Solution structures of a G-quadruplex bound to linear- and cyclic-dinucleotides

Winnerdy, Fernaldo Richtia; Das, Poulomi; Heddi, Brahim; Phan, Anh Tuấn

2019

Winnerdy, F. R., Das, P., Heddi, B., & Phan, A. T. (2019). Solution structures of a G-quadruplex bound to linear- and cyclic-dinucleotides. *Journal of the American Chemical Society*, 141(45), 18038-18047. doi:10.1021/jacs.9b05642

<https://hdl.handle.net/10356/140171>

<https://doi.org/10.1021/jacs.9b05642>

This document is the Accepted Manuscript version of a Published Work that appeared in final form in *Journal of the American Chemical Society*, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see <https://doi.org/10.1021/jacs.9b05642>

Downloaded on 26 Apr 2023 16:49:45 SGT

Solution structures of a G-quadruplex bound to a linear- and cyclic-dinucleotide (cGAMP)

Fernaldo Richtia Winnerdy¹, Poulomi Das,¹ Brahim Heddi^{1,2} and Anh Tuấn Phan^{1,*}

¹ School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore

² Laboratoire de Biologie et de Pharmacologie Appliquée, CNRS UMR 8113, Ecole Normale Supérieure Paris-Saclay, Cachan 94235, France

KEYWORDS: *dinucleotide, cyclic-dinucleotide, G-quadruplex, ligand, structure, cGAMP, NMR*

ABSTRACT: Cyclic dinucleotides have emerged as important secondary messengers and cell signalling molecules that regulates several cell responses. A guanine-deficit G-quadruplex structure formation by a sequence containing $(4n-1)$ guanines— n denoting the number of G-tetrad layer—was previously reported. Here, a $(4n-1)$ G-quadruplex structure is shown to be capable of binding guanine-containing dinucleotides in micromolar affinity. The guanine base of the dinucleotides interacts with a vacant G-triad, forming four additional Hoogsteen hydrogen bonds to complete a G-tetrad. Solution structures of two complexes, both comprised of a $(4n-1)$ G-quadruplex structure—one bound to a linear dinucleotide and the other to a cyclic dinucleotide (cGAMP)—are solved using NMR spectroscopy. The latter suggests sufficiently strong interaction between the guanine base of the dinucleotide and the vacant G-triad, which acts as an anchor point of binding. The detailed binding interfaces from the two solution structures provide useful information for specific ligand and/or drug design. The results also infer that other guanine-containing metabolites of a similar size have the capability of binding G-quadruplexes, potentially affecting the expression of the metabolites and functionality of the bound G-quadruplexes.

INTRODUCTION

It is of great interest to understand ligand binding towards different nucleic acid secondary structures¹. Specifically, the interaction of ligands towards a distinct four-stranded DNA or RNA structure termed G-quadruplex (G4) has been continuously studied.²⁻⁶ G4 is a secondary structure of DNA or RNA which consists of stacking G-tetrads and stabilizing cations, where each G-tetrad comprises of four Hoogsteen-associated guanine bases forming a planar structure⁷⁻¹². G4-forming motifs¹³⁻¹⁴ have been extensively studied and shown to be abundant in regions with key genomic functions such as telomeric, 5'-UTR, gene promoter, and replication initiation regions.¹⁵⁻¹⁶ The G4 structure has been established to have physiological roles in telomerase dysfunction¹⁷, gene expression¹⁸, DNA damage¹⁹ and genomic instability²⁰⁻²¹ among others¹⁵⁻¹⁶. A vast range of G4 structure variations was established, such as formations of bulges²²⁻²⁴, double-helix loops²⁵⁻²⁷, stacking triads²⁸⁻³⁰, non-canonical tetrad planes³¹⁻³² and left-handed structures³³⁻³⁴. Thus far, the primary function of most studied G4 ligands was to stabilize the quadruplex structure upon binding²⁻⁶. However, there is a lack of research on G4 ligands that promote or even required in the formation of the G4 structure itself. Our previous study showed the formation of a $(4n-1)$ G4 structure³⁵⁻³⁶— n denotes the number of G-tetrad layers; $4n-1$ denotes the number of guanines—which contains a perfect vacant site for a guanine base filling (Figure 1C).

Guanine derivatives such as guanosine or guanine triphosphate (GTP) were shown to bind DNA sequences with $(4n-1, n = 3)$ guanine residues, supplementing it towards a complete G-quadruplex structure with three G-tetrads³⁶⁻³⁸. This type of interaction is highly probable *in vivo* due to the abundance of guanine-derivatives³⁹ as well as frequent occurrence of guanine-deficit G4 motifs in both eukaryote and prokaryote cells.

One of the most prominent guanine-containing metabolites in cells are cyclic-dinucleotides⁴⁰. Cyclic-dinucleotides are 2-nt ribonucleic acids with an additional bond linkage that connect the phosphodiester backbone to form a cyclic loop. A few types of cyclic-dinucleotides have been extensively studied⁴¹. For instance, cyclic di-GMP was found to be a secondary messenger in bacteria⁴², where it serves as a gene expression regulator⁴³. Furthermore, together with extracellular quorum-sensing, the intracellular cyclic-dinucleotide signalling administers bacterial complex processes such as multicellularity, biofilm formation, and virulence⁴⁴. In human cells, cyclic GMP-AMP (or cGAMP), particularly the one with 2'-5' phosphodiester linkage in the sugar of the guanine (known as 2'3'cGAMP), was established to be a required metabolite that activates endoplasmic-reticulum protein STING, which has a crucial role in DNA sensing and immunology⁴⁰. In addition, 3'3'cGAMP was found to play a role in bacterial chemotaxis and colonization as a second messenger.⁴⁵ Most regulatory roles require a precise concentration level of the metabolites, therefore any molecules

that can alter the concentration by sequestering the cyclic-dinucleotides could play important role for the overall cell response. In addition, the binding of cell metabolites to a (4*n*-1) G4 has been applied in the design of metabolite sensing^{36-37, 46}, and it could also potentially lead to riboswitch-like regulatory roles⁴⁷.

Herein, we report the solution structures of two complexes involving a (4*n*-1) G4-forming sequence named *delta1*, d(TTGGTGGGTGGGTGGGT). Its interactions with two selected ligands: (i) a linear dinucleotide d(AG) and (ii) a 3'3' cyclic-dinucleotide guanine-adenine-monophosphate (cGAMP), were investigated. The solution structures confirmed the expected binding of the guanine base of the ligands to the vacant site³⁵⁻³⁸, with both complexes incorporating a complete 4*n*, 12-guanine G4 cores. Additionally, adenine-thymine interaction was detected in the *delta1*-d(AG) complex, strengthening the binding. The binding affinities of multiple dinucleotides were determined, providing additional information on the possibility of *in vivo* interaction. Detailed molecular interaction between the dinucleotides and the (4*n*-1) G4 *delta1* provides insights towards improving G4 ligand sensitivity and specificity, and thus is beneficial for drug design.

MATERIALS AND METHODS

Bioinformatics Search. The basic algorithm for the search is exactly one G₂ tract and three G₃ tracts, excluding the sequences which potentially form a complete G4 via a guanine base in the loop or a single-base bulge adjacent to the G₂ tract.²² The algorithm (see Supporting Information) was matched against the *hg38* human reference genome.

Sample Preparation. Unlabelled oligonucleotides (*delta1*), as well as linear dinucleotide d(AG) were chemically synthesized using an ABI 394 DNA/RNA synthesizer. Necessary reagents were purchased from Glen Research. Samples were purified with a Poly-Pak II cartridge (Glen Research) and dialyzed successively against water, 25 mM KCl solution and water. The samples typically contained 10 mM KCl and 10 mM potassium phosphate buffer (KPi) (pH 7.0), supplemented with 1 mM EDTA and 0.1 mM NaN₃. The oligonucleotides were quenched (by heating to 95°C and fast cooling to room temperature) before any spectroscopic measurements to trap the kinetically favourable structures. 3'3' Cyclic-GAMP, 2'3' cyclic-GAMP and cyclic-diGMP were purchased from InvivoGen and a stock solution was prepared (1.4 mM in Limulus ameobocyte lysate (LAL) reagent water).

Ultraviolet (UV) Absorption Spectroscopy. UV spectra were recorded on a JASCO V-650 spectrophotometer using a 1-cm path-length quartz cuvette. The DNA concentration was ~10 μM. UV absorbance at 295 nm was recorded as a function of temperature (from 95°C to 20°C and back to 95°C) and used to define the thermal stability of *delta1*. The heating and cooling rates were set to 0.5°C per minute. Average between cooling and heating transition points was taken to be the melting temperature of the sequence.

Circular Dichroism (CD) Spectroscopy. CD spectra were recorded at 20°C on a JASCO-815 spectropolarimeter

using a 1-cm path-length quartz cuvette. The DNA concentration was ~10 μM. For each experiment, an average of three scans was recorded between 220 and 320 nm. Baseline correction was applied for all spectra and the data were zero-corrected at 320 nm.

Nuclear Magnetic Resonance (NMR) Spectroscopy.

Experiments were performed on 600 MHz Bruker AVANCE II spectrometer between 10 and 25°C. DNA concentrations were between 0.1-1.5 mM. 2D NMR spectra were obtained using [DNA]:[ligand] concentration ratio of 1:1.2 and 1:1.5 for d(AG) and 3'3'cGAMP, respectively. Various salt concentrations (from 10 to 100 mM K⁺) were tested and similar NMR spectra were observed, indicating the formation of the same G-quadruplex structure. All spectra were processed and analyzed using TopSpin 3.5 and Sparky 3.1⁴⁸ softwares.

Structure Calculation.

NOE distance restraints. Inter-proton distances for *delta1*-d(AG) and *delta1*-cGAMP complexes were obtained from NOESY experiments performed in H₂O and D₂O at various mixing times (100, 250 and 300 ms). For non-exchangeable protons, the peaks were classified as strong, medium, medium-weak and weak corresponding to the distance restraints of (2.7 ± 0.8) Å, (3.8 ± 0.9) Å, (4.6 ± 1.2) Å and (5.5 ± 1.7) Å respectively. Distances from exchangeable protons were classified as strong, medium and weak corresponding to the distance restraints of (4.0 ± 1.2) Å, (4.8 ± 1.4) Å and (5.5 ± 1.7) Å respectively. Distances involving thymine methyl protons were altered to be directed towards the methyl carbon with a 0.5 Å looser restraints as compensation.

Dihedral restraints. Dihedral angle restraints were imposed to the dihedral angle formed by O4'-C1'-N9-C4 of guanine residues. *Anti*-guanine residues were restricted to an angle of (240 ± 70)° or (240 ± 40)° for the outer tetrad and inner tetrad guanines respectively. *Syn* guanine residues were restricted to an angle of (60 ± 70)° for the guanine residue from d(AG). Additional dihedral angle restraints were implemented on the 3'3'cGAMP molecule (complete list of the angles is shown in Table S1) based on several crystal structures (residue 4BW, PDB IDs: 5VDT, 4YAZ, 4YB1, 6AEL and 5CFM).⁴⁹⁻⁵² Dihedral angles with standard deviation values less than 25° were used as constraints in the structure calculation.

Hydrogen-bond restraints. Hoogsteen hydrogen bonds between guanines were restrained using H21-N7, N2-N7, H1-O6 and N1-O6 distances, which were set to (2.0 ± 0.2) Å, (2.9 ± 0.3) Å, (2.0 ± 0.2) Å and (2.9 ± 0.3) Å respectively. The hydrogen bond restraints from the Hoogsteen interaction between adenine of d(AG) and T17 were set to (2.0 ± 0.2) Å for N7-H3, (2.9 ± 0.3) Å for N7-N3, (2.0 ± 0.2) Å for H62-O4 and (2.9 ± 0.3) Å for N6-O4.

Planarity restraints. Planarity restraints were used for the G3-G6-G10-G14, G4-G7-G11-G15 and G*-G8-G12-G16 tetrads for both complexes, as well as for the A*-T17 base pair of the linear dinucleotide complex.

Generation of topology and parameter tables. The structure of 3'3'cGAMP molecule was extracted from a crystal

Figure 1. (A) 1D NMR spectra of *T4* (top) and *delta1* (bottom) showing 10 peaks in the imino proton region. The DNA sequences are given above the respective spectrum. Each peak is marked with a black dot. (B) Schematics of G-quadruplex folds of *T4* (top) and *delta1* (bottom). Black arrows indicate 5'-3' backbone progression and red lines indicate single thymine propeller loops. The thymine bulge is indicated in the *T4* schematic. (C) Illustration of vacant G-triad. Indicated in high transparency is possible filling guanine that would complete the G-tetrad, adding four more Hoogsteen hydrogen bonds. The black sphere in the middle indicates a monovalent cation.

structure (residue 4BW, PDB ID: 5CFM)⁵² and used as initial structure for the calculation. Hydrogen atoms were added into the crystal structure using the 'add hydrogen' function from YASARA software⁵³. The resulting PDB file was used as the input for the generation of topology, parameter files and partial charges⁵⁴ with ACPYPE program.⁵⁵

Distance geometry simulated annealing. Initial extended conformations of both complexes were generated using the XPLOR-NIH program⁵⁶ by supplying the available standard RNA topology and parameter tables as well as the generated 3'3'cGAMP topology and parameter table. Each system was then subjected to distance geometry simulated annealing by incorporating distance, dihedral, hydrogen bond and planarity restraints. 100 structures were generated and subjected to further refinement.

Distance-restrained molecular dynamics refinement. The 100 structures obtained from each simulated annealing step were refined with a distance-restrained molecular dynamics protocol incorporating all distance restraints. The system was heated from 300 to 1000 K in 14 ps and allowed to equilibrate for 6 ps, during which force constants for the distance restraints were kept at 2 kcal.mol⁻¹.Å⁻². The force constants for non-exchangeable proton and exchangeable proton restraints were then increased to 16 kcal.mol⁻¹.Å⁻² and 8 kcal.mol⁻¹.Å⁻² respectively in 20 ps before another equilibration at 1000 K for 50 ps. Next, the system was cooled down to 300 K in 42 ps, after which an equilibration was performed for 18 ps. Coordinates of the molecule were saved every 0.5 ps during the last 10.0 ps and averaged. The average structure obtained was then subjected to minimization until the gradient of energy was less than 0.1 kcal.mol⁻¹. Dihedral (50 kcal.mol⁻¹.rad⁻²) and planarity (1 kcal.mol⁻¹.Å⁻² for tetrads) restraints were maintained

throughout the course of refinement. Ten-lowest energy structures were selected.

Isothermal Titration Calorimetry (ITC). ITC experiments were performed using ITC-200. The initial concentration of oligonucleotide (*delta1*) was set to be 600 μM, unless specified otherwise. The ligands stock concentrations were ~2-4 mM, with the titration volume of 2 μL. The data were fitted using MicroCal-PEAQ-ITC analysis software, by specifying the concentrations of the oligonucleotides and the stock ligands. All other parameters, including the number of binding sites (which can be translated to folded fraction of the sample), the reaction enthalpy (ΔH) and the dissociation constant (K_D) were set as fitting parameters.

NMR Thermal Denaturation. A series of 1D ¹H NMR experiments were performed for both the free *delta1* sample and *delta1*-d(AG) sample in 5 mM KPi and 1 mM KCl buffer. In both cases, the concentration of *delta1* was 100 μM, with 300 μM d(AG) added to the latter sample. The experiments were performed in a 5°C interval starting from 25°C to 50°C (the upper limit for the cryoprobe set by the NMR facility). The quantitative determination of the melting temperature was obtained by integrating a representative peak in the imino proton region of each series of spectra. The peak area is considered to be proportional to the concentration of the folded species. The area under the peak was plotted against the temperature; the resulting plot was fitted using a function derived from Boltzmann distribution function (see Figure S14) to obtain the approximate melting temperature values.

Data Deposition. The coordinates of the ten lowest-energy structures of the *delta1*-d(AG) and *delta1*-cGAMP complexes have been deposited in the Protein Data Bank (PDB ID: 6K3X and 6K3Y)

Figure 2. Chemical structures of d(AG) and 3'3'cGAMP (top), as well as NMR titration experiments (bottom). Dinucleotides were added into 100 μ M *delta1* up to the 2-fold equivalent ratio. Both set of spectra showed 10 initial (free form) peaks and 12 final (bound form) peaks. Peaks of the free and bound forms are indicated by black dots and red asterisks respectively.

RESULTS

Abundance of (4n-1) G-quadruplex motifs with a vacant site. There has been an escalating number of putative G4-forming motifs in the last decade, ever since its first introduction of $\sim 376,000$ sequences.¹³⁻¹⁴ For instance, the G4-hunter scoring algorithm indicated an increase in potential G4-forming motifs by 2-10 folds,⁵⁷ which was also affirmed by the results of high-throughput sequencing experiments stating the existence of $\sim 716,000$ G4-forming motifs.⁵⁸ Here, the number of 3-layer and vacant-guanine (4n-1, n=3) G4 motifs was determined by bioinformatic analyses, where the sequence pattern of exactly one G₂ tract plus three G₃ tracts each separated by 1-7 nt loops was matched against the human genome (see Supporting Information). Sequences with a possibility of forming a complete G₃ tract via a guanine base in the loop or a single-base bulge adjacent to the G₂ tract were excluded from the result (see Supplementary Information, Figure S1). We found $\sim 230,000$ of such motifs, which is comparable to the number of (4n, n $\geq$ 3) G4 motifs of $\sim 376,000$ based on an analogous searching technique. The number of guanine-vacant G4 motifs could very well be expanded if more generous constraints were used. For instance, cases of more than three-layer G4 (n > 3) were not counted, as well as the situations with guanine base(s) in the connecting loops.

Formation of a (4n - 1) G-quadruplex structure, *delta1*. The 18-nt sequence *T95-2T*, d(TTGGGTGGGTGGGTGGGT) is frequently used as a reference G4-forming sequence⁵⁹. It forms a monomeric propeller-type parallel G4 in K⁺ solution; NMR imino proton spectrum of *T95-2T* displays 12 well-resolved peaks⁵⁹ (Figure

S2). A G-to-T substitution at position 4 of *T95-2T* results in a sequence named *T4* that forms a so called (4n-1) G-quadruplex containing a vacant site in the G-tetrad core (Figure 1B, top); NMR imino proton spectrum of *T4* displays 10 well-resolved peaks (Figure 1A, top)³⁶. However, the *T4* sequence contains a destabilizing thymine bulge in the first G-tract²². Removal of this bulge results in the *delta1* sequence, d(TTGGGTGGGTGGGTGGGT). The NMR spectrum of *delta1*, displaying 10 well-resolved imino protons, is very similar to that of *T4*, suggesting the formation of the same (4n-1) G4 fold (Figure 1A, B bottom). This observation is consistent with the formation of two G-tetrads and a G-triad. The melting temperature of *delta1* in 10 mM KCl and 10 mM KPi buffer was measured to be 48.1°C using UV spectroscopy (data not shown), higher than that of *T4* (45.6°C in 35 mM KCl and 10 mM KPi buffer)³⁶. The CD spectrum showed a positive peak at 260 nm and a negative trough at 240 nm (Figure S3), consistent with the formation of a parallel G4 structure⁶⁰⁻⁶². The *delta1* sequence was used for all further measurements as a model system due to its simple and well-resolved NMR spectrum. However, it is expected that the results could be translated into other natural vacancy-bearing G4 sequences.³⁸

Binding of dinucleotides to the *delta1* G-quadruplex structure involves Hoogsteen hydrogen bonds. We observed the interactions between several guanine-containing dinucleotides (linear and cyclic) and *delta1* by NMR spectroscopy, for which addition of dinucleotides showed emergence of new imino proton peaks (Figure 2). Each tested dinucleotide: both linear (d(AG), d(GA), d(TG) and d(GT); Figure S4A) and cyclic (3'3'cGAMP, 2'3'cGAMP and cdiGMP; Figure S4B), indeed showed a binding. Here, we focused on one of each ligand type as a representative: a linear

dinucleotide d(AG) and a cyclic dinucleotide 3'3'cGAMP (Figure 2, top), mainly due to their well-resolved NMR spectra. Both the ligands were titrated separately into 100 μ M *delta1* up to 1:2 [DNA]:[ligand] concentration ratio (Figure 2, bottom). About 1.2 equivalence ratio of d(AG) was sufficient to completely bind *delta1* (data not shown), while $\sim$ 2 equivalence ratio is required for cGAMP to achieve the same result, indicating a lower affinity of the latter ligand to *delta1*. The imino proton spectra of the complexes showed 12 peaks in the 10 to 12 ppm range (marked by red asterisks in Figure 2, bottom). The peak shifts are expected in any type of binding due to alteration of the chemical environment. However, the appearances of two additional imino proton peaks indicate that the ligand interacts with the vacant G-triad via Hoogsteen hydrogen bonding, transforming the vacant G-triad into a standard G-tetrad with a total of 8 hydrogen bonds (Figure 1C). This interaction mode would involve both Hoogsteen and Watson-Crick faces of the guanine base, as previously proposed for the interaction between guanine mononucleotides (e.g. GTP) and G4-motifs.^{36, 38, 63}

NMR structure determination of *delta1*-dinucleotide complexes. Assignments of guanine imino (H1) and aromatic (H8) protons of the *delta1*-d(AG) complex were obtained by site-specific 2% 15 N enrichment method⁶⁴ and long-range through-bond coupling method⁶⁵ respectively (Figure S5A, B). H1-H8 connectivity patterns indicative of G-tetrad formations were observed in the NOESY spectrum recorded in H₂O (Figure 3A, top). It showed that the core structure consisted of three G-tetrads, G3•G6•G10•G14, G4•G7•G11•G15, and G*•G8•G12•G16 (G* denotes the guanine residue from the dinucleotide) as well as that all four G-tracts were oriented parallel to each other. H8_(n)-H1'_(n)-H8_(n+1) sequential walk was recognized in the NOESY spectrum recorded in D₂O (Figure 3A, middle), showing broken connections only on the thymine loops. Cross-peak of G16(H1')-T17(H8) was detected indicating stacking of the thymine terminal to the 3'-end G-tetrad. Medium intensities of H8-H1' intra-residue cross-peaks—characteristic of *anti*-guanine conformations—were observed in NOESY at 100-ms mixing time for all the *delta1* guanines, while a strong intensity of the corresponding cross-peak for the guanine (G*) in d(AG) was detected, indicating a *syn*-guanine conformation (Figure 3A, middle). The proposed schematic structure is shown in Figure 3A.

Similar experimental procedures were performed to elucidate the solution structure of the *delta1*-cGAMP complex. The assignments of the guanine imino and aromatic protons were obtained based on the 1D and NOESY NMR spectra comparisons with the previous complex. The 1D NMR imino proton spectrum, H1-H8 connectivity patterns and NOESY sequential walk are presented (Figure 3B). Similar G-quadruplex core was observed with G-tracts oriented parallel to each other. Missing cross-peak of G16(H1')-T17(H8) indicated non-stacking thymine terminal on the 3'-end. The proposed schematic structure of the *delta1*-cGAMP complex is shown in Figure 3B.

NMR solution structures of *delta1* complexes was solved based on NMR constraints (Table 1). The 10 lowest-energy

structures of the complexes are presented in Figure 4A, B and Figure 5A, B.

***Syn*-guanine and A•T base pair in the *delta1*-d(AG) complex.** The *delta1*-d(AG) complex involves the completion of the 3' vacant G-triad of *delta1* into a standard G-tetrad (Figure 1C). The completing guanine (G*) coming from the ligand d(AG) was ascertained to be in *syn*-conformation based on the strong intensity of NOE intra-residue cross-peak between its aromatic (H8) and sugar (H1') protons (Figure 3A). Other distance information was acquired from the NOESY spectra, such as proximity of G* aromatic proton (H8) with G4 sugar protons (H1'/H2'/H2''/H3') and proximity of G4 imino proton (H1) with G* sugar proton (H2'), further supporting the proposition of *syn*-conformation of G* (Figure S6). In addition, hydrogen bond between the A* residue from the dinucleotide and the T17 residue was inferred via the observation of T17 imino proton peak in 1D NMR spectrum (Figure S7A) and an NOE cross-peak between T17(H3) and A*(H8) protons (Figure S7B). This cross-peak established the interaction between A* and T17 to be Hoogsteen base-pairing as opposed to the Watson-Crick base-pairing (Figure S7C). The insertion of the dinucleotide d(AG) into the (4n-1) G4 structure *delta1* and the interaction between them were illustrated in detail (Figure 4C, D), accompanied by surface representation (Figure 4E, F).

Table 1. Statistics of the computed structures of *delta1*-d(AG) and *delta1*-cGAMP complexes

delta1-d(AG)			delta1-cGAMP	
A. NMR Restraints				
Distance restraints	D2O	H2O	D2O	H2O
Intra-residue	339	0	361	0
Inter-residue	91	57	78	39
Other restraints				
Hydrogen bond	52		48	
Dihedral angle	13		23	
Planarity	4		4	
B. Structure Statistics				
NOE violations				
Number (>0.2 Å)	0.700 ± 0.700		0.000 ± 0.000	
Deviations from the ideal geometry				
Bond lengths (Å)	0.003 ± 0.000		0.004 ± 0.000	
Bond angles (°)	0.671 ± 0.006		1.270 ± 0.020	
Impropers (°)	0.357 ± 0.002		0.382 ± 0.082	
Pairwise heavy atom RMSD value (Å)				
G-tetrad core	0.539 ± 0.120		0.499 ± 0.106	
All heavy atom	1.463 ± 0.548		2.400 ± 0.294

Figure 3. NMR analysis of the (A) *delta1*-d(AG) and (B) *delta1*-cGAMP complexes. On the very top are the 1D NMR spectra for both complexes with assignments. Red indicates the imino protons of 5'-G-tetrad; blue for the middle G-tetrad; and green for 3'-G-tetrad. Second from the top are NOESY spectra (mixing time, 200 ms) recorded in H₂O. The characteristic H1-H8 cyclic connectivity patterns of G-tetrads are indicated. Third from the top are the 300-ms mixing time NOESY spectra recorded in D₂O. The H8_(n)-H1'_(n)-H8_(n+1) sequential walk are indicated, with red highlights showing the intra-residue cross peak of the filling guanine (G*) as well as its accompanying adenine (A*). Indicated by asterisks are the broken sequential connectivity. On the very bottom are the proposed schematics of the complexes. Cyan indicates *anti*-guanine; magenta *syn*-guanine; blue guanine; orange thymine; and green adenine.

Guanine base interaction is sufficient to maintain the *delta1*-cGAMP complex formation. The overall structure of the *delta1*-cGAMP complex is almost identical with that of *delta1*-d(AG). However, while the 5'-end side of the two complexes are indistinguishable, the 3'-end side differs as T17 was observed to not stack on the 3'-end G-tetrad and no interactions were observed between the A* residue of cGAMP and the T17 residue. The broad linewidths of aromatic protons of the bound cGAMP (Figure S8A) and the observation of intense exchange peak between the bound and free forms (Figure S8B) indicated a relatively fast exchange between the two states. Despite the broad signals, the cross-peaks between the guanine base protons of cGAMP and protons of *delta1* were still observed (Figure S8C). No other inter-molecular cross-peaks were observed, suggesting that the sole interaction of the guanine base with the vacant G-triad was sufficient to facilitate the ligand binding. The insertion of the guanine base (G*) of cGAMP into *delta1* and the lack of interaction between A* and *delta1* were illustrated in a representative structure (Figure 5C-F). The average dihedral angle (χ) of the guanine moiety in cGAMP molecule was computed to be 161.4°, corresponding to 8 *anti*-conformation guanines from the 10 lowest-energy structures, in agreement with previous computational and NMR study of free cyclic-dinucleotides⁶⁶. The average stacking distance between cGAMP bases (between G* and A*) was observed to be ~7.2 Å, approximately 2 times the corresponding distance in the linear dinucleotide (Figure S9). These observations rationalize the lack of interaction and cross-peaks between A* and T17 in the *delta1*-cGAMP complex structure.

Ionic dependency of the *delta1*-d(AG) formation. 1D ¹H NMR spectra of *delta1* in two different buffers (sodium buffer: 20 mM NaPi and 70 mM NaCl; lithium buffer: 10 mM LiPi and 10 mM LiCl) revealed a completely unfolded state of *delta1* (Figure S10). However, we found that upon addition of d(AG)—which only contain residual (< 1 mM) K⁺ ions—the G4 structures were immediately formed. The results indicated that the complex formation is robust, regardless of the presence of an excess amount of Na⁺ or Li⁺ ions.

Effect of a guanine-containing ligand on the kinetics of G4 formation. Semi-quantitative determination of folded fractions of the sequence were examined by performing 1D ¹H NMR experiments on two quenched samples: the free *delta1* and the *delta1*-d(AG) complex. The resulting NMR spectra revealed that the imino proton signals of the complex are cleaner, indicating a lower presence of other conformations (Figure S11). Given the same initial concentration of *delta1* in the two samples, the ratio of signal intensities for G4 structures with and without the ligand was calculated to be approximately 1.2 times. This suggested that the guanine-containing ligand could increase the folding rate and stability of the guanine-deficit G4s.

Calorimetric measurements revealed micromolar affinity of dinucleotides to *delta1*. ITC experiments were performed to measure the binding affinity of *delta1* with

Figure 4. (A) Side and (B) bottom view of ten lowest-energy structures of the *delta1*-d(AG) complex. Grey indicates phosphodiester backbone, cyan indicates anti-guanine; magenta *syn*-guanine; orange thymine; and green adenine. (C) Illustration of the insertion of d(AG) into *delta1* structure. The dinucleotide is shown with full opacity while the *delta1* structure is shown with 50% transparency. (D) Detailed view on the stacking between the A*•T17 base pair and the 3' G-tetrad (G*•G8•G12•G16). (E) Side and (F) bottom view of surface representation of *delta1*-d(AG) complex. Grey indicates *delta1* molecule and red indicates d(AG) molecule.

various ligands. The oligonucleotide concentration used was 600 μM, with the ligand stock concentration being generally ~2-4 mM. The results showed a dissociation constant (K_D) in the micromolar range (Figure S12, Table 2). The highest binding affinity was measured to be ~7 μM, corresponding to the d(AG) ligand, likely due to the additional A-T base pair formed in the *delta1*-d(AG) complex. Other dinucleotides, including the 3'3'-cGAMP were found to bind *delta1* with a K_D in ~30-70 μM range.

Interestingly, the dinucleotides with a 3'-guanine interacted better with the vacant site compared to those with a 5'-guanine (Table 2). However, this conclusion for *delta1* might not be valid for other (4*n*-1) G4 contexts. Moreover, we learnt that the binding of dGTP was undetectable in this concentration range, consistent with earlier study on the binding affinity of GTP to a G4-forming oligonucleotide, which resulted in a dissociation constant of 270 μM.⁶³ The overall data revealed a significantly stronger interaction of dinucleotides compared to mononucleotides, particularly dGTP. Indeed, the result agreed with an NMR titration

Figure 5. (A) Side and (B) bottom view of ten lowest-energy structures of the *delta1*-cGAMP complex. Grey indicates phosphodiester backbone, cyan indicates anti-guanine; blue guanine; orange thymine; and green adenine. (C) Illustration of the insertion of cGAMP into *delta1* structure. The dinucleotide is shown with full opacity while the *delta1* structure is shown with 50% transparency. (D) Detailed view on the A* and T17 residues with the 3' G-tetrad (G*•G8•G12•G16). (E) Side and (F) bottom view of surface representation of *delta1*-cGAMP complex. Grey indicates *delta1* molecule and blue indicates cGAMP molecule.

Table 2. Dissociation constant values of *delta1*-ligand complexes as measured by isothermal calorimetry.

Substrate	Ligand	K _D
<i>delta1</i>	d(AG)	7.4 ± 0.3
	d(GA)	69.5 ± 2.5
	d(TG)	34.9 ± 0.9
	d(GT)	71.2 ± 2.9
	3'3'cGAMP	61.7 ± 3.2
	dGTP	N/A

experiment (Figure S13) as well as a previous report indicating that dGTP binds relatively poorly to a vacancy-bearing G4 due to its abundant negative charges.⁴⁶

Thermal denaturation experiments suggest a G4 stabilization effect by the d(AG) ligand. A series of 1D ¹H NMR spectra were recorded for both *delta1* and *delta1*-d(AG) complex samples (3 equivalent amounts of ligand was added in the latter). Each series contained six spectra

which were recorded at different temperatures, ranging from 25°C to 50°C in 5°C interval. In the aromatic region (Figure S14A), qualitative observations can be made on the diminishing signal of the folded species and the elevating signal of the unfolded species, with the latter only observable in the case of free *delta1*. Comparing the two series of spectra, the free *delta1* series has significantly higher signal intensity from the unfolded species at 50°C compared to the bound *delta1* series, indicative of the higher melting temperature of bound *delta1* molecules.

The imino proton region showed peak broadening and decreasing peak intensity of upon increasing temperature (Figure S14B). In the *delta1*-d(AG) sample, we observed that the 12 peaks corresponding to the complex remained up to 50°C, while no appearance of the 10-peak free *delta1* species was observed, indicating that the complex is maintained at this temperature. More quantitative results were obtained by integrating representative peaks of the two series of spectra. The fraction of the folded species was plotted against temperature to derive the melting temperature of both free and bound *delta1* (see Materials and Methods, Figure S14C). The results indicated a ~15°C increase ($T_{m(\text{bound})} = 63.9^\circ\text{C}$, $T_{m(\text{free})} = 48.7^\circ\text{C}$) in the melting temperature of the *delta1*-d(AG) complex compared to the free *delta1*.

DISCUSSION

Guanine-G-triad binding mode is confirmed and generalizable. The high-resolution structures for both complexes confirm the theory of the guanine base binding mode into the vacant G-triad. The interaction indeed involves Hoogsteen hydrogen-bond formation, increasing the total number of hydrogen bonds between bases from 4 in the vacant G-triad to 8 in the resulting G-tetrad, therefore stabilizing the structure as seen by the NMR melting results. In fact, an increase in the melting temperature of a (4n-1) G4 in presence of guanine derivatives was previously observed via FRET-melting experiments.³⁸ The structure of *delta1*-cGAMP showed that a single guanine base binding interaction is sufficient to hold the cGAMP molecule tightly at room temperature. Results from previous studies suggested the same mode of binding for other guanine-containing small molecules with smaller or similar size such as guanosine, guanine, dGMP and dGTP³⁶⁻³⁷. We believe that this concept could be translated to various dinucleotides and small-molecule metabolites.

G4 structure triggering by guanine-containing ligands. The folding kinetics of some G4s in cells were shown to be equivalent or slower than many cellular processes⁶⁷. Kinetically favourable intermediate structures, such as triplex and G4 containing G-triad, could be formed during the folding. Theoretically, a fast-folding process of G4 in presence of guanine-containing dinucleotide or any other guanine derivatives could push towards a complex formation involving the described binding mode. In other words, the guanine-containing ligands could trigger the formation of G-quadruplexes on DNA sequences that would otherwise be in single-stranded or other types of structures. This hypothesis is supported by experimental data (Figure S6), which

showed increased folded fraction of G4 by approximately 1.2 times when the *delta1* sequence was quenched in the presence of the ligand d(AG). Moreover, the ~716,000 putative G4 sequences predicted by the sequencing technique⁵⁸ might include (4*n*-1) G4-forming sequences, but not the ones which would only form G4s in presence of guanine-containing ligands. Hence, controlling the ligand concentration might indirectly affect the formations of G-quadruplexes. We believe that the concept is not limited to only (4*n*-1) guanine sequences but also to include 4*n* guanine or even sequences with excess guanines.

Designing higher specificity G-quadruplex ligands.

The knowledge on the detailed interaction of the guanine base in the vacant G-triad allows several modes of ligand designs. For instance, attaching a guanine base on certain low-affinity G4 ligands might strengthen their bindings, by utilizing the vacant site as an anchor point. In addition, the position of the vacant site—which can be controlled by modifying the G4-forming sequences—may conceivably determine the location of the ligand binding.

Regulation of metabolic rate by (4*n*-1) G-quadruplex sequences. The concentration of common guanine-based metabolite (e.g. GTP) in cells was found to be in the millimolar range³⁹, which is two to three orders of magnitude higher than the dissociation constants of the *delta1*-dinucleotide complexes. Some of these guanine-derivative metabolites are compatible with the (4*n*-1) G-quadruplex structure, including a few cyclic-dinucleotides which were established to be important cell-signalling molecules in both prokaryotes and eukaryotes^{44, 68}. Taking the ~230,000 expected occurrences of (4*n*-1, *n*=3) G4s as a lower limit of guanine-deficient G4s and the ~100 (μm)³ average volume of mammalian cell nuclei,⁶⁹ the approximate concentration of (4*n*-1) G4 DNA sequences inside the cell nuclei was estimated to be > 3.2 μM, which is just slightly below their binding affinity to dinucleotides (*K_D* ~ 7-70 μM). As for guanine-containing ligands, processes such as nuclease digestions of DNA and RNA⁷⁰ and oligomerization of mononucleotides⁷¹ could be possible sources of di- and trinucleotides inside cells. For cyclic dinucleotides, synthesis of cGAMP from cytosolic DNA with mammalian cGAS enzyme was previously reported.⁴⁵ Furthermore, the second-messenger cyclic dinucleotides from bacteria cells could also find its way into the host cells,⁴¹ providing another source of cyclic dinucleotide ligands. These cyclic dinucleotides might undergo enzymatic hydrolyses, producing either mononucleotides or linear dinucleotides.⁷² To provide a gauge on the abundance of cyclic-dinucleotides, the concentration of cyclic-dinucleotide 2'3'cGAMP of ~1.5 μM was reported in L929 cell lysates.⁷³ Based on those estimations, binding between G-quadruplexes and cGAMP are possible, and potentially generalizable to other types of cyclic-dinucleotide metabolites. These binding processes directly influence the partition levels between bound-metabolites and free-metabolites which might be important in regulating certain metabolic rates in cells⁷⁴—as also observed in several RNA riboswitches⁴⁷—which potentially propagate into controlling gene expression or immune system response.

Sequestering metabolites and G-quadruplex activity. Previous research showed binding between different small molecules/metabolites and G4 DNA/RNA. For instance, G4 structures formed by G₄C₂ repeats were shown to sequester and activate hemes, which enhanced their oxidative processes upon formation of G4-heme complexes⁷⁵. On the other hand, the binding of natural cell metabolites such as flavin was speculated to affect G4 functionality which might translate into gene expressions⁷⁶. With the establishment of the complex formation between G4s and cGAMP, the above concepts could therefore be extended to cyclic dinucleotides, as well as other guanine-containing metabolites.

CONCLUSION

Solution structures of two complexes of a (4*n*-1) G-quadruplex *delta1* were described, one involved a linear dinucleotide d(AG) and the other a cyclic dinucleotide 3'3'cGAMP. Both structures have a similar binding mode where the guanine base from the dinucleotides filled the empty pocket of the G-triad from *delta1*, forming a complete G-tetrad. An A•T Hoogsteen base pair was identified in the first structure involving d(AG), while no additional interaction was detected in the second structure with cGAMP. Guanine base interaction with the (4*n*-1) G-quadruplex was shown to be sufficient in maintaining the small-molecule binding. The comprehensive structure elucidation may contribute towards designing higher specificity G-quadruplex ligands. The binding affinity of various dinucleotides was observed to be in the ~7-70 μM range compatible with *in vivo* binding, and thus they are possibly involved in biological functions.

ASSOCIATED CONTENT

Supporting Information

Supplementary Figures S1-S14

AUTHOR INFORMATION

Corresponding Author

* phantuan@ntu.edu.sg

ACKNOWLEDGMENT

This work was supported by the Singapore National Research Foundation Investigatorship (NRF-NRFI2017-09) and Nanyang Technological University (NTU Singapore) grants to A.T.P. The authors acknowledge the use of NMR facilities at NTU Institute of Structural Biology.

REFERENCES

1. Ren, J.; Chaires, J. B., Sequence and structural selectivity of nucleic acid binding ligands. *Biochemistry* **1999**, *38* (49), 16067-75.
2. Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H., Inhibition of human telomerase by a G-quadruplex-interactive compound. *J Med Chem* **1997**, *40* (14), 2113-6.
3. Monchaud, D.; Teulade-Fichou, M. P., A hitchhiker's guide to G-quadruplex ligands. *Org Biomol Chem* **2008**, *6* (4), 627-36.

4. Haider, S. M.; Neidle, S.; Parkinson, G. N., A structural analysis of G-quadruplex/ligand interactions. *Biochimie* **2011**, *93* (8), 1239-51.
5. Neidle, S., Quadruplex nucleic acids as targets for anticancer therapeutics. *Nat Rev Chem* **2017**, *1* (5), 0041.
6. Asamitsu, S.; Obata, S.; Yu, Z.; Bando, T.; Sugiyama, H., Recent Progress of Targeted G-Quadruplex-Preferred Ligands Toward Cancer Therapy. *Molecules* **2019**, *24* (3), 429.
7. Gellert, M.; Lipsett, M. N.; Davies, D. R., Helix formation by guanylic acid. *Proc Natl Acad Sci U S A* **1962**, *48*, 2013-8.
8. Sen, D.; Gilbert, W., Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **1988**, *334* (6180), 364-6.
9. Davis, J. T., G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew Chem Int Ed Engl* **2004**, *43* (6), 668-98.
10. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S., Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res* **2006**, *34* (19), 5402-15.
11. Patel, D. J.; Phan, A. T.; Kuryavyi, V., Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res* **2007**, *35* (22), 7429-55.
12. Mergny, J. L.; Sen, D., DNA Quadruple Helices in Nanotechnology. *Chem Rev* **2019**, *119* (10), 6290-6325.
13. Huppert, J. L.; Balasubramanian, S., Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* **2005**, *33* (9), 2908-16.
14. Todd, A. K.; Johnston, M.; Neidle, S., Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res* **2005**, *33* (9), 2901-7.
15. Maizels, N.; Gray, L. T., The G4 genome. *PLoS Genet* **2013**, *9* (4), e1003468.
16. Rhodes, D.; Lipps, H. J., G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res* **2015**, *43* (18), 8627-37.
17. Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M., Inhibition of telomerase by G-quartet DNA structures. *Nature* **1991**, *350* (6320), 718-20.
18. Cahoon, L. A.; Seifert, H. S., An alternative DNA structure is necessary for pilin antigenic variation in *Neisseria gonorrhoeae*. *Science* **2009**, *325* (5941), 764-7.
19. Fleming, A. M.; Burrows, C. J., 8-Oxo-7,8-dihydroguanine, friend and foe: Epigenetic-like regulator versus initiator of mutagenesis. *DNA Repair (Amst)* **2017**, *56*, 75-83.
20. Lopes, J.; Piazza, A.; Bermejo, R.; Kriegsman, B.; Colosio, A.; Teulade-Fichou, M. P.; Foiani, M.; Nicolas, A., G-quadruplex-induced instability during leading-strand replication. *EMBO J* **2011**, *30* (19), 4033-46.
21. Paeschke, K.; Capra, J. A.; Zakian, V. A., DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell* **2011**, *145* (5), 678-91.
22. Mukundan, V. T.; Phan, A. T., Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences. *J Am Chem Soc* **2013**, *135* (13), 5017-28.
23. Kerkour, A.; Marquevielle, J.; Ivashchenko, S.; Yatsunyk, L. A.; Mergny, J. L.; Salgado, G. F., High-resolution three-dimensional NMR structure of the KRAS proto-oncogene promoter reveals key features of a G-quadruplex involved in transcriptional regulation. *J Biol Chem* **2017**, *292* (19), 8082-8091.
24. Sengar, A.; Vandana, J. J.; Chambers, V. S.; Di Antonio, M.; Winnerdy, F. R.; Balasubramanian, S.; Phan, A. T., Structure of a (3+1) hybrid G-quadruplex in the PARP1 promoter. *Nucleic Acids Res* **2019**, *47* (3), 1564-1572.
25. Lim, K. W.; Phan, A. T., Structural basis of DNA quadruplex-duplex junction formation. *Angew Chem Int Ed Engl* **2013**, *52* (33), 8566-9.
26. Kang, H. J.; Cui, Y.; Yin, H.; Scheid, A.; Hendricks, W. P. D.; Schmidt, J.; Sekulic, A.; Kong, D.; Trent, J. M.; Gokhale, V.; Mao, H.; Hurley, L. H., A Pharmacological Chaperone Molecule Induces Cancer Cell Death by Restoring Tertiary DNA Structures in Mutant hTERT Promoters. *J Am Chem Soc* **2016**, *138* (41), 13673-13692.
27. Onel, B.; Carver, M.; Wu, G.; Timonina, D.; Kalarn, S.; Larriva, M.; Yang, D., A New G-Quadruplex with Hairpin Loop Immediately Upstream of the Human BCL2 P1 Promoter Modulates Transcription. *J Am Chem Soc* **2016**, *138* (8), 2563-70.
28. Kettani, A.; Bouaziz, S.; Wang, W.; Jones, R. A.; Patel, D. J., Bombyx mori single repeat telomeric DNA sequence forms a G-quadruplex capped by base triads. *Nat Struct Biol* **1997**, *4* (5), 382-9.
29. Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J., Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. *Nat Chem Biol* **2005**, *1* (3), 167-73.
30. Matsugami, A.; Xu, Y.; Noguchi, Y.; Sugiyama, H.; Katahira, M., Structure of a human telomeric DNA sequence stabilized by 8-bromoguanosine substitutions, as determined by NMR in a K⁺ solution. *FEBS J* **2007**, *274* (14), 3545-3556.
31. Caceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A., A thymine tetrad in d(TGGGGT) quadruplexes stabilized with Tl⁺/Na⁺ ions. *Nucleic Acids Res* **2004**, *32* (3), 1097-102.
32. Kocman, V.; Plavec, J., Tetrahelical structural family adopted by AGCGA-rich regulatory DNA regions. *Nat Commun* **2017**, *8*, 15355.
33. Chung, W. J.; Heddi, B.; Schmitt, E.; Lim, K. W.; Mechulam, Y.; Phan, A. T., Structure of a left-handed DNA G-quadruplex. *Proc Natl Acad Sci U S A* **2015**, *112* (9), 2729-33.
34. Bakalar, B.; Heddi, B.; Schmitt, E.; Mechulam, Y.; Phan, A. T., A Minimal Sequence for Left-Handed G-Quadruplex Formation. *Angew Chem Int Ed Engl* **2019**, *58* (8), 2331-2335.
35. Martin-pintado, N.; Heddi, B.; Teuku, A.K., Serimbetov, Z., and Phan, A.T. In *G-quadruplexes with (4n-1) guanines in the G-tetrad core: formation of a G-triad.H₂O complex*, 5th International Meeting on Quadruplex Nucleic Acids: G4thering in Bordeaux, Bordeaux, France, May 26-28; Bordeaux, France, 2015.
36. Heddi, B.; Martin-Pintado, N.; Serimbetov, Z.; Kari, T. M.; Phan, A. T., G-quadruplexes with (4n - 1) guanines in the G-tetrad core: formation of a G-triad.water complex and implication for small-molecule binding. *Nucleic Acids Res* **2016**, *44* (2), 910-6.
37. Pang, S. M.; Heddi, B.; Martin-pintado, N.; Serimbetov, Z., and Phan, A.T. In *Interaction between guanosine and a G-quadruplex with a vacant site*, 5th International Meeting on Quadruplex Nucleic Acids: G4thering in Bordeaux, Bordeaux, France, May 26-28; Bordeaux, France, 2015.
38. Li, X. M.; Zheng, K. W.; Zhang, J. Y.; Liu, H. H.; He, Y. D.; Yuan, B. F.; Hao, Y. H.; Tan, Z., Guanine-vacancy-bearing G-quadruplexes responsive to guanine derivatives. *Proc Natl Acad Sci U S A* **2015**, *112* (47), 14581-6.
39. Bennett, B. D.; Kimball, E. H.; Gao, M.; Osterhout, R.; Van Dien, S. J.; Rabinowitz, J. D., Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol* **2009**, *5* (8), 593-9.
40. Wu, J.; Sun, L.; Chen, X.; Du, F.; Shi, H.; Chen, C.; Chen, Z. J., Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science* **2013**, *339* (6121), 826-30.
41. Danilchanka, O.; Mekalanos, J. J., Cyclic dinucleotides and the innate immune response. *Cell* **2013**, *154* (5), 962-970.
42. D'Argenio, D. A.; Miller, S. I., Cyclic di-GMP as a bacterial second messenger. *Microbiology* **2004**, *150* (Pt 8), 2497-502.
43. Tischler, A. D.; Camilli, A., Cyclic diguanylate [c-di-GMP] regulates *Vibrio cholerae* biofilm formation. *Mol Microbiol* **2004**, *53* (3), 857-69.

44. Camilli, A.; Bassler, B. L., Bacterial small-molecule signaling pathways. *Science* **2006**, *311* (5764), 1113-6.
45. Diner, E. J.; Burdette, D. L.; Wilson, S. C.; Monroe, K. M.; Kellenberger, C. A.; Hyodo, M.; Hayakawa, Y.; Hammond, M. C.; Vance, R. E., The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. *Cell Rep* **2013**, *3* (5), 1355-61.
46. Li, X. M.; Zheng, K. W.; Hao, Y. H.; Tan, Z., Exceptionally Selective and Tunable Sensing of Guanine Derivatives and Analogues by Structural Complementation in a G-Quadruplex. *Angew Chem Int Ed Engl* **2016**, *55* (44), 13759-13764.
47. Nudler, E.; Mironov, A. S., The riboswitch control of bacterial metabolism. *Trends Biochem Sci* **2004**, *29* (1), 11-7.
48. Goddard, T. D.; Ferrin, T. E., Visualization software for molecular assemblies. *Curr Opin Struct Biol* **2007**, *17* (5), 587-95.
49. Ren, A.; Wang, X. C.; Kellenberger, C. A.; Rajashankar, K. R.; Jones, R. A.; Hammond, M. C.; Patel, D. J., Structural basis for molecular discrimination by a 3',3'-cGAMP sensing riboswitch. *Cell reports* **2015**, *11* (1), 1-12.
50. Hall, J.; Ralph, E. C.; Shanker, S.; Wang, H.; Byrnes, L. J.; Horst, R.; Wong, J.; Brault, A.; Dumlao, D.; Smith, J. F.; Dakin, L. A.; Schmitt, D. C.; Trujillo, J.; Vincent, F.; Griffor, M.; Aulabaugh, A. E., The catalytic mechanism of cyclic GMP-AMP synthase (cGAS) and implications for innate immunity and inhibition. *Protein Sci* **2017**, *26* (12), 2367-2380.
51. Kato, K.; Nishimasu, H.; Oikawa, D.; Hirano, S.; Hirano, H.; Kasuya, G.; Ishitani, R.; Tokunaga, F.; Nureki, O., Structural insights into cGAMP degradation by Ecto-nucleotide pyrophosphatase phosphodiesterase 1. *Nat Commun* **2018**, *9* (1), 4424.
52. Kranzusch, P. J.; Wilson, S. C.; Lee, A. S.; Berger, J. M.; Doudna, J. A.; Vance, R. E., Ancient Origin of cGAS-STING Reveals Mechanism of Universal 2',3' cGAMP Signaling. *Mol Cell* **2015**, *59* (6), 891-903.
53. Krieger, E.; Koraimann, G.; Vriend, G., Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. *Proteins* **2002**, *47* (3), 393-402.
54. Kleywegt, G. J.; Henrick, K.; Dodson, E. J.; van Aalten, D. M., Pound-wise but penny-foolish: How well do micromolecules fare in macromolecular refinement? *Structure* **2003**, *11* (9), 1051-9.
55. Sousa da Silva, A. W.; Vranken, W. F., ACPYPE - AnteChamber PYthon Parser interfAcE. *BMC Res Notes* **2012**, *5*, 367.
56. Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M., The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **2003**, *160* (1), 65-73.
57. Bedrat, A.; Lacroix, L.; Mergny, J. L., Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res* **2016**, *44* (4), 1746-59.
58. Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S., High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat Biotechnol* **2015**, *33* (8), 877-81.
59. Do, N. Q.; Phan, A. T., Monomer-dimer equilibrium for the 5'-5' stacking of propeller-type parallel-stranded G-quadruplexes: NMR structural study. *Chemistry* **2012**, *18* (46), 14752-9.
60. Kypr, J.; Kejnovska, I.; Renciuik, D.; Vorlickova, M., Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res* **2009**, *37* (6), 1713-25.
61. Karsiotis, A. I.; Hessari, N. M.; Novellino, E.; Spada, G. P.; Randazzo, A.; Webb da Silva, M., Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew Chem Int Ed Engl* **2011**, *50* (45), 10645-8.
62. Del Villar-Guerra, R.; Trent, J. O.; Chaires, J. B., G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew Chem Int Ed Engl* **2018**, *57* (24), 7171-7175.
63. Curtis, E. A.; Liu, D. R., Discovery of widespread GTP-binding motifs in genomic DNA and RNA. *Chem Biol* **2013**, *20* (4), 521-32.
64. Phan, A. T.; Patel, D. J., A site-specific low-enrichment (15)N,(13)C isotope-labeling approach to unambiguous NMR spectral assignments in nucleic acids. *J Am Chem Soc* **2002**, *124* (7), 1160-1.
65. Phan, A. T., Long-range imino proton-13C J-couplings and the through-bond correlation of imino and non-exchangeable protons in unlabeled DNA. *J Biomol NMR* **2000**, *16* (2), 175-8.
66. Wang, B.; Wang, Z.; Javornik, U.; Xi, Z.; Plavec, J., Computational and NMR spectroscopy insights into the conformation of cyclic di-nucleotides. *Sci Rep* **2017**, *7* (1), 16550.
67. Zhang, A. Y.; Balasubramanian, S., The kinetics and folding pathways of intramolecular G-quadruplex nucleic acids. *J Am Chem Soc* **2012**, *134* (46), 19297-308.
68. Ablasser, A.; Goldeck, M.; Cavlar, T.; Deimling, T.; Witte, G.; Rohl, I.; Hopfner, K. P.; Ludwig, J.; Hornung, V., cGAS produces a 2'-5'-linked cyclic dinucleotide second messenger that activates STING. *Nature* **2013**, *498* (7454), 380-4.
69. Alberts, B.; Johnson, A. D.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P., *Molecular Biology of the Cell: Sixth International Student Edition*. Garland Science, Taylor and Francis Group: 2014.
70. Markham, R.; Smith, J. D., The structure of ribonucleic acids. II. The smaller products of ribonuclease digestion. *Biochem J* **1952**, *52* (4), 558-65.
71. Kozlov, I. A.; Orgel, L. E., [Nonenzymatic matrix synthesis of RNA from monomers]. *Mol Biol (Mosk)* **2000**, *34* (6), 921-30.
72. Romling, U.; Liang, Z. X.; Dow, J. M., Progress in Understanding the Molecular Basis Underlying Functional Diversification of Cyclic Dinucleotide Turnover Proteins. *J Bacteriol* **2017**, *199* (5).
73. Bose, D.; Su, Y.; Marcus, A.; Raulet, D. H.; Hammond, M. C., An RNA-Based Fluorescent Biosensor for High-Throughput Analysis of the cGAS-cGAMP-STING Pathway. *Cell Chem Biol* **2016**, *23* (12), 1539-1549.
74. Ottaway, J. H., Sequestration of metabolites: insights into metabolic control. *Biochem Soc Trans* **1979**, *7* (5), 1161-7.
75. Grigg, J. C.; Shumayrikh, N.; Sen, D., G-quadruplex structures formed by expanded hexanucleotide repeat RNA and DNA from the neurodegenerative disease-linked C9orf72 gene efficiently sequester and activate heme. *PLoS One* **2014**, *9* (9), e106449.
76. Merkle, T.; Sinn, M.; Hartig, J. S., Interactions between Flavins and Quadruplex Nucleic Acids. *Chembiochem* **2015**, *16* (17), 2437-40.

Insert Table of Contents artwork here