13422–13429 *Nucleic Acids Research, 2014, Vol. 42, No. 21* *Published online 11 November 2014*
doi: 10.1093/nar/gku1122

Crystal structure of a DNA/Ba^{2+} G-quadruplex containing a water-mediated C-tetrad

Diana Zhang[1], Terry Huang[2], Philip S. Lukeman[3] and Paul J. Paukstelis[1,*]

[1]Department of Chemistry & Biochemistry, Center for Biomolecular Structure and Organization, Maryland NanoCenter, University of Maryland, College Park, MD 20742, USA, [2]Chemistry and Biochemistry Department, California State Polytechnic University, 3801 West Temple Avenue, Pomona, CA 91768, USA and [3]Chemistry Department, St. John's University, 8000 Utopia Parkway, Queens, NY 11439, USA

Received August 12, 2014; Revised October 20, 2014; Accepted October 25, 2014

ABSTRACT

We have determined the 1.50 Å crystal structure of the DNA decamer, d(CCA^{CNV}KGCGTGG) (^{CNV}K, 3-cyanovinylcarbazole), which forms a G-quadruplex structure in the presence of Ba^{2+}. The structure contains several unique features including a bulged nucleotide and the first crystal structure observation of a C-tetrad. The structure reveals that water molecules mediate contacts between the divalent cations and the C-tetrad, allowing Ba^{2+} ions to occupy adjacent steps in the central ion channel. One ordered Mg^{2+} facilitates 3′-3′ stacking of two quadruplexes in the asymmetric unit, while the bulged nucleotide mediates crystal contacts. Despite the high diffraction limit, the first four nucleotides including the ^{CNV}K nucleoside are disordered though they are still involved in crystal packing. This work suggests that the bulky hydrophobic groups may locally influence the formation of non-Watson–Crick structures from otherwise complementary sequences. These observations lead to the intriguing possibility that certain types of DNA damage may act as modulators of G-quadruplex formation.

INTRODUCTION

A significant proportion of our understanding of DNA structure derives from directed studies of Watson–Crick duplexes due to their importance in genetic inheritance and gene expression. However, there has been a growing appreciation for the structural polymorphism of DNA, and the role of alternate DNA structures in cellular function (1). The most well-studied of these DNA structures are G-quadruplexes (2). G-quadruplexes are columnar structures formed from core G-tetrad motifs (3). This planar motif is composed of four guanines associated by hydrogen bonding interactions; the motif is stabilized by base stacking between multiple G-tetrad planes and by electrostatic interactions with cations positioned in a central channel of the G-tetrad core (4,5). G-quadruplexes have been found *in vivo* at chromosomal telomeres (6–10) and at many gene promoters (11), which has led to significant interest in finding ways to control or modulate their formation (12).

Structural and biophysical studies of G-rich DNA and RNA sequences initially led to a common motif for intramolecular G-quadruplexes being composed of four runs of three or more guanosine residues, separated by loops of varying lengths. The $G_{3+}L_1G_{3+}L_2G_{3+}L_3G_{3+}$ sequence motif is commonly used for identifying potential G-quadruplex structures in bioinformatic studies (11); however, structural studies of G-quadruplexes have revealed that non-G-tetrads and different structural organization of unpaired residues can lead to greater complexity in G-quadruplex-forming sequences. Non-G-tetrads have been observed within (or associated with) both DNA and RNA G-quadruplexes. Uridine was the first nucleoside to be reported in a non-G-tetrad (13), and has subsequently been found in a variety of quadruplex solution and crystal structures, suggesting it is well accommodated at the ends of RNA quadruplexes (14–17). Both symmetric (18) and non-symmetric T-tetrads (19) have been reported, though the symmetric form appears to be less well accommodated within DNA structures due to steric clashes between O2 and adjacent methyls (19). Several forms of A-tetrads in DNA quadruplexes have been observed in solution for truncated sequences of human telomeres (20), while the crystal structure of an A-tetrad containing RNA has also been determined (16). Unlike the other non-G-tetrads, C-tetrads have only been observed in DNA solution studies (21,22). These studies demonstrated that a C-tetrad could be accommodated between G-tetrad planes, through the formation of N4-O2 hydrogen bonds between neighbors. This C-tetrad contained an enlarged central channel space and it was unclear if or how the C-tetrad interacted with any central cations. One initial suggestion was protonation of the cytosine N3 position could result in a delocalized positive charge in the C-tetrad

*To whom correspondence should be addressed. Tel: 301.405.9933; Fax: 301.314.0386; Email: paukstel@umd.edu

(22). This is consistent with the ability for cytosine to form other alternate base pairs dependent on N3 protonation, include the i-motif which is dependent on hemiprotonation in the C-C(+) parallel base pair (23,24). A variety of hetero-quadruplex base pairs have also been observed (25–28).

In quadruplexes, loops connect the corners of an intramolecular quadruplex column, while bulges connect adjacent guanines within the column. G-quadruplexes often contain unpaired nucleotides in loop regions, though recent evidence suggests that bulged nucleotides can be accommodated at various positions in G-quadruplex structures. Bulges have been observed in the crystal structure of several related intermolecular RNA quadruplexes (29), and a recent solution study has shown that bulges with different identities can be well tolerated at multiple positions (30). The fact that any of the standard nucleotides can form non-G-tetrads and the fact that profoundly different structural organization of bulged residues has been observed, lead us to conclude that there is significant diversity in G-quadruplex-forming sequences and structures.

Here we add to the sequence and structural diversity of DNA quadruplexes by describing a high-resolution crystal structure of a DNA quadruplex that contains both internal non-G-tetrads and a bulged nucleotide. The motivation of our study was to explore how a cyanovinylcarbazole (^{CNV}K) nucleoside substitution could be incorporated into an otherwise self-complementary decameric DNA sequence. The ^{CNV}K nucleoside is an effective general photoreversible crosslinker of complementary DNA and RNA strands (when placed opposite to a pyrimidine and with certain sequence preferences adjacent to the ^{CNV}K), (31,32) which has found diverse uses (33–36). Surprisingly, the crystals we obtained prior to photocrosslinking formed a tetrameric quadruplex instead of a DNA duplex structure.

This DNA quadruplex has a number of unique features. This is the first crystal structure of a DNA G-quadruplex structure with Ba^{2+} in the central ion channel. Significantly, it is also the first crystal structure containing a C-tetrad. The structure reveals that water molecules mediate contacts between the divalent cations and the C-tetrad. One ordered Mg^{2+} facilitates 3′-3′ stacking of two quadruplexes, while a single bulged nucleotide in each strand mediates crystal contacts. Surprisingly, despite the high diffraction limit, the first four nucleotides are fully or partially disordered, including the ^{CNV}K monomer. This study highlights the structural plasticity of DNA and its ability to undertake non-canonical structures, even in the context of a highly self-complementary sequence. The demonstration of a C-tetrad under these environmental conditions may further expand the sequence repertoire of G-quadruplex-forming structures. Though it remains unclear how significant the ^{CNV}K nucleoside is to the formation of the quadruplex, this structure raises the possibility that non-standard or nucleobase lesions may promote quadruplex formation in otherwise non-traditional G-quadruplex-forming sequences.

MATERIALS AND METHODS

Synthesis and purification

The DNA decamer was synthesized trityl-on on a solid-phase DNA synthesizer (ABI 394) using ultramild phosphoramidites, columns and reagents (Glen Research, Sterling, VA USA). Standard phosphoramidite protocols for the 1 μmol scale were used, except the coupling time per monomer was extended to 180 s. The sample was cleaved from the column using ∼3 ml 33% aqueous ammonia at room temperature and deprotected overnight in the same solution. The sample solution was then evaporated to ∼1/4 of the volume under reduced pressure.

The crude sample was desalted using a G-10 column, evaporated and resuspended in water prior to loading on a preparative XBridge C18 column (Waters, 2.5 μm particle size, 4.6 mm ID, 75 mm length—equipped with a pre-column) controlled by an HPLC system (Agilent 1100) with in-line ultraviolet detection. Elution was performed with a mixed solvent system run at 72°C with an integrated column heater. Solvents used in gradient elution were 0.1 M triethylamine acetate (Solvent A) and HPLC-grade Acetonitrile (Solvent B). Typically, gradients running from A:B 95:5 to 70:30 over 45′ were used, with a wash of 5:95 for 5′ at the end of each run to remove hydrophobic impurities from the column. The fractions containing decamer as determined by HPLC were pooled, detritylated at room temperature for 1 h in 80% acetic acid and desalted using a G-10 column, and characterized by MALDI-MS (Supplementary Figure S1).

Crystallization

Crystals were grown by sitting drop vapor diffusion. DNA decamer oligonucleotide (650 μM in water) was mixed in a 1:1 ratio with the crystallization solution (0.01 M magnesium chloride hexahydrate, 0.002 M barium chloride dihydrate, 0.05 M MOPS pH 7.0, 30% v/v 1,4-dioxane). No other divalent cations were included in the sample or crystallization conditions. Crystals grew as thin square plates to a maximum size of 100 μm within 2–3 weeks at 22°C. Crystals were flash cooled directly in liquid nitrogen, with dioxane from the crystallization solution serving as a cryoprotectant.

Data collection and structure determination

Data were collected at Advanced Photon Source, Argonne National Labs Sector 22-BM at a X-ray energy of 12398.4 eV (1.0 Å). Data were processed with HKL2000 (37). Barium atoms substructure detection was performed with HySS followed by density modification by RESOLVE as part of the Phenix macromolecular crystallography package (38). Model building was performed in Coot (39). Refinement was performed with Refmac5 using Babinet bulk solvent correction without non-crystallographic symmetry restraints (40). Data collection and refinement statistics are given in Supplementary Table S1.

RESULTS AND DISCUSSION

Crystal organization

The decameric oligonucleotide crystallized with two molecules in the asymmetric unit. Both molecules reside on the 4-fold symmetry axis to create two nearly identical tetrameric parallel-stranded quadruplexes (RMSD 0.47 Å for all DNA atoms). Each quadruplex is composed of

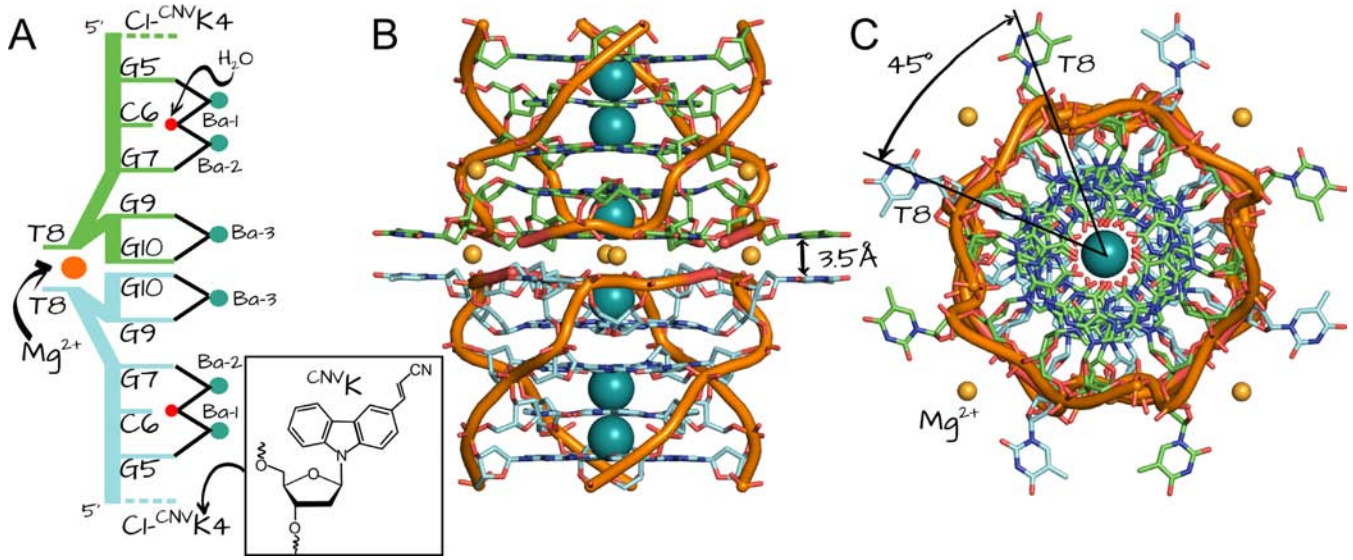

Figure 1. A G-quadruplex crystal structure formed from d(CCA^{CNV}KGCGTGG). (**A**) Schematic representation of the two molecules in the asymmetric unit. The two DNA molecules (green, light blue) are stacked in a tail-to-tail arrangement and are stabilized by Mg^{2+} (orange sphere) interactions with two phosphates from each monomer. Dashes represent disordered residues; black lines indicate coordination interactions between nucleobases and solvent molecules (water, red; Ba^{2+}, cyan). The ^{CNV}K nucleoside is shown in the inset box. (**B**) Cartoon representation seen perpendicular to the 4-fold symmetry axis. Bulged T8 residues are offset by 3.5 Å. (**C**) A view down the 4-fold symmetry axis that is coincident with the central ion channel of the quadruplex. The 45° offset between T8 residues allows stacking interactions between translated quadruplexes.

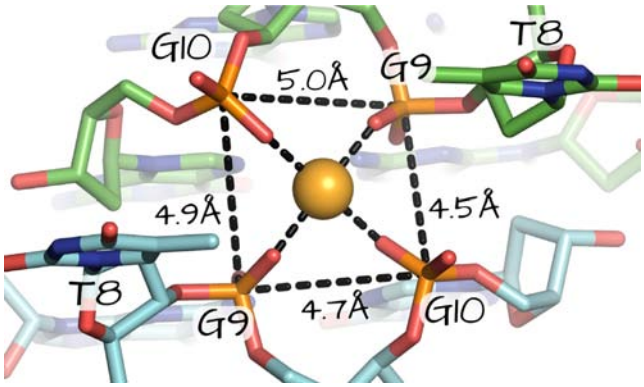

Figure 2. Mg^{2+} coordination between DNA strands. One Mg^{2+} (orange sphere) is coordinated between G9 and G10 phosphates of both monomers in the asymmetric unit (black dashes). Two water molecules (not shown) complete the octahedral coordination. Phosphate-phosphate distances are shown.

four identical symmetry-related chains that are co-axially stacked in a 3′-to-3′ arrangement (Figure 1A). Similar to other crystal structures (14,29), non-core nucleotide residues stabilize adjacent quadruplexes through stacking interactions perpendicular to the central ion channel axis. T8 bulges out of the quadruplex core and is the only unpaired ordered nucleotide in this structure. The T8 nucleobases are co-planar with their respective G10 quadruple pairs, and are offset by 45° with respect to its quadruplex stacking partner (Figure 1C). This results in T8 nucleobases sticking out every 45° from the central 3′-to-3′ stack junction, with an alternating 3.5 Å offset. This allows adjacent quadruplex stacks to interleave through T8 stacking interactions to form a planar 2D array of stacked quadruplex sheets (Supplementary Figure S2).

Monomer interactions

The two monomers in the asymmetric unit interact through nucleobase stacking and ion-mediated contacts. G10 from each strand stacks with the equivalent residue in the other monomer through 6-ring stacking interactions, consistent with previously described 3′-to-3′ stacking interactions and predicted geometric and molecular dynamic simulations (41). The intrachain phosphate distances between G9 and G10 are relatively short (P-P distance 5.0 and 4.7 Å for chains A and B, respectively), as are the interchain distances between these same phosphates (P-P distances: 4.5 Å between G9 of chain A and G10 of chain B, and 4.9 Å G9 of chain B and G10 of chain A). One Mg^{2+} ion coordinates these four phosphates and facilitates 3′-3′ stacking (Figure 2). The octahedral coordination around this Mg^{2+} ion is completed by two water molecules, with one water proximal and one distal to the G-quadruplex core. The proximal water is positioned to make additional contacts between O1P of the G9 phosphates, while also being within the van Der Waal radii (2.3 Å) with H8 from a symmetry-related G10 nucleobase. The O3′ of the two G10 monomers are within hydrogen bonding distance (2.65 Å) and may further stabilize the stacked monomer interactions.

Tetrad pairings

The quadruplexes in the structure consist of five parallel quadruple base pairs. G5, G7, G9 and G10 form parallel G-tetrad structures with all of the nucleobases in the *anti* conformation. Intercalated between the G5 and G7 tetrads is a tetrad formed from C6. The C6 nucleobases in the tetrad

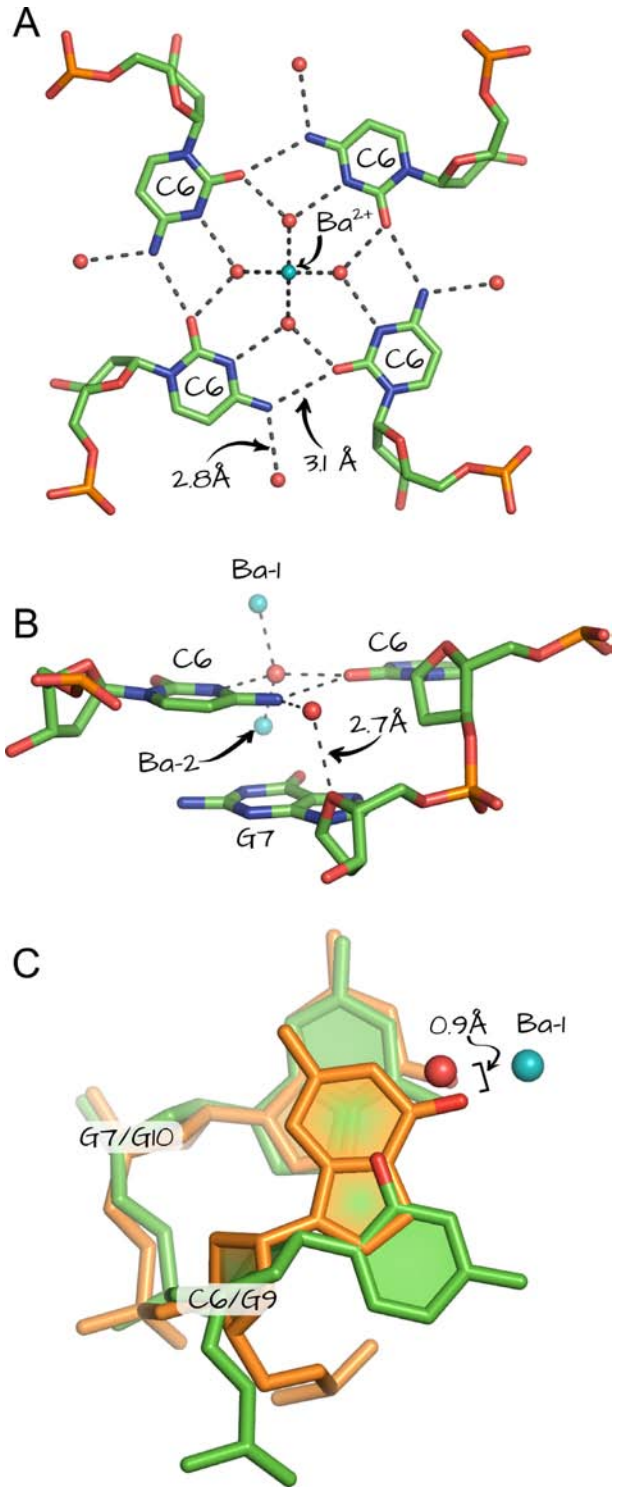

Figure 3. A C-tetrad coordinated to the central ion channel. (**A**) The C-tetrad forms around the central ion channel through direct hydrogen bonding and water-mediated interactions. Water molecules are shown as red spheres. Hydrogen bonds and coordination interactions are shown as dashes. The central water molecule mediates hydrogen bonding between nucleobases in the tetrad, while simultaneously coordinating interactions with two Ba^{2+} ions in the central channel. (**B**) The exterior water molecule forms bridging hydrogen bonds between N4 of each C6 and O4′ of G7 from a partner strand. (**C**) Superposition of the G9-G10 dinucleotide (yellow carbons) on the C6-G7 dinucleotide (green carbons). The central water molecule associated with C6 is positionally equivalent to the O6 of G9.

interact both directly and through solvent-mediated interactions (Figure 3A). Direct interaction between C6 residues is through a long (3.11 Å) N4-O2 H-bond, as has been described previously for C-tetrads (21,22). One central water molecule is associated between each C6 residue. This water forms hydrogen bonds between N3 of one C6 and O2 of an adjacent C6 partner, while also being coordinated between two divalent cations in the central ion channel (Figure 3B; see below). Additionally, each N4 donates a proton in a hydrogen bond with a second water molecule that in turn donates a proton to G7 O4′ of a partner strand (Figure 3B).

The water molecules associated with this C-tetrad are in the same position as the O6 of the guanosine nucleobases in a typical G-tetrad. The distances between the water and two metal ions are 2.77 and 2.81 Å, almost identical to the distance between G(O6) and central divalent cations observed in this and other structures (average distance in this structure, 2.72 Å). The positional equivalence is clear from the least-squares superposition of the G10 tetrad plane on G7. This equivalence effectively places G9 over the C6 tetrad (Figure 3C). In this overlap, the G9 O6 is only 0.9 Å away from the C-tetrad-associated water. Interestingly, despite this positional equivalence, divalent cations are generally only found between every other tetrad step, while in this case the cations occupy adjacent steps. This suggests that these water molecules may not be functionally equivalent to guanosine O6 with respect to their interaction with the divalent cations in the central ion channel.

Central ion channel

The central ion channel contains three barium ions per molecule in the asymmetric unit. For reference, they are numbered with respect to the 5′ ends of the quadruplexes (Ba-1, Ba-2, Ba-3). The position and identity of these barium ions were confirmed by anomalous X-ray diffraction. Despite a relatively modest anomalous signal for barium at the data collection energy, it was sufficient to generate exceptional SAD-phased experimental electron density maps (Supplementary Figure S3) and to confirm the positions of the divalent cation in the central ion channel through anomalous difference electron density maps (Figure 3A).

Each G-tetrad plane is associated with one Ba^{2+} ion. The G-tetrad planes from G9 and G10 are coordinated by Ba-3, while the G5 and G7 tetrads are associated with Ba-1 and Ba-2, respectively. Similar to other G-tetrad structures with central divalent cations, the ions are positioned between tetrad planes, leading to a slight buckling of the G-tetrad planes toward the cations. This is most apparent between the G7 and G9 tetrads, and it may be further influenced by several factors, including the bulged T8 residue between them and the presence of a water molecule in the central ion channel that is coordinated to Ba-2 (Figure 4A). As described above, the water-mediated C6-tetrad is coordinated between Ba-1 and Ba-2, placing these two divalent cations in adjacent steps only 3.86 Å apart.

The adjacent Ba-1 and Ba-2 metal pair in the central channel shows distinct coordination geometries. Coordination for Ba-1 between the G5 O6 and the C6-bound waters shows bipyramidal antiprismatic geometry, while the C6-directed waters and the G7 O6 show a nearly square

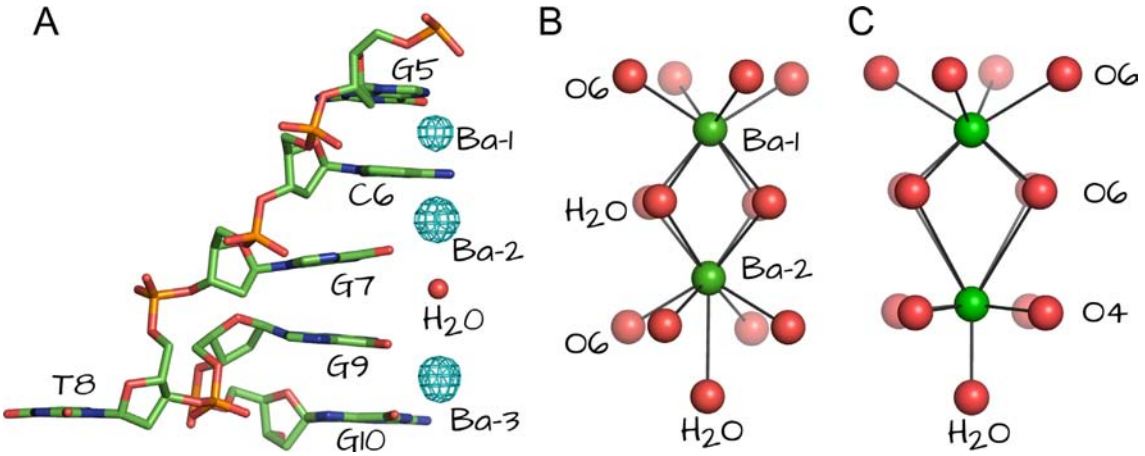

Figure 4. Barium in the central ion channel. (**A**) Stick model of one quadruplex forming strand shown with anomalous difference density (cyan) contoured at 5σ. A water molecule coordinated to Ba-2 is shown as a red sphere. G7 and G9 show buckling toward Ba-2 and Ba-3, respectively. (**B**) Oxygen coordination around the adjacent barium ions (green) in the structure described here. (**C**) Oxygen coordination around adjacent barium ions in an RNA quadruplex structure (PDBID: 1J6S). All oxygen atoms (depth-cued) are in red and their source, either nucleobase (guanosine O6, or uracil O4) or water is indicated.

prismatic coordination geometry around Ba-2 with an additional capping solvent molecule (Figure 4B). A similar arrangement of barium ions was observed in an RNA quadruplex (16), though in that case the distance between adjacent Ba ions is larger (4.42 Å) and the Ba-2 equivalent is directly coordinated by the O4 atoms of a highly buckled U-tetrad, resulting in a collapsed prismatic geometry (Figure 4C).

To our knowledge, these are only two examples of quadruplex structures determined with divalent cations in adjacent positions within the central ion channel, and both examples are associated with pyrimidine tetrads. This structure suggests that the water-mediated C-tetrad interactions may be favorable for stabilizing adjacent divalent cations. Waters associated with the C-tetrad donate both of their protons to different cytosines. Donating both protons increases the electron density at the water oxygen's lone pairs, making this oxygen suitable for shielding charge between the adjacent barium ions. Cytosine is unique in that it can accept both water protons between N3 and O2 in this arrangement, and the positioning of these more electronegative waters is likely better than the keto oxygen of G at coordinating adjacent divalent cations. Further, the geometry of this C-tetrad is more planar than the highly buckled U-tetrads. This suggests that C-tetrads could be more easily accommodated within a standard G-quadruplex, as the stacking interactions with adjacent tetrad planes could be maintained.

Comparison to solution structure C-tetrads

This crystal structure is consistent with the overall structures described in solution studies of C-tetrad-containing quadruplexes (21,22), but clarifies several critical features. The most striking difference is the water-mediated interaction that bridges direct contacts between the nucleobases and between the nucleobases and the central ion channel. Previous solution studies indicated that C-tetrad structures could be sandwiched between flanking G-tetrads; however, no analysis was provided for how the C-tetrad might interact with the central ion channel. For d(TGGCGGC), modest pH-dependent changes in NMR chemical shifts for the N3 of the tetrad cytosines suggested these cytosines were protonated in the quadruplex, despite lacking the typical resonances of protonated N3 (22). This led to the suggestion of a delocalized proton present in the central ion channel that remained associated with the C-tetrad. Though there may be differences between how a C-tetrad interacts in the context of a monovalent central ion channel, this structure clearly shows how water molecules can be positioned in the larger C-tetrad hole and how they can mediate interactions between the nucleobases and ions. The observation of a C-tetrad in solution and now a crystal structure suggests that this may be a predictable motif, and adds to the growing list of homocytidine nucleobase interactions (23,42,43).

Disordered residues

Despite the high diffraction limit of these crystals, the four 5′-most nucleotides of both monomers in the asymmetric unit were disordered. There was strong difference density for the ^{CNV}K4 nucleotide (Figure 5A) and additional planes of difference density separated by approximately 3.2 Å, along the 4-fold symmetry axis between the quadruplex layers (Figure 5B). We attempted to build the ^{CNV}K4 residue, but there was substantial overlap of the ^{CNV}K groups (Figure 5A). We explored the possibility that the bulkiness of the ^{CNV}K group precluded stacking of all four ^{CNV}K4 nucleobases along the 4-fold symmetry axis, and that only two of these residues were responsible for symmetry-averaged density. However, our attempts to build the ^{CNV}K4 nucleotides with partial occupancy did not improve the refinement and they were omitted from the final model. Though the first four residues could not be built, presumably the ∼3.2 Å separation of difference density along the 4-fold axis corresponds to nucleobases of the missing nucleotides. Thus, it appears that these nucleotides are still involved in crystal packing. The tail-to-tail stacked quadruplexes are vertically stacked, with the missing residue providing the head-

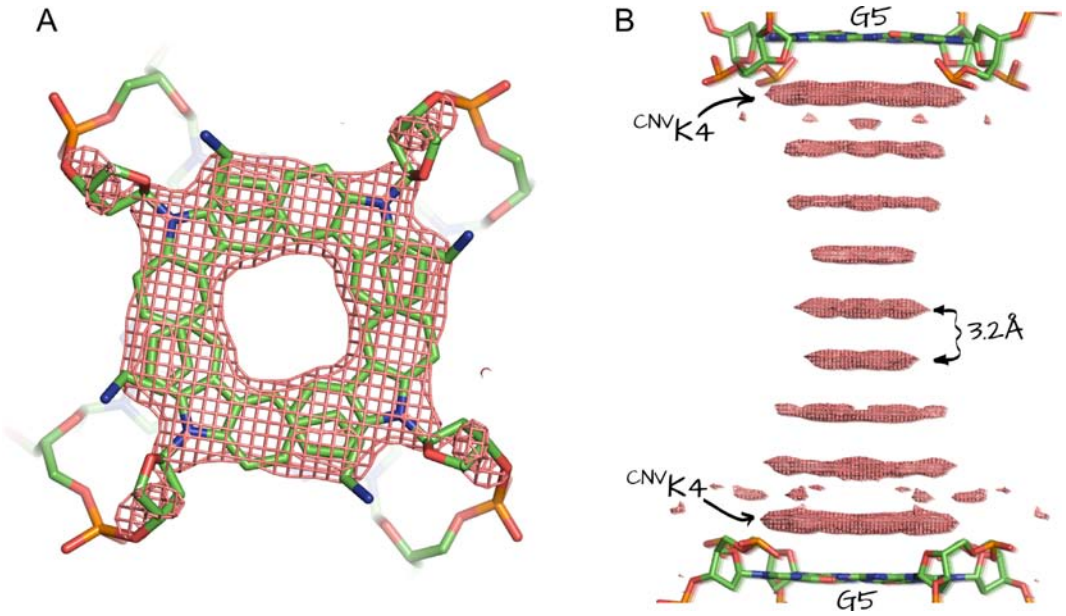

Figure 5. Disordered nucleotides. (**A**) Difference density (mF_o-dF_c) contoured at 2.7 σ (pink) adjacent to G5 after completed refinement of residues 5–10. The ^{CNV}K4 monomer is shown fit in this density. The 4-fold symmetry resulted in significant clashes between ^{CNV}K4 residues as fit, suggesting that the observed difference density may be the result of partial occupancy. (**B**) Complete difference density (3.0 σ contour) between layers of vertically stacked quadruplexes. The planar density is spaced ∼3.2 Å apart and extends between the quadruplex layers. Presumed ^{CNV}K4 density regions corresponding to (A) are indicated.

to-head interactions to keep the different layers of quadruplexes in the crystal aligned. This is reminiscent of order-disorder transitions observed in some protein crystals (44), though in this case the disorder appears to occur within a segment of each DNA strand, as opposed to arising from discrete layers of macromolecules. We reprocessed and re-determined the structure in spacegroup P1, and observed identical regions of disorder and partial difference density. This suggests that the residues that mediate crystal contacts are likely to be in multiple conformations; however, we cannot rule out that the difference density belongs to additional oligonucleotide strands.

Biological implications for non-standard nucleobases in quadruplex formation

One of the remarkable features of this structure is the ability for a nearly self-complementary DNA oligonucleotide to form a multi-strand G-quadruplex. Though the crystal structure does not directly answer how this might occur, there are clearly several factors that can be considered. One intriguing possibility is that the ^{CNV}K monomer favors formation of quadruplex structure. The base pairing in this decamer sequence is only disrupted by the ^{CNV}K and this could lead to duplex instability. Further, because ^{CNV}K is significantly more hydrophobic than a standard nucleobase, it is conceivable that hydrophobic interactions between ^{CNV}K nucleosides from different strands prime this particular sequence to take on a conformation and orientation that is conducive to forming a quadruplex. Though it remains to be directly determined how much of an impact the ^{CNV}K monomer has on quadruplex formation, this raises the interesting possibility that nucleobase modifications, particularly adducts that result in the addition of bulky hydrophobic groups (benzo(a)pyrene, aflatoxin B1, aristolactam, N-hydroxyl-aminofluorene) may locally favor non-duplex DNA structures, such as quadruplexes, through stacking interactions. This is consistent with the prevailing mechanism for how G-quadruplex-stabilizing ligands interact with terminal tetrads (12). Genomic G-quadruplex structures have been implicated as a signal for DNA damage (45–47), but to our knowledge the possibility that DNA damage may be responsible in some cases for inducing G-quadruplex formation has yet to explored.

Expanding the DNA quadruplex sequence space

DNA has proved to be an extraordinary material for the rational design and assembly of nanostructures (48). Though most designs have relied on the predictability of Watson–Crick base pairing, there has been a growing appreciation for the structural diversity afforded by non-canonical DNA motifs (43,49). The predictability and thermal stability of inter and intramolecular G-quadruplexes have allowed them to be used in the construction of DNA structures and devices (50–55). Unfortunately, the predictability of G-quadruplex formation also highlights the major drawback: a lack of sequence diversity (54). Unlike Watson–Crick duplex DNA which guides assembly through complementary base pairing, assembling multiple different G-quadruplex-forming sequences simultaneously may be challenging due to their inherent sequence requirements. The characterization of the C-tetrad described here, as well as other non-G forming tetrads may open the door for improved control over formation of multiple quadruplex-forming structures for biomaterial and nanotechnological applications.

ACCESSION NUMBERS

Structure factors and coordinates have been deposited in the Protein Data Bank (PDBID 4U92).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online

ACKNOWLEDGMENTS

P.S.L. thanks Professor Ned Seeman for use of his DNA synthesizer, Dr Clare Waggoner at CSU Long Beach for MALDI-MS and Professors Niles Pierce and Kenzo Fujimoto for generous gifts of ^{CNV}K phosphoramidite.

FUNDING

California State Polytechnic University; CSUPERB; Research Corporation [Cottrell College Science Award #10484]; St. John's University; Army Research Office [Award #W911NF1110371 to P.S.L.]; NSF CAREER Award [DMR1149665 to P.J.P.]. Funding for open access charge: University of Maryland startup funds.
Conflict of interest statement. None declared.

REFERENCES

1. Murat,P. and Balasubramanian,S. (2014) Existence and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev.*, **25**, 22–29.
2. Davis,J.T. (2004) G-quartets 40 years later: from 5′-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed. Engl.*, **43**, 668–698.
3. Neidle,S. (1999) *Oxford Handbook of Nucleic Acid Structure*. 1 edn . Oxford University Press, Oxford.
4. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
5. Phan,A.T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J.*, **277**, 1107–1117.
6. Blackburn,E.H. (2001) Switching and signaling at the telomere. *Cell*, **106**, 661–673.
7. Schaffitzel,C., Berger,I., Postberg,J., Hanes,J., Lipps,H.J. and Plückthun,A. (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc. Natl Acad. Sci. U.S.A.*, **98**, 8572–8577.
8. Parkinson,G.N., Lee,M.P.H. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.
9. Biffi,G., Tannahill,D., McCafferty,J. and Balasubramanian,S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.*, **5**, 182–186.
10. Henderson,A., Wu,Y., Huang,Y.C., Chavez,E.A., Platt,J., Johnson,F.B., Brosh,R.M., Sen,D. and Lansdorp,P.M. (2014) Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.*, **42**, 860–869.
11. Huppert,J.L. and Balasubramanian,S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.*, **35**, 406–413.
12. Balasubramanian,S. and Neidle,S. (2009) G-quadruplex nucleic acids as therapeutic targets. *Curr. Opin. Chem. Biol.*, **13**, 345–353.
13. Cheong,C. and Moore,P.B. (1992) Solution structure of an unusually stable RNA tetraplex containing G- and U-quartet structures. *Biochemistry*, **31**, 8406–8414.
14. Pan,B., Xiong,Y., Shi,K. and Sundaralingam,M. (2003) Crystal structure of a bulged RNA tetraplex at 1.1 a resolution: implications for a novel binding site in RNA tetraplex. *Struct. Lond. Engl. 1993*, **11**, 1423–1430.
15. Pan,B., Xiong,Y., Shi,K. and Sundaralingam,M. (2003) An eight-stranded helical fragment in RNA crystal structure: implications for tetraplex interaction. *Struct. Lond. Engl. 1993*, **11**, 825–831.
16. Pan,B., Xiong,Y., Shi,K., Deng,J. and Sundaralingam,M. (2003) Crystal structure of an RNA purine-rich tetraplex containing adenine tetrads: implications for specific binding in RNA tetraplexes. *Struct. Lond. Engl. 1993*, **11**, 815–823.
17. Xu,Y., Ishizuka,T., Kimura,T. and Komiyama,M. (2010) A U-tetrad stabilizes human telomeric RNA G-quadruplex structure. *J. Am. Chem. Soc.*, **132**, 7231–7233.
18. Patel,P.K. and Hosur,R.V. (1999) NMR observation of T-tetrads in a parallel stranded DNA quadruplex formed by Saccharomyces cerevisiae telomere repeats. *Nucleic Acids Res.*, **27**, 2457–2464.
19. Cáceres,C., Wright,G., Gouyette,C., Parkinson,G. and Subirana,J.A. (2004) A thymine tetrad in d(TGGGGT) quadruplexes stabilized with Tl+/Na+ ions. *Nucleic Acids Res.*, **32**, 1097–1102.
20. Patel,P.K., Koti,A.S. and Hosur,R.V. (1999) NMR studies on truncated sequences of human telomeric DNA: observation of a novel A-tetrad. *Nucleic Acids Res.*, **27**, 3836–3843.
21. Patel,P.K., Bhavesh,N.S. and Hosur,R.V. (2000) NMR observation of a novel C-tetrad in the structure of the SV40 repeat sequence GGGCGG. *Biochem. Biophys. Res. Commun.*, **270**, 967–971.
22. Patel,P.K., Bhavesh,N.S. and Hosur,R.V. (2000) Cation-dependent conformational switches in d-TGGCGGC containing two triplet repeats of Fragile X Syndrome: NMR observations. *Biochem. Biophys. Res. Commun.*, **278**, 833–838.
23. Gehring,K., Leroy,J.L. and Guéron,M. (1993) A tetrameric DNA structure with protonated cytosine.cytosine base pairs. *Nature*, **363**, 561–565.
24. Leroy,J.L., Gueron,M., Mergny,J.L. and Helene,C. (1994) Intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif. *Nucleic Acids Res.*, **22**, 1600–1606.
25. Salisbury,S.A., Wilson,S.E., Powell,H.R., Kennard,O., Lubini,P., Sheldrick,G.M., Escaja,N., Alazzouzi,E., Grandas,A. and Pedroso,E. (1997) The bi-loop, a new general four-stranded DNA motif. *Proc. Natl Acad. Sci.*, **94**, 5515–5518.
26. Webba da Silva,M. (2003) Association of DNA quadruplexes through G:C:G:C tetrads. Solution structure of d(GCGGTGGAT). *Biochemistry*, **42**, 14 356–14 365.
27. Webba da Silva,M. (2005) Experimental demonstration of T:(G:G:G:G):T hexad and T:A:A:T tetrad alignments within a DNA quadruplex stem. *Biochemistry*, **44**, 3754–3764.
28. Viladoms,J., Escaja,N., Frieden,M., Gómez-Pinto,I., Pedroso,E. and González,C. (2009) Self-association of short DNA loops through minor groove C:G:G:C tetrads. *Nucleic Acids Res.*, **37**, 3264–3275.
29. Pan,B., Shi,K. and Sundaralingam,M. (2006) Base-tetrad swapping results in dimerization of RNA quadruplexes: implications for formation of the i-motif RNA octaplex. *Proc. Natl Acad. Sci. U.S.A.*, **103**, 3130–3134.
30. Mukundan,V.T. and Phan,A.T. (2013) Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences. *J. Am. Chem. Soc.*, **135**, 5017–5028.
31. Yoshimura,Y. and Fujimoto,K. (2008) Ultrafast reversible photo-cross-linking reaction: toward in situ DNA manipulation. *Org. Lett.*, **10**, 3227–3230.
32. Yoshimura,Y., Ohtake,T., Okada,H. and Fujimoto,K. (2009) A new approach for reversible RNA photocrosslinking reaction: application to sequence-specific RNA selection. *Chembiochem*, **10**, 1473–1476.
33. Choi,H.M.T., Chang,J.Y., Trinh,L.A., Padilla,J.E., Fraser,S.E. and Pierce,N.A. (2010) Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat. Biotechnol.*, **28**, 1208–1212.
34. Tagawa,M., Shohda,K., Fujimoto,K. and Suyama,A. (2011) Stabilization of DNA nanostructures by photo-cross-linking. *Soft Matter*, **7**, 10931.
35. Vieregg,J.R., Nelson,H.M., Stoltz,B.M. and Pierce,N.A. (2013) Selective nucleic acid capture with shielded covalent probes. *J. Am. Chem. Soc.*, **135**, 9691–9699.
36. Ma,Z., Lee,R.W., Li,B., Kenney,P., Wang,Y., Erikson,J., Goyal,S. and Lao,K. (2013) Isothermal amplification method for next-generation sequencing. *Proc. Natl Acad. Sci.*, **110**, 14 320–14 323.
37. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. In: Carter,CW Jr. and

Sweet,RM (eds). *Methods in Enzymology*. Academic Press, New York, Vol. **276**, pp. 307–326.
38. Afonine,P.V., Grosse-Kunstleve,R.W., Echols,N., Headd,J.J., Moriarty,N.W., Mustyakimov,M., Terwilliger,T.C., Urzhumtsev,A., Zwart,P.H. and Adams,P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.*, **68**, 352–367.
39. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
40. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.*, **53**, 240–255.
41. Lech,C.J., Heddi,B. and Phan,A.T. (2013) Guanine base stacking in G-quadruplex nucleic acids. *Nucleic Acids Res.*, **41**, 2034–2046.
42. Robinson,H., van der Marel,G.A., van Boom,J.H. and Wang,A.H. (1992) Unusual DNA conformation at low pH revealed by NMR: parallel-stranded DNA duplex with homo base pairs. *Biochemistry*, **31**, 10 510–10 517.
43. Muser,S.E. and Paukstelis,P.J. (2012) Three-dimensional DNA crystals with pH-responsive noncanonical junctions. *J. Am. Chem. Soc.*, **134**, 12 557–12 564.
44. Renko,M., Taler-Verčič,A., Mihelič,M., Žerovnik,E. and Turk,D. (2014) Partial rotational lattice order–disorder in stefin B crystals. *Acta Crystallogr. D Biol. Crystallogr.*, **70**, 1015–1025.
45. D' Adda di Fagagna,F., Reaper,P.M., Clay-Farrace,L., Fiegler,H., Carr,P., Von Zglinicki,T., Saretzki,G., Carter,N.P. and Jackson,S.P. (2003) A DNA damage checkpoint response in telomere-initiated senescence. *Nature*, **426**, 194–198.
46. Tauchi,T., Shin-Ya,K., Sashida,G., Sumi,M., Nakajima,A., Shimamoto,T., Ohyashiki,J.H. and Ohyashiki,K. (2003) Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways. *Oncogene*, **22**, 5338–5347.
47. Rodriguez,R., Müller,S., Yeoman,J.A., Trentesaux,C., Riou,J.F. and Balasubramanian,S. (2008) A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.*, **130**, 15 758–15 759.
48. Seeman,N.C. (2003) DNA in a material world. *Nature*, **421**, 427–431.
49. Paukstelis,P.J., Nowakowski,J., Birktoft,J.J. and Seeman,N.C. (2004) Crystal structure of a continuous three-dimensional DNA lattice. *Chem. Biol.*, **11**, 1119–1126.
50. Pitchiaya,S. and Krishnan,Y. (2006) First blueprint, now bricks: DNA as construction material on the nanoscale. *Chem. Soc. Rev.*, **35**, 1111–1121.
51. Alberti,P., Bourdoncle,A., Saccà,B., Lacroix,L. and Mergny,J.L. (2006) DNA nanomachines and nanostructures involving quadruplexes. *Org. Biomol. Chem.*, **4**, 3383–3391.
52. Ghodke,H., Krishnan,R., Vignesh,K., Kumar,G. V., Narayana,C. and Krishnan,Y. (2007) The I-tetraplex building block: rational design and controlled fabrication of robust 1D DNA scaffolds through non-Watson-Crick Interactions. *Angew. Chem. Int. Ed. Engl.*, **119**, 2700–2703.
53. Modi,S., Nizak,C., Surana,S., Halder,S. and Krishnan,Y. (2013) Two DNA nanomachines map pH changes along intersecting endocytic pathways inside the same cell. *Nat. Nanotechnol.*, **8**, 459–467.
54. Yatsunyk,L.A., Piétrement,O., Albrecht,D., Tran,P.L.T., Renčiuk,D., Sugiyama,H., Arbona,J.M., Aimé,J.P. and Mergny,J.L. (2013) Guided assembly of tetramolecular G-quadruplexes. *ACS Nano*, **7**, 5701–5710.
55. Yatsunyk,L.A., Mendoza,O. and Mergny,J.L. (2014) 'Nano-oddities': unusual nucleic acid assemblies for DNA-based nanostructures and nanodevices. *Acc. Chem. Res.*, **47**, 1836–1844.