-N(R)_2$, $-C(O)R$, $-CO_2R$, $-C(O)C(O)R$, $-C(O)CH_2C(O)R$, $-S(O)R$, $-S(O)_2R$, $-C(O)N(R)_2$, $-SO_2N(R)_2$, $-OC(O)R$, $-N(R)C(O)R$, $-N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, $-C=NOR$, $-N(R)C(O)N(R)_2$, $-N(R)SO_2N(R)_2$, $-N(R)SO_2R$, $-OC(O)N(R)_2$, and an optionally substituted C_{1-12} aliphatic;

each R is independently selected from the group consisting of hydrogen or an optionally substituted group selected from C_{1-6} aliphatic, phenyl, a 3-7 membered saturated or partially unsaturated carbocyclic ring, a 3-7 membered saturated or partially unsaturated monocyclic heterocyclic ring having 1-2 heteroatoms independently selected from nitrogen, oxygen, or sulfur, and a 5-6 membered heteroaryl ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur; or:

two R groups on the same nitrogen are taken together with their intervening atoms to form an optionally substituted 3-7 membered saturated, partially unsaturated, or heteroaryl ring having 1-4 heteroatoms independently selected from nitrogen, oxygen, or sulfur; and

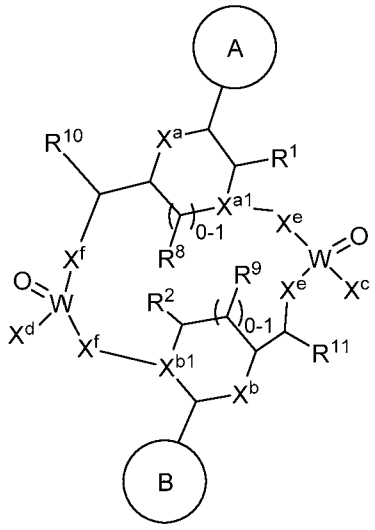
R^a is an oxygen protecting group or R;

with the proviso that the compound is other than:



or a pharmaceutically acceptable salt thereof.

3. A compound of Formula II:

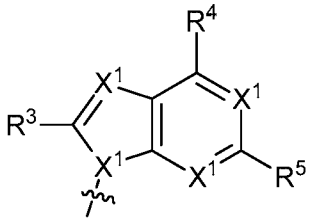
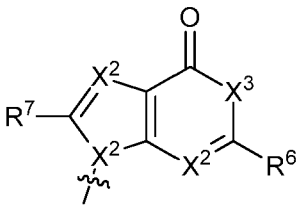
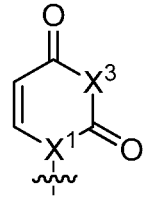
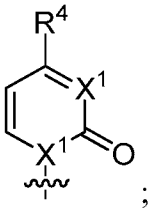


II

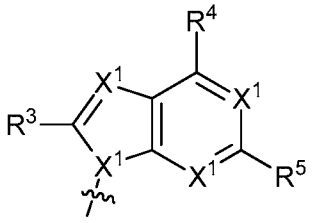
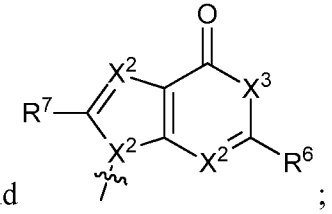
or a pharmaceutically acceptable salt thereof,

wherein:

Ring A is selected from the group consisting of:


,

,

, and

;

Ring B is selected from the group consisting of:


and

;

each X^1 and X^2 is independently –CR- or –N-;

X^3 is $–C(R)_2$-, -O-, or –NR-;

X^a and X^b are independently $–C(R)_2$-, -C(R)=C(R)-, -O-, -S-, -S(O)-, $-S(O)_2$-, or -N(R)-;

X^{a1} and X^{b1} are independently –C(R)- or –N-;

X^c and X^d, when present, are independently optionally substituted oxygen, optionally substituted sulfur, a substituted nitrogen atom, BH_3, or optionally substituted C_{1-12} aliphatic;

each X^e and X^f is independently –O-, -S-, or -N(R)-;

each W is independently P or S;

each R^1 and R^2 is independently selected from the group consisting of hydrogen, halogen, $–NO_2$, –CN, $–OR^a$, –SR, $–N(R)_2$, -C(O)R, $-CO_2R$, -C(O)C(O)R, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, -OC(O)R, -N(R)C(O)R, $–N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, and optionally substituted C_{1-12} aliphatic or C_{1-4} alkoxy-C_{1-4} alkyl;

each R^3, R^4, R^5, R^6, R^7, R^{10}, and R^{11} is independently selected from the group consisting of hydrogen, halogen, $–NO_2$, –CN, –OR, –SR, $–N(R)_2$, -C(O)R, $-CO_2R$, -C(O)C(O)R, $–C(O)CH_2C(O)R$, –S(O)R, $–S(O)_2R$, $–C(O)N(R)_2$, $-SO_2N(R)_2$, -OC(O)R, -N(R)C(O)R, $–N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, –C=NOR, $-N(R)C(O)N(R)_2$, $–N(R)SO_2N(R)_2$, $–N(R)SO_2R$, $-OC(O)N(R)_2$, or an optionally substituted group selected from C_{1-12} aliphatic, phenyl, a 3-7 membered saturated or partially unsaturated monocyclic

carbocyclic ring, a 7-10 membered saturated or partially unsaturated bicyclic carbocyclic ring, a 3-7 membered saturated or partially unsaturated heterocyclic ring having 1-2 heteroatoms independently selected from nitrogen, oxygen, or sulfur, a 7-10 membered saturated or partially unsaturated bicyclic heterocyclic ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur, and a 5-6 membered heteroaryl ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur;

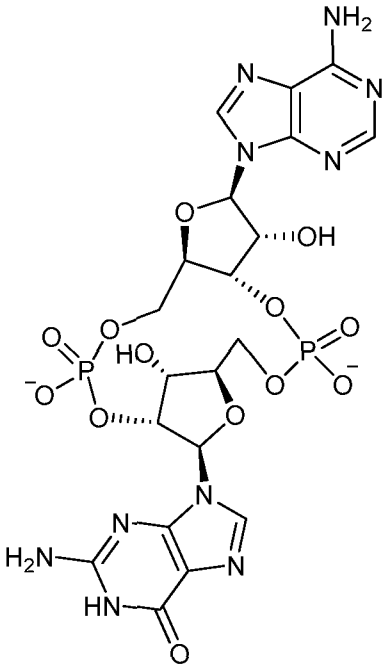
each R^8 and R^9, when present, is independently selected from the group consisting of hydrogen, halogen, $-NO_2$, $-CN$, $-OR^a$, $-SR$, $-N(R)_2$, $-C(O)R$, $-CO_2R$, $-C(O)C(O)R$, $-C(O)CH_2C(O)R$, $-S(O)R$, $-S(O)_2R$, $-C(O)N(R)_2$, $-SO_2N(R)_2$, $-OC(O)R$, $-N(R)C(O)R$, $-N(R)N(R)_2$, $-N(R)C(=NR)N(R)_2$, $-C(=NR)N(R)_2$, $-C=NOR$, $-N(R)C(O)N(R)_2$, $-N(R)SO_2N(R)_2$, $-N(R)SO_2R$, $-OC(O)N(R)_2$, and optionally substituted C_{1-12} aliphatic;

each R is independently selected from the group consisting of hydrogen or an optionally substituted group selected from C_{1-6} aliphatic, phenyl, a 3-7 membered saturated or partially unsaturated carbocyclic ring, a 3-7 membered saturated or partially unsaturated monocyclic heterocyclic ring having 1-2 heteroatoms independently selected from nitrogen, oxygen, or sulfur, and a 5-6 membered heteroaryl ring having 1-3 heteroatoms independently selected from nitrogen, oxygen, or sulfur; or:

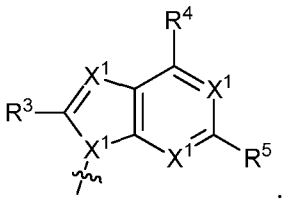
two R groups on the same nitrogen are taken together with their intervening atoms to form an optionally substituted 3-7 membered saturated, partially unsaturated, or heteroaryl ring having 1-4 heteroatoms independently selected from nitrogen, oxygen, or sulfur; and

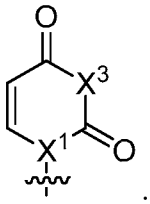
R^a is an oxygen protecting group or R;

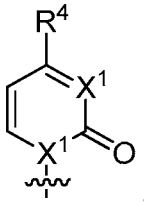
with the proviso that the compound is other than:

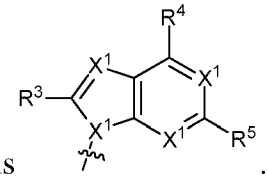


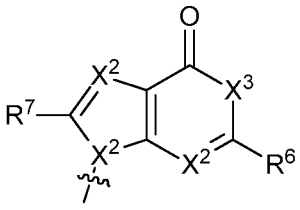
.

4. The compound of any one of the preceding claims, wherein Ring A is  .

5. The compound of any one of the preceding claims, wherein Ring A is  .

6. The compound of any one of the preceding claims, wherein Ring A is  .

7. The compound of any one of the preceding claims, wherein Ring B is .

8. The compound of any one of the preceding claims, wherein Ring B is 

9. The compound of any one of the preceding claims, wherein each X^1 is –N-.

10. The compound of any one of preceding claims, wherein each X^2 is –N-.

11. The compound of any one of the preceding claims, wherein X^3 is –NR-.

12. The compound of any one of the preceding claims, wherein X^a is –O-.

13. The compound of any one of the preceding claims, wherein X^b is –O-.

14. The compound of any one of the preceding claims, wherein X^{a1} and X^{b1} are –C(R)-.

15. The compound of any one of the preceding claims, wherein X^c and X^d are both oxygen.

16. The compound of any one of the preceding claims, wherein X^e and X^f are both oxygen.

17. The compound of any one of the preceding claims, wherein R^1 is selected from the group consisting of hydrogen, halogen, $–OR^a$, -SR, $-N(R)_2$, and optionally substituted C_{1-12} aliphatic or C_{1-4} alkoxy-C_{1-4} alkyl.

18. The compound of claim 17, wherein R^1 is –OH.

19. The compound of any one of the preceding claims, wherein R^2 is selected from the group consisting of hydrogen, halogen, $-OR^a$, -SR, $-N(R)_2$, and optionally substituted C_{1-12} aliphatic or C_{1-4} alkoxy-C_{1-4} alkyl.

20. The compound of claim 19, wherein R^2 is –OH.

21. The compound of claim 17 or 19, wherein R^a is R.

22. The compound of any one of the preceding claims, wherein each R^3, R^5, and R^7 is independently selected from the group consisting of hydrogen, halogen, and optionally substituted C_{1-6} aliphatic.

23. The compound of claim 22, where each of R^3, R^5, and R^7 is hydrogen.

24. The compound of any one the preceding claims, wherein each of R^4 and R^6 is independently selected from the group consisting of hydrogen, halogen, and $-N(R)_2$.

25. The compound of claim 24, wherein R^4 and R^6 are each $–NH_2$.

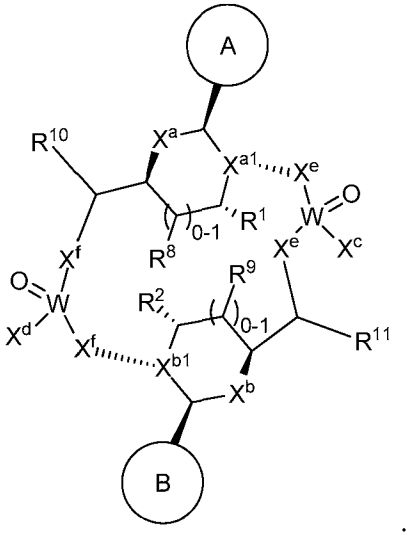
26. The compound of any one of the preceding claims, wherein R^8 and R^9 are hydrogen.

27. The compound of any one of the preceding claims, wherein R^{10} and R^{11} are hydrogen.

28. The compound of any one of the preceding claims, wherein W is P.

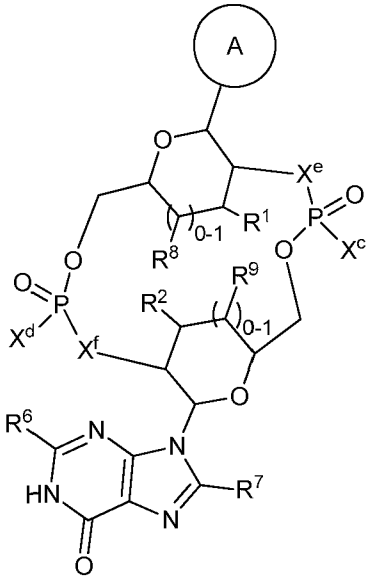
29. The compound of any one of the preceding claims, wherein R is hydrogen.

30. The compound of any one of claims 1, 2, or 4-29, wherein the compound is of formula I-a:

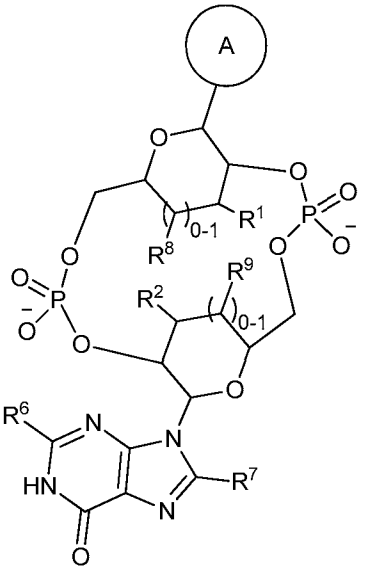


I-a

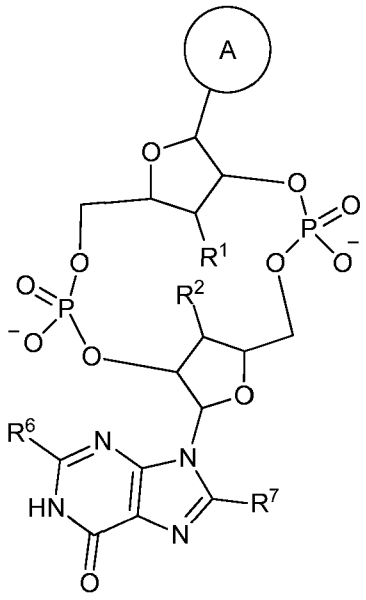
31. The compound of any one of claims 1, 2, or 4-29, wherein the compound is of formula III, IV, V, VI, VII, VIII, or IX:



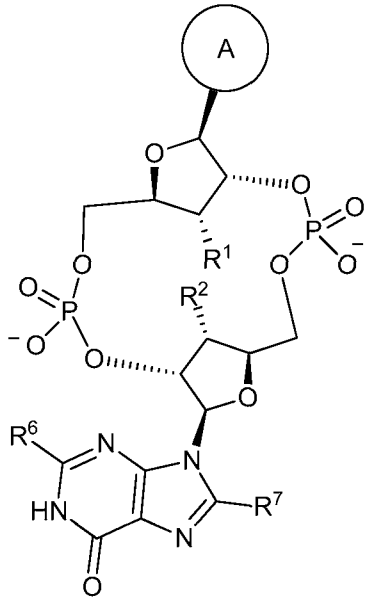
III



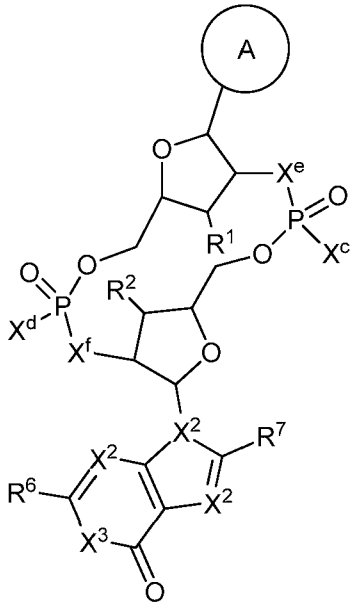
IV



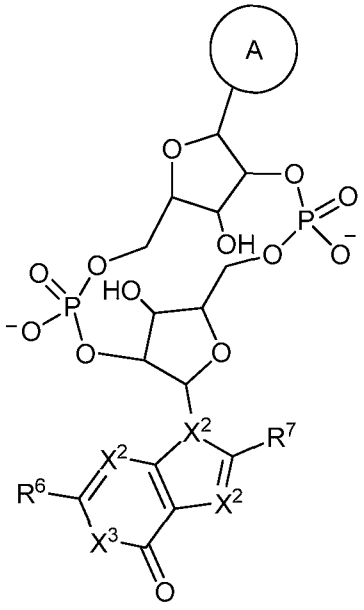
V



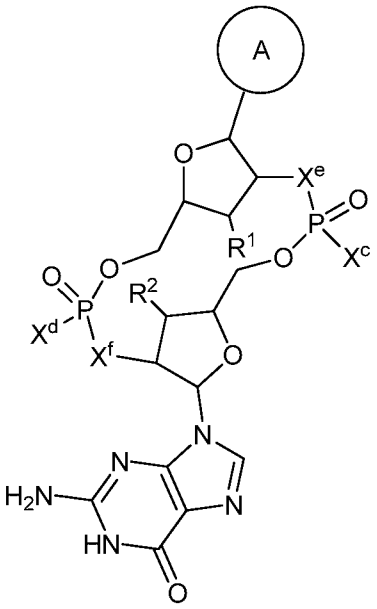
VI



VII



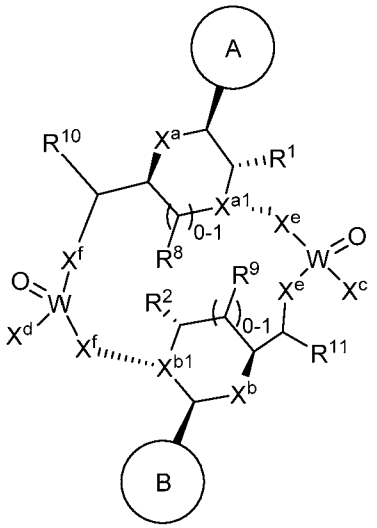
VIII



IX

or a pharmaceutically acceptable salt thereof.

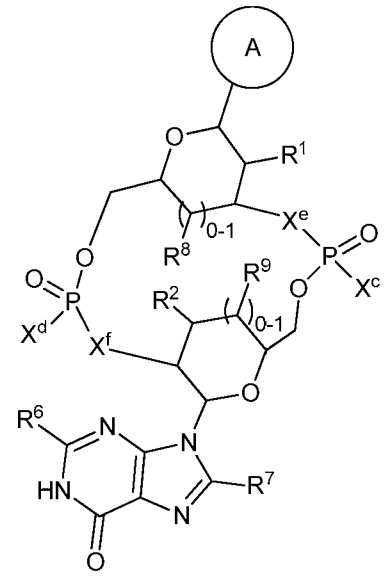
32. The compound of any one of claims 1 or 3-29, wherein the compound is of formula II-a:



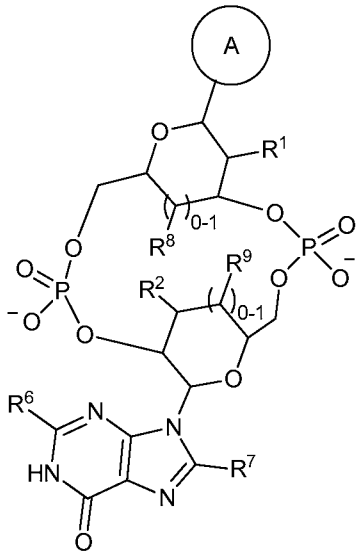
II-a

or a pharmaceutically acceptable salt thereof.

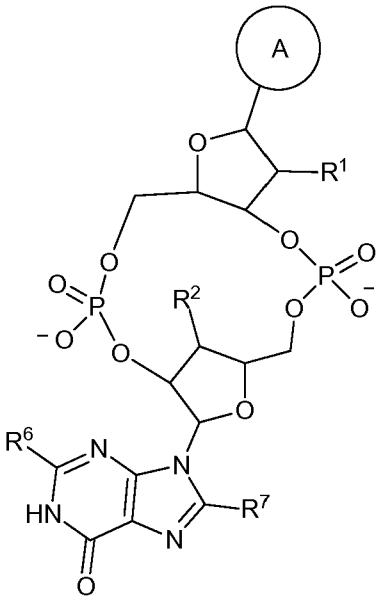
33. The compound of any one of claims 1 or 3-29, wherein the compound is of formula X, XI, XII, XIII, XIV, XV, or XVI:



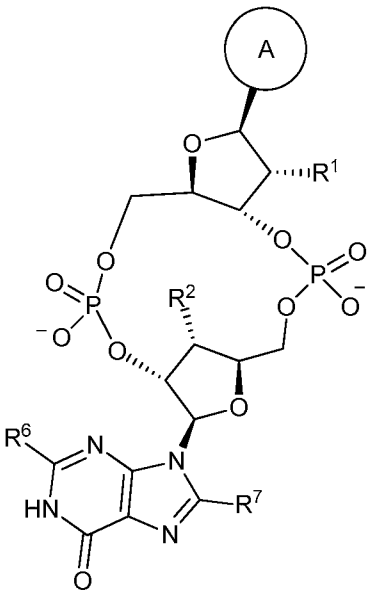
X



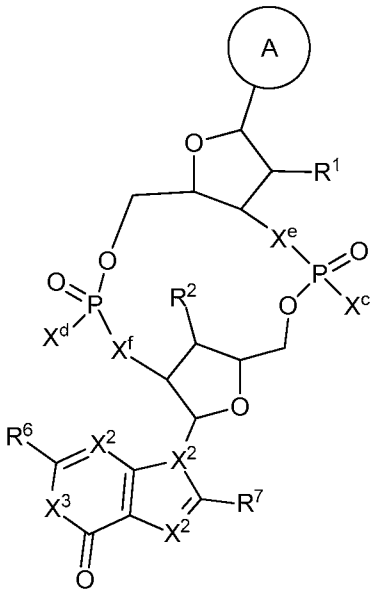
XI



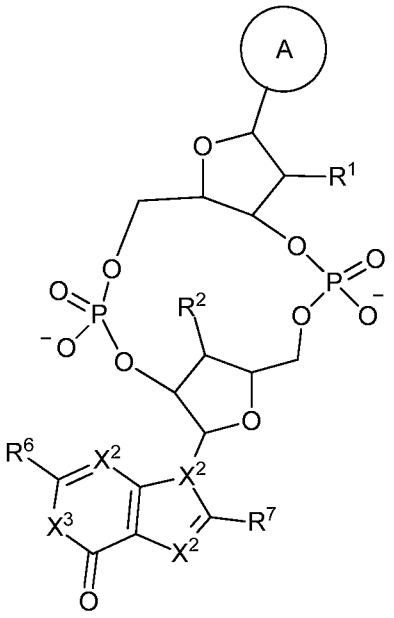
XII



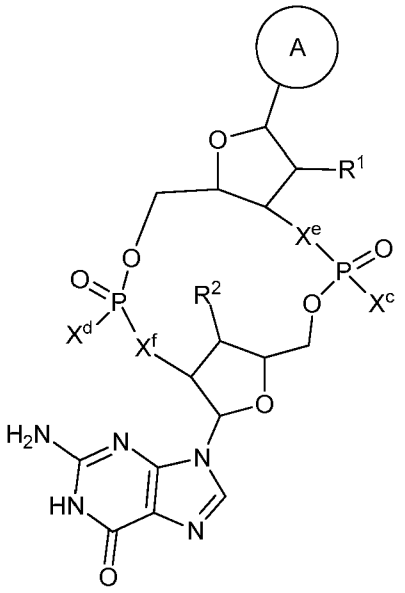
XIII



XIV



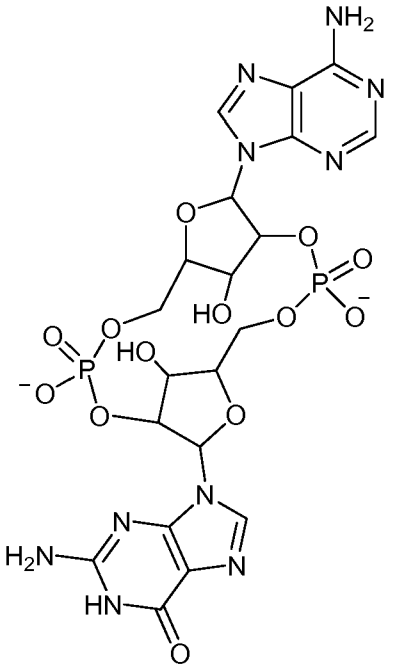
XV



XVI

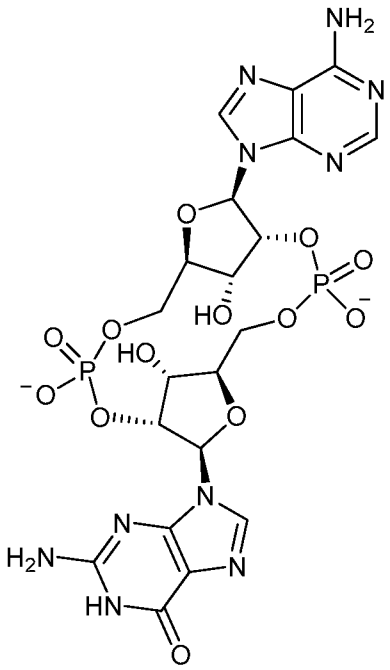
or a pharmaceutically acceptable salt thereof.

34. A compound:



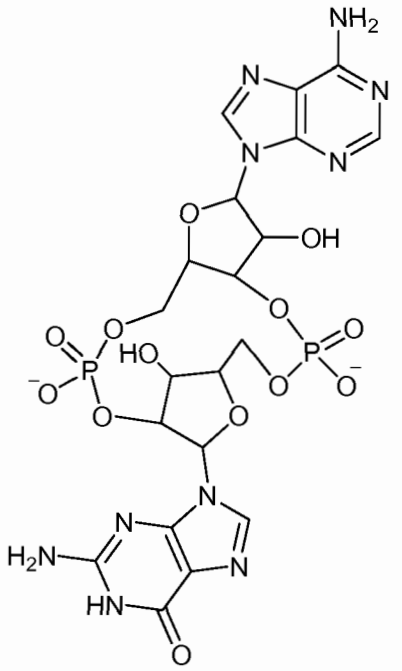
or a pharmaceutically acceptable salt thereof.

35. The compound of claim 34, wherein the compound is:



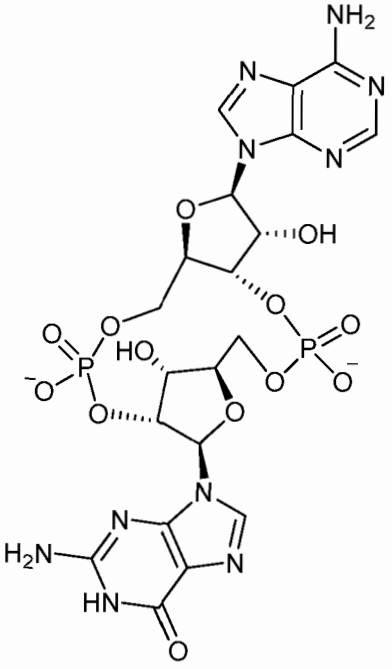
or a pharmaceutically acceptable salt thereof.

36. A compound:



or a pharmaceutically acceptable salt thereof.

37. The compound of claim 36, wherein the compound is:



or a pharmaceutically acceptable salt thereof.

38. The compound of any one of claims 34-37, in isolated form.

39. The compound of claim 38, in purified form.

40. A pharmaceutical composition comprising a compound of any one of claims 1-39 and a pharmaceutically acceptable carrier.

41. A method for treatment or prevention of an immune disease, disorder, or condition comprising administering to a subject in need thereof a compound of any one of claims 1-39 or the pharmaceutical composition of claim 40.

42. The method of claim 41, wherein the immune disease, disorder, or condition is an autoimmune disease, disorder, or condition.

43. The method of claim 41 of 42, wherein the immune disease, disorder, or condition is selected from the group consisting of systemic lupus erythematosus, Sjögren's syndrome, and Aicardi-Goutières syndrome.

44. The method of claim 41, wherein the immune disease, disorder, or condition is characterized by inflammation.

45. The method of claim 41, wherein the immune disease, disorder, or condition is caused by, sustained by, or related to cGAS activation.

46. The method of claim 41, wherein the immune disease, disorder, or condition is caused by, sustained by, or related to STING activation.

47. The method of claim 41, wherein the administration of the compound results in a decrease of cGAS activity.

48. The method of claim 41, wherein the administration of the compound results in a decrease of STING activity.

49. The method of any one of claims 41-48, wherein the compound is prepared by chemical synthesis.

50. A method of inhibiting cGAS comprising contacting cGAS with a compound of any one of claims 1-39.

51. A method of inhibiting STING comprising contacting STING with a compound of any one of claims 1-39.

52. A method of designing or characterizing a cGAS modulator comprising the steps of:
a) providing an image of a cGAS crystal that includes at least one potential interaction site;
b) docking in the image at least one moiety that is a potential cGAS modulator structural element; and
c) assessing one or more features of a potential moiety-interaction site interaction.

53. The method of claim 52, wherein the at least one potential interaction site includes a site selected from the group consisting of Ser199, Ser420, Lys402, Glu211, Asp213, Asp307, Tyr421, Arg364, and combinations thereof.

54. The method of claim 52 or 53, wherein the at least one potential interaction site includes a site selected from the group consisting of Tyr421, Thr197, Ser366, Ser368, Arg364, and combinations thereof.

55. The method of claim 52, 53, or 54, wherein the at least one potential interaction site includes a site selected from the group consisting of Tyr421, Asp213, Asp307, Arg364, and combinations thereof.

56. The method of claim 52, wherein the at least one potential interaction site includes Arg161.

57. The method of any one of claims 52-55, wherein the modulator is a compound of any one of claims 1-39.

58. The method of any one of claims 52-57, wherein the one or more features include at least one feature selected from the group consisting of: spatial separation between the moiety and the potential interaction site; energy of the potential moiety-interaction site interaction, and/or combinations thereof.

59. The method of any one of claims 52-57, further comprising a step of providing an image of a potential cGAS modulator comprising the moiety docked with the image of the cGAS crystal.

60. The method of claim 59, further comprising the step of comparing the image with that of a cGAS crystal including a bound known modulator, substrate, or product.

61. A system comprising a computer or computer readable medium in which a cGAS crystal structure, or coordinates thereof, is embedded and/or displayed.

62. A method of designing and/or characterizing an cGAS modulator, which method comprises steps of:
(i) using the system of claim 51 to assess one or more structural features of the cGAS modulator; and
(ii) performing one or more *in vitro, in vivo* or cell-based assays to characterize the cGAS modulator.

63. The method of any one of claims 52-60 or 62, further comprising the step of performing a competition experiment between the cGAS modulator and a known cGAS modulator, substrate, or product.

64. The method of claim 63, wherein the method further comprises the step of defining the three-dimensional shape of the inhibitor.

65. An inhibitor of cGAS characterized in that it binds in the binding pocket having a three-dimensional structure characterized by the structure coordinates of any one of Tables 1, 2, 3, 4, 5, 6, or 7.

66. A designed cGAS inhibitor, comprising the crystallography coordinates of any one of Tables 1, 2, 3, 4, 5, 6, or 7, wherein the crystallography coordinates are within about a root mean square deviation of not more than about 1.5 Å from the backbone atoms of the amino acids according to Tables 1, 2, 3, 4, 5, 6, or 7.

67. A computer system containing a set of information to perform a design or characterization of an cGAS inhibitor having a user interface comprising a display unit, the set of information comprising:
(i) logic for inputting an information regarding a binding of a cGAS protein to a moiety known to bind cGAS protein;
(ii) logic for design a candidate cGAS inhibitor based on the binding of the cGAS protein to the moiety known to bind cGAS protein;
(iii) logic for determining an information regarding a binding of the cGAS protein to the candidate cGAS inhibitor; and
(iv) logic for making a conclusion regarding a cGAS inhibitory properties of the candidate cGAS inhibitor based on the determination of step (iii).

68. A computer-readable storage medium containing a set of information for a general purpose computer having a user interface comprising, a display unit, the set of information comprising:
(i) logic for inputting an information regarding a binding of a cGAS protein to a chemical known to binding cGAS protein;
(ii) logic for design a candidate cGAS inhibitor based on the binding of the cGAS protein to the chemical known to bind cGAS protein;

(iii) logic for determining an information regarding a binding of the cGAS protein to the candidate cGAS inhibitor; and
(iv) logic for making a conclusion regarding a cGAS inhibitory properties of the candidate cGAS inhibitor based on the determination step of step (iii).

69. An electronic signal or carrier wave that is propagated over the internet between computers comprising a set of information for a general purpose computer having a user interface comprising a display unit, the set of information comprising a computer-readable storage medium containing a set of information for a general purpose computer having a user interface comprising a display unit, the set of information comprising:
(i) logic for inputting an information regarding a binding of a cGAS protein to a chemical known to bind cGAS protein;
(ii) logic for designing a candidate cGAS inhibitor based on the binding of the cGAS protein to the chemical known to bind cGAS protein;
(iii) logic for determining an information regarding a binding of the cGAS protein to the candidate cGAS inhibitor; and
(iv) logic for making a conclusion regarding a cGAS inhibitory properties of the candidate cGAS inhibitor based on the determination of step (iii).

70. A method of modulating activity of an cGAS polypeptide, the method comprising contacting the cGAS polypeptide with a cGAS modulator of claim 52 which modulating agent is not a known modulator, substrate, or product of cGAS.

71. The method of claim 70, wherein the modulating agent is a compound of any one of claims 1-39.

72. A crystalline or crystallizable composition comprising or consisting of an cGAS polypeptide.

73. A method of making the composition of claim 72.

74. A method of using the composition of claim 72.

75. A method of reducing the level of interferon in a subject comprising contacting said subject with an antagonist of cGAS.

76. The method of claim 75, wherein the antagonist is an analog, mimic or variant of an isomer of the cGAMP family of molecules.

77. A method of altering cytokine profile in an individual, comprising contacting the individual with a cGAS antagonist which is an analog, mimic or variant of an isomer of the cGAMP family of molecules.

78. A kit comprising one or more cGAMP compounds and optionally one or more cGAS molecules.