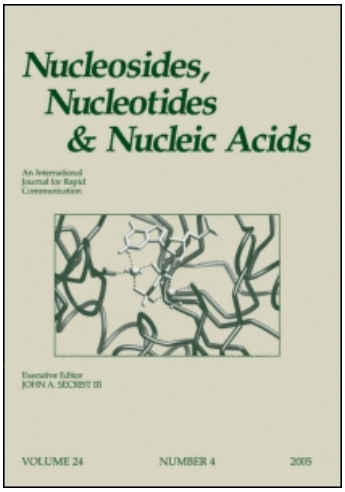

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Crystal Structure of d(gcGXGAgc) with X = G: a Mutation at X is Possible to Occur in a Base-Intercalated Duplex for Multiplex Formation

Jiro Kondo[a]; Tanashaya Ciengshin[a]; Ella Czarina Magat Juan[a]; Yoshiteru Sato[a]; Kenta Mitomi[a]; Satoru Shimizu[a]; Akio Takénaka[a]
[a] Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan

To cite this Article Kondo, Jiro , Ciengshin, Tanashaya , Juan, Ella Czarina Magat , Sato, Yoshiteru , Mitomi, Kenta , Shimizu, Satoru and Takénaka, Akio(2006) 'Crystal Structure of d(gcGXGAgc) with X = G: a Mutation at X is Possible to Occur in a Base-Intercalated Duplex for Multiplex Formation', Nucleosides, Nucleotides and Nucleic Acids, 25: 4, 693 — 704
To link to this Article: DOI: 10.1080/15257770600686543
URL: http://dx.doi.org/10.1080/15257770600686543

Nucleosides, Nucleotides, and Nucleic Acids, 25:693–704, 2006

ISSN: 1525-7770 print / 1532-2335 online
DOI: 10.1080/15257770600686543

Taylor & Francis
Taylor & Francis Group

CRYSTAL STRUCTURE OF d(gcGXGAgc) WITH X=G: A MUTATION AT X IS POSSIBLE TO OCCUR IN A BASE-INTERCALATED DUPLEX FOR MULTIPLEX FORMATION

Jiro Kondo, Tanashaya Ciengshin, Ella Czarina Magat Juan, Yoshiteru Sato, Kenta Mitomi, Satoru Shimizu, and Akio Takénaka □ *Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan*

□ *DNA fragments with the sequences d(gcGX[Y]$_n$Agc) (n = 1, X = A, and Y = A, T, or G) form base-intercalated duplexes, which is a basic unit for formation of multiplexes such as octaplex and hexaplex. To examine the stability of multiplexes, a DNA with X = Y = G and n = 1 was crystallized under conditions different from those of the previously determined sequences, and its crystal structure has been determined. The two strands are coupled in an anti-parallel fashion to form a base-intercalated duplex, in which the first and second residues form Watson-Crick type G:C pairs and the third and sixth residues form a sheared G:A pairs at both ends of the duplex. The G_4 and G_5 bases are stacked alternatively on those of the counter strand to form a long G column of G_3-G_4-G_5*-G_5-G_4*-G_3*, the central four Gs being protruded. In addition, the three duplexes are associated to form a hexaplex around a mixture of calcium and sodium cations on the crystallographic threefold axis. These structural features are similar to those of the previous crystals, though slightly different in detail. The present study indicates that mutation at the 4th position is possible to occur in a base-intercalated duplex for multiplex formations, suggesting that DNA fragments with any sequence sandwiched between the two triplets gcG and Agc can form a multiplex.*

Keywords DNA; Base-intercalated duplex; DNA hexaplex; X-ray structure; Repetitive DNA sequence

Received 9 February 2006; accepted 16 February 2006.

We thank Atsushi Nakagawa for facilities and help during data collection. This work was supported in part by Grants-in-Aid for Scientific Research (No. 12480177 and 14035217) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

This article is dedicated to Professor Eiko Ohtsuka on the occasion of her 70th birthday.

Present address for Jiro Kondo Institut de Biologie Moléculaire et Cellulaire du CNRS, Université Louis Pasteur, 15 rue René Descartes, Strasbourg, France.

Address correspondence to Akio Takénaka, Department of Life Science, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuda, Midori-ku, Yokohama 226-8501, Japan. E-mail: atakenak@bio.titech.ac.jp

INTRODUCTION

We have reported that DNA fragments with the sequence d(gc<u>G</u>X[Y]$_n$<u>A</u>gc)* (X = A; Y = A, T, or G; and n = 1) form a base-intercalated duplex[1–4] that is a basic unit for further formation of several types of multiplexes. Such multiplex formations suggest that similar repetitive sequences found in human genome may have some biological significance.[4] At both ends of the duplex, two Watson-Crick G:C base pairs are formed. The critical point of the duplex, however, is that the subsequent G_3 residue forms a base pair with the A_6 residue of the counter strand, in which the two hydrogen bonds are formed through N3(G)...H-N6(A) and N2-H(G)...N7(A). This pairing occurs between a Sugar-edge and a Hoogsteen-edge, typical of the sheared type (*trans* Sugar-edge/Hoogsteen type) with shorter C1′...C1′ distance. As a result, the central parts of the two strands are forced to come together closer, and the A_4 and X_5 residues are not involved in any base pair interactions. Instead those residues form an intercalated A_4-X_5*-X_5-A_4* (*indicates residues in the counter strand) base-base stacked column. In addition, the central column is protruded from the duplex so that three or four duplexes can be associated to form a multiplex, in which several cations and anions facilitate the multiplex formation. It seems that the fourth A residues are just stacked between the upper and lower residues with no specific roles in the structures of d(gc<u>G</u>A[Y]$_1$<u>A</u>gc) (Y = A, T or G), and that it can be replaced with any other base. If so, it is expected that base-intercalated duplexes of repetative sequences would have a variety of biological significance. To examine the versatility of the fourth residue, the crystal structure of d(gc<u>G</u>G[G]$_1$<u>A</u>gc) (hereafter GGA-Br), which is mutated at the 4th and/or 5th positions of the previous sequences[1–4] has been determined.

MATERIALS AND METHODS

Preparation

GGA-Br was synthesized on a DNA synthesizer, the second residue being modified with 2′-deoxy-5-bromocytidine for phasing in X-ray analysis. This octamer was purified by HPLC and gel filtration. Crystals were obtained at 293 K by the hanging-drop vapor diffusion method in droplets mixed 1 μl of 1.0 mM DNA and 1 μl of the reservoir solution containing 50 mM sodium cacodylate (pH 7.0), 10 mM spermine tetrahydrochloride, 20 mM calcium chloride, 100 mM sodium chloride, and 10% (v/v) 2-methyl-2,4-pentanediol. Crystals were smaller than those that are ordinarily used for X-ray diffraction studies. They were mounted in nylon cryoloops (Hampton Research, CA)

*Small and underlined characters indicate that they can form Watson-Crick G:C and sheared G:A base pairs, respectively, when the fragments are aligned in an anti-parallel fashion for self-assembly or folding.

with solutions a solution containing 35% (v/v) 2-methyl-2,4-pentanediol as a cryoprotectant and stored in liquid nitrogen prior to X-ray experiments.

Data Collection

X-Ray data were collected at 100 K with synchrotron radiation ($\lambda = 0.900$ Å) at BL44XU of SPring-8 (Harima, Japan), using a DIP2040 detector (Mac Science, Japan) set 500 mm from the crystal. Diffraction patterns were processed in a resolution range from 32 to 3.6 Å with the program *MOSFLM*.[5] Intensity data were then put on a relative scale and merged into the independent reciprocal space using the program *SCALA* in *CCP4* suite.[6] The crystal belongs to the space group $P6_3$ with unit cell dimensions of $a = b = 36.4$ Å and $c = 62.8$ Å. There were 525 unique reflections with 93.6% completeness and R_{merge} of 11.2%.

Structure Determination and Refinement

The unit cell dimensions are very close to those of d(gcGA[Y]$_1$Agc) (Y = A, T, or G), and the space group is similar but different ($P6_3$ and $P6_322$).[1,2] Therefore, molecular replacement (program *AMoRe*[7]) with the crystal structure of d(gcGA[A]$_1$Agc)[2] gave a unique solution with $Cc = 81.3\%$ and R-factor = 32.4% in a resolution range from 15 to 4.5 Å, from which initial phases of diffraction data were successively derived. The initial electron density map was improved by density modification followed by solvent flattening (solvent content 43.8%) with the program *CNS*.[8] The atomic parameters of the structure were refined with the program *CNS* through a combination of rigid-body refinement, simulated-annealing, crystallographic conjugate gradient minimization refinements, and B-factor refinements. As the resolution of the data was not so high, structural restraints were applied on all base-base interactions throughout refinements. Furthermore, a non-crystallographic twofold symmetry restraint was applied between the two crystallographically independent DNA strands of the duplex. After several cycles of structure refinements, five strong densities were observed on the crystallographic threefold axis. Although it was difficult to identify the types of atoms due to the low-resolution data, calcium cations were assigned by taking specific features of ions into account, and these calcium ions were included in the final refinement (details are discussed below).

The final R-factor was 29.4% for 10–3.6 Å resolution data ($R_{\text{free}} = 32.9$ for 15.0% of the observed data[9]). All local helical parameters including torsion angles and pseudorotation phase angles of ribose rings were calculated using the program *3DNA*.[10] The statistics of structure refinement are summarized in Table 1. A $2|F_o| - |F_c|$ map, shown in Figure 1, was drawn with the program *PyMOL* (DeLano Scientific LLC, CA). Other figures were drawn with the program *RASMOL*.[11]

TABLE 1 Crystal Data, Statistics of Data Collection and Statistics of Structure Refinement

Crystal data	
Space group	$P6_3$
Unit cell (Å)	$a = b = 36.4$, $c = 62.8$
Z^a	1
Data collection	
Beamline	BL44XU of Spring-8
Wavelength (Å)	0.900
Resolution (Å)	32-3.6
Observed reflections	33,760
Unique reflections	525
Completeness (%)	93.6
in the outer shell (%)	95.3
R_{merge} (%)^b	11.2
in the outer shell (%)	29.3
Structure refinement	
Resolution range (Å)	10-3.6 ($F_o > 3\sigma$)
Used reflections	458
R-factor (%)^c	29.4
R_{free}^{d}	32.9
Number of DNA atoms	334
Number of ions	$5Na^+$, $2Ca^{2+}$
Number of water molecules	2
R.m.s. deviation	
Bond lengths (Å)	0.009
Bond angles (°)	1.2
Improper angles (°)	1.0

[a] Number of dsDNA in the asymmetric unit.

[b] $R_{merge} = 100 \times \Sigma_{hklj} |I_{hklj} - \langle I_{hklj} \rangle| / \Sigma_{hklj} \langle I_{hklj} \rangle$.

[c] R-factor $= 100 \times \Sigma \| F_o | - | F_c \| / \Sigma | F_o |$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively.

[d] Calculated using a random set containing 10% of observations that were not included throughout refinement.[9]

RESULTS AND DISCUSSION

Structure of the Base-Intercalated Duplex of GGA-Br

As shown in Figure 1, an overall $2|F_o| - |F_c|$ map well defines the structure of GGA-Br even though the data was obtained at low resolution. The crystal data shows isomorphism with the crystals of d(gcGA[Y]$_1$Agc) (Y = A, T, or G) obtained under conditions containing hexammine cobalt chloride.[1,2] As the present crystallization condition did not contain hexammine cobalt chloride, the space group is slightly different ($P6_3$ for the present crystal and $P6_322$ for the previous one). However, as expected, GGA-Br also forms base-interculated duplexes similar to those of d(gcGA[Y]$_1$Agc) (Y = A, T, or G).[1–4] Schematic diagram and a stereo-view of GGA-Br are shown in Figure 2. Two DNA octamers related by a non-crystallographic two-fold symmetry are aligned in an anti-parallel fashion are associated to each other to form a

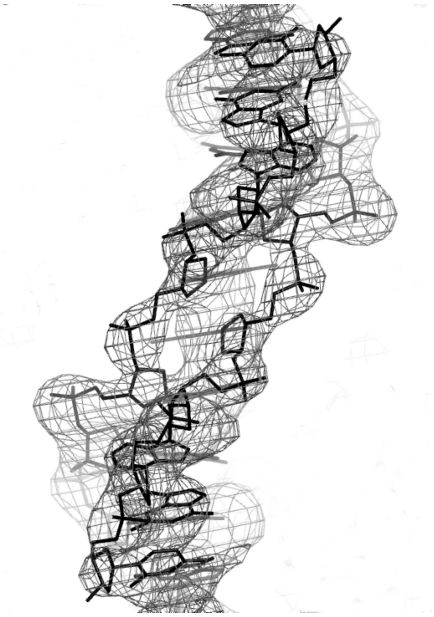

FIGURE 1 An overall $2|F_o| - |F_c|$ map of the base-intercalated duplex of GGA-Br. Density is contoured at 1σ level. Formations of sheared G_3:A_6* and A_6:G_3* base pairs make two phosphate backbones closer in the central region of the duplex.

base-intercalated duplex. As shown schematically in Figure 2 (a), the duplex consists of two parts, the stem and the central regions. In the stem region of the duplex, two Watson-Crick base pairs, G_1:C_8* and C_2:G_7*, are followed by a sheared G_3:A_6* base pair (see Figure 3). On the other hand, no base

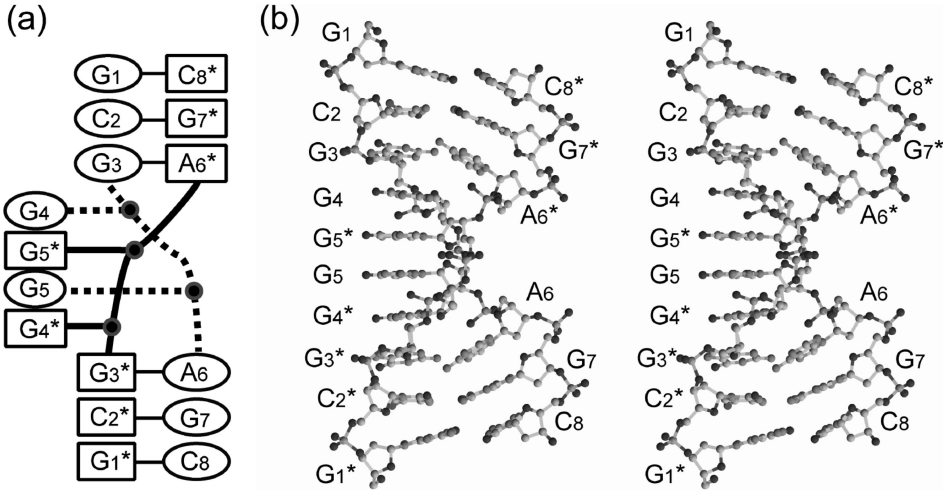

FIGURE 2 Schematic diagram (a) and a stereo-view (b) of the base-intercalated duplex of GGA-Br. * Indicates a residue in the counter strand.

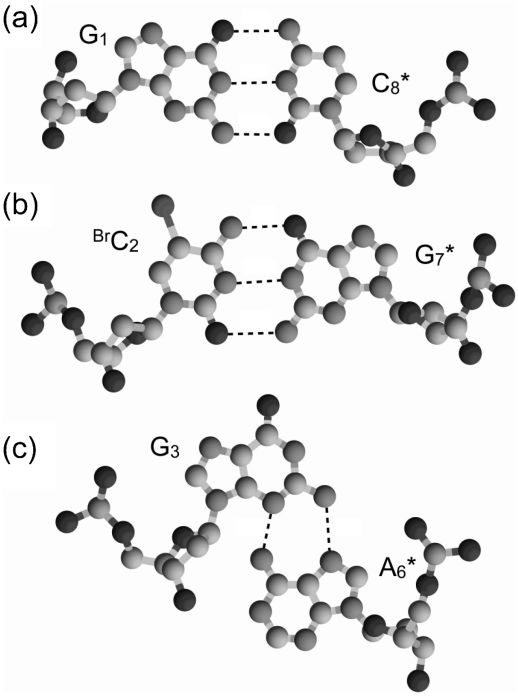

FIGURE 3 Base pairs found in the stem region of the base-intercalated duplex of GGA-Br. Base pairs for one end of the duplex are shown due to the non-crystallographic twofold symmetry. Broken lines indicate possible hydrogen bonds. * Indicates a residue in the counter strand.

pair formations occur in the central region of the duplex. Six Gs make a long poly-guanine stacked column G_3-G_4-G_5*-G_5-G_4*-G_3*, and expose their Watson-Crick and Hoogsteen edges into the solvent region. These structural features are quite similar to those of the previous sequences.[1–4]

Significance of Sheared G_3:A_6* and A_6:G_3* Base Pairs

The G_3 residue forms a pair with the A_6 residue of the counter strand in non-Watson-Crick geometry, through hydrogen bonds between N2(G_3) and N7(A_6*) and between N3(G_3) and N6(A_6*), as shown in Figure 3 (*c*). This type of pairing has been found in some functional RNA molecules, such as hammerhead ribozymes.[12–14] The C1′-C1′ distances of these pairs (8.2–8.3 Å) are shorter than that (10.7 Å) of the standard Watson-Crick pairs. The two bases are not coplanar and their propeller twist angles become larger (22.5–22.7°). However, the two residues adopt a conformation close to the C2′-*endo* pucker. In consequence, it can be concluded that the sheared G:A base pair formations in the two stems are essential to make the two phosphate backbones closer and to stabilize the central G_4-G_5*-G_5-G_4* intercalated stacking.

Residues in the Central Region

Similar to those in the base-intercalated duplexes of d(gcGA[Y]$_1$Agc) (Y = A, T, or G),[1–4] the central 4th and 5th residues, G_4 and G_5, are not involved in any base-pair formations. Their base moieties are stacked on those of the other strand, so that G_4 is intercalated between G_3 and G_5*, and G_5 is intercalated between G_5* and G_4* (see Figure 2). The G_4 residue adopts C4′-*exo* (close to C3′-*endo*) pucker, which is necessary to make an open space for intercalation of the G_5* base between the G_4 and G_5 residues. As discussed above, the G_3:A_6^* pairs serve as scaffolds, which support the subsequent G_4-G_5*-G_5-G_4* intercalated stacking. The residues in the polyguanine stacked G_3-G_4-G_5*-G_5-G_4*-G_3* column expose their Watson-Crick and Hoogsteen edges into the solvent region to interact with the surrounding molecules. These findings confirm that the 4th residue can be replaced with any other bases in the formation of the base-intercalated duplex.

Hexaplex Formation

The GGA-Br crystal, which is obtained under a condition containing sodium and calcium ions, has the crystallographic threefold symmetry along the *c*-axis, around which three base-intercalated duplexes are assembled to form a DNA hexaplex (see Figures 4 and 5). The overall structure of this hexaplex is very similar to those of d(gcGA[Y]$_1$Agc) (Y = A, T, or G) obtained under conditions containing hexammine cobalt chloride.[1,2] No direct hydrogen-bonding contacts are observed between the base-intercalated duplexes.

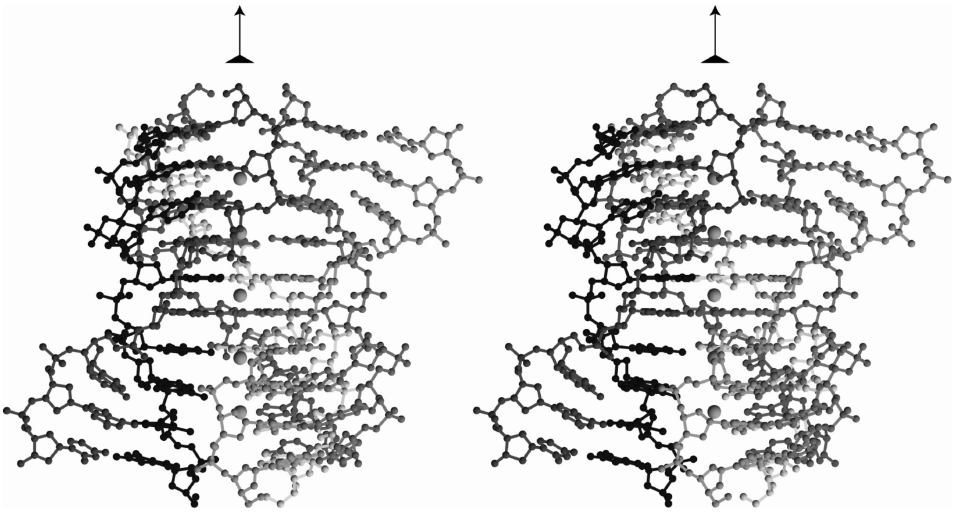

FIGURE 4 A stereo-view of the hexaplex composed of the three base-intercalated duplexes found in the GGA-Br crystals. Five ions are observed on the threefold axis.

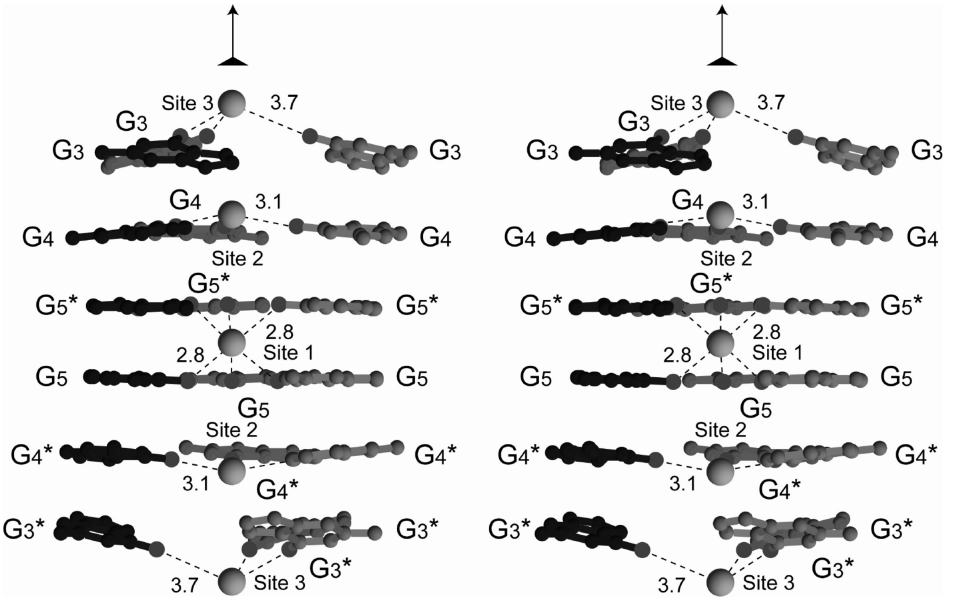

FIGURE 5 A stereo-view of the central region of the GAA-Br hexaplex. Three base-intercalated duplexes are assembled around the crystallographic threefold axis. Values indicate atomic distances (Å) from ions to O6 atoms of guanine residues.

Role of Ions

Due to low diffraction resolution, only a few solvent densities were found. As mentioned in materials and methods, five strong densities are found on the crystallographic threefold axis (see Figure 5). Although it is difficult to identify the types of atoms, it is possible to postulate their nature by taking the crystallization condition into account. In crystallization, sodium chloride and calcium chloride were used as cement. The space around the crystallographic three-fold axis is surrounded by the negatively charged O6 atoms of G_3, G_4, G_5*, G_5, G_4*, and G_3* residues. Therefore it is plausible to assume that a mixture of calcium and sodium cations can occupy the long negative tunnel.

At the center of the negative tunnel in the hexaplex (see Figure 6), six O6 atoms of G_5 residues surround the site 1. The distance (2.8 Å) between the site 1 and the O6 atoms is slightly longer than those of normal Na—O and Ca—O bonds (2.29–2.50 Å and 2.33–2.46 Å, respectively).[15] This sequence is similar to that of the hexaplex of the previous sequence d(gcGA[G]$_1$Agc), in which a sodium cation occupies the site.[1] Likewise, the center of the present sequence must be occupied by a sodium cation.

The major difference is observed at the 4th residue mutated for the present investigation. In the previous d(gcGA[G]$_1$Agc) hexaplex, the central point of the three A_4 residues (corresponding to site 2) is a positive space

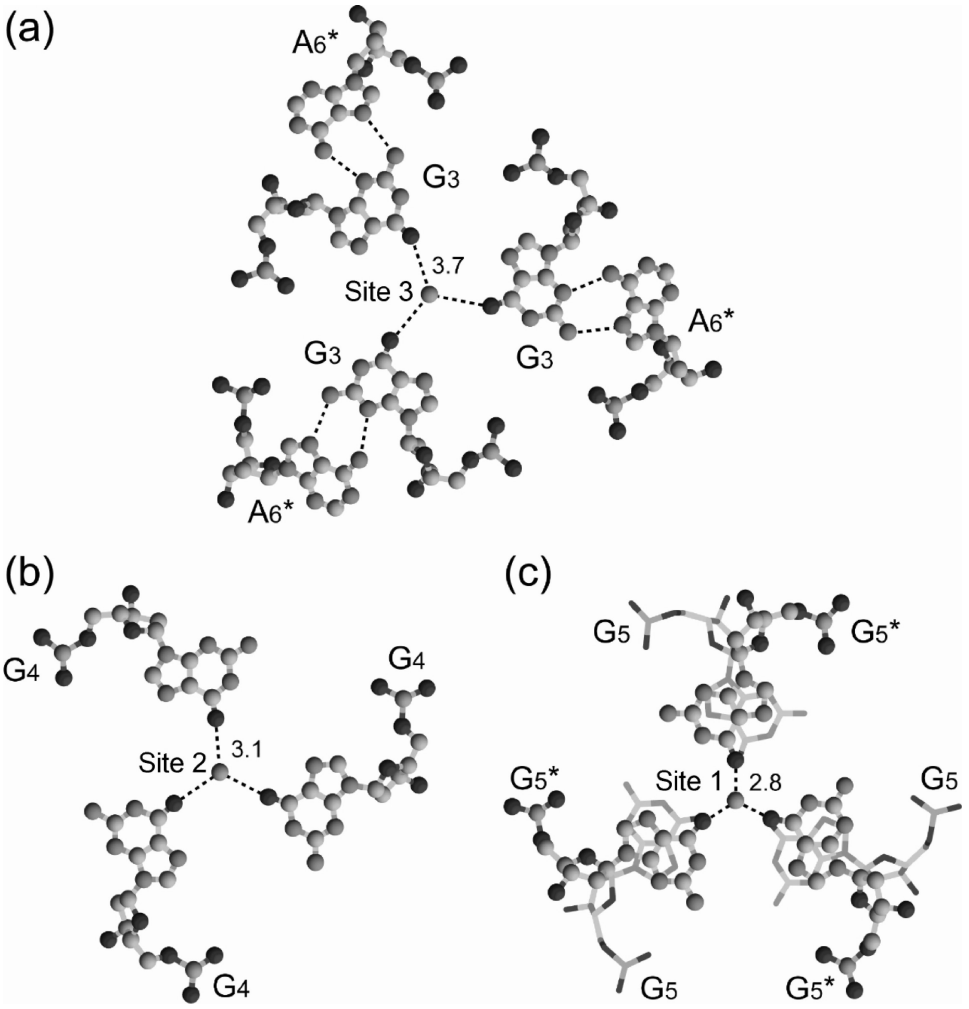

FIGURE 6 Assemblies of the three G_3 (a), G_4 (b), and G_5 (c) residues around the crystallographic threefold axis in the hexaplex of GGA-Br. In (c), the site 1 is just a mid-point between the stacked G_5-triplets. Values indicate atomic distances (Å) from ions to O6 atoms of guanine residues.

surrounded by the amino groups, at which a chloride ion is reasonably accommodated to stabilize the hexaplex.[1] In contrast, the present mutant GGA-Br has G residues that make this space negative. Therefore, it is reasonable to consider that the site 2 is occupied by a mixture of calcium and sodium cations.

Another difference is found at the site 3. In the previous d(gcGA[G]$_1$Agc) hexaplex, Hoogsteen edges of the three G_3 residues, which form G:A sheared pairs, are facing each other around the three fold-axis, so that three N7 atoms and three O6 atoms make a negative space with an octahedral shape. At this site, a hexammine cobalt cation fits well into the space through

hydrogen-bonds between six ammonias bound (coordinated) to the cobalt atom and the O6 and N7 atoms.[1] In GGA-Br, however, the crystallization condition does not contain hexammine cobalt cation. Instead, a mixture of calcium and sodium cations may occupy the site 3. The long distance between the center and O6 atoms (3.7 Å) suggests that the cation is hydrated. Since the hexaplex has a non-crystallographic twofold symmetry, there are two sites with the same features.

The present mutant also has a high potential to form an octaplex with *I*-motif of G-quartet under low potassium concentration, similar to that observed in the previous sequence d(gcGA[G]$_1$Agc),[4] because the G_5 residue is essential for octaplex formation is conserved. It is known that quadruplex formation with G-quartet is stabilized by sodium, potassium and calcium cations.[16,17] Despite the fact that calcium and sodium cations were used for crystallization of the present mutant, the multiplex is hexa-assembly. Therefore, potassium cation is essential for octaplex fomation, but a mixture of calcium and sodium cations may disturb this process.

Biological Significance of the Sequence d(gcGX[Y]$_n$Agc)

From analyses of human genome sequence,[18,19] it was revealed that a large amount of repetitive sequences (transposon-derived sequences, simple sequence repeats (SSRs), segmental duplications, and repetitive sequences in heterochromatin) are dispersed throughout the human genome, whereas the coding sequences account only 1.2% of the euchromatin. It can be expected that such repetitive DNA sequences are in single-stranded states at a moment during dynamic biological processes such as replication, transcription, recombination and cell development. They may form complicated three-dimensional structures for specific biological functions. For example, the G-rich DNA repeats found in some SSRs and in the telomere are able to form quadruplexes with G-quartet.[20] We reported that the DNA sequence d(gcGX[Y]$_n$Agc) (X = A or G; Y = A, T, or G; n = 1) forms a base-intercalated duplex. These structures suggest that the sheared G:A base pairs that flank X[Y]$_n$ inserts facilitate the formation of the base-intercalated duplexes. Such sequences are dispersed specially throughout the repetitive sequences in eukaryotic genome. Examples are d(ccGA[G]$_4$Agg), d(gGA[G]$_2$Ac) and d(tgGA[G]$_3$Aca) sequences in *Homo sapiens* variable number of tandem repeats[21–23] and the *Drosophila* centromeric sequence d(GTACGG[G]$_1$ACCGA)$_n$.[24] An NMR study suggested that the sequence may form a stable fold-back structure,[25] which is very similar to the base-intercalated duplex found in the present GGA-Br crystal. These structural features suggest that a large amount of repetitive sequences in genome, which have been considered as "junk," would have specific functions supported by the specific three-dimensional structures.

The present study shows that mutation at the 4th position is possible to occur in a base-intercalated duplex for multiplex formations. This fact suggests that DNA fragments with any sequence sandwiched between the two triplets gcG and Agc can form a multiplex. Investigations on repeats in human genome having such sequences may lead to discovery of new biological functions.

Database Deposition

The structure has been deposited in the Protein Data Bank (accession code 2FZA).

REFERENCES

1. Kondo, J.; Umeda, U.; Fujita, K.; Sunami, T.; Takénaka, A. X-Ray analyses of d(GCGAXAGC) containing G and T at X: the base-intercalated duplex is still stable even in point mutants at the fifth residue. Journal of Synchrotron Radiation **2004**, 11, 117–120.
2. Sunami, T.; Kondo, J.; Hirao, I.; Watanabe, K.; Miura, K.; Takénaka, A. Structure of d(GCGAAAGC) (hexagonal form): a base-intercalated duplex as a stable structure. Acta Crystallographica **2004**, D60, 90–96.
3. Sunami, T.; Kondo, J.; Hirao, I.; Watanabe, K.; Miura, K.; Takénaka, A. Structures of d(GCGAAGC) and d(GCGAAAGC) (tetragonal form): a switching of partners of the sheared G.A pairs to form a functional G.AxA.G crossing. Acta Crystallographica **2004**, D60, 422–431.
4. Kondo, J.; Adachi, W.; Umeda, S.; Sunami, T.; Takénaka, A. Crystal structures of a DNA octaplex with I-motif of G-quartets and its splitting into two quadruplexes, suggesting a folding mechanism of eight tandem repeats. Nucleic Acids Research **2004**, 32, 2541–2549.
5. Leslie, A.G.W. Joint CCP4 + ESF-EAMCB, *Newsletter on Protein Crystallography No. 26*, 1992.
6. Collaborative Computational Project, Number 4. The CCP4 Suite: Programs for Protein Crystallography. Acta Crystallographica **1994**, D50, 760–763.
7. Navaza, J. AMoRe: an automated package for molecular replacement. Acta Crystallographica **1994**, A50, 157–163.
8. Brünger, A.T.; Adams, P.D.; Clore, G.M.; DeLano, W.L.; Gros, P.; Grosse-Kunstleve, R.W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N.S.; Read, R.J.; Rice, L.M.; Simonson T.; Warren G.L. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallographica **1998**, D54, 905–921.
9. Brünger, A.T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structure. Nature **1992**, 355, 472–475.
10. Lu, X.J.; Olson, W.K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucleic Acids Research **2003**, 31, 5108–5121.
11. Sayle, R.A.; Milner-White, E.J. RASMOL: biomolecular graphics for all. Trends in Biochemical Sciences **1995**, 20, 374.
12. Pley, H.W.; Flaherty, K.M.; McKay, D.B. Three-dimensional structure of a hammerhead ribozyme. Nature **1994**, 372, 68–74.
13. Scott, W.G.; Finch, J.T.; Klug, A. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. Cell **1995**, 81, 991–1002.
14. Doudna, J.A. Hammerhead ribozyme structure: U-turn for RNA structural biology. Structure **1995**, 3, 747–750.
15. Marcus, Y. Ionic radii in aqueous solutions. Chemical Reviews **1988**, 88, 1475–1498.
16. Pilch, D.S.; Plum, G.E.; Breslauer, K.J. The thermodynamics of DNA structures that contain lesions or guanine tetrads. Current Opinion in Structural Biology **1995**, 5, 334–342.
17. Hardin, C.C.; Watson, T.; Corregan, M.; Bailey, C. Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG). Biochemistry **1992**, 31, 833–841.

18. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature **2004**, 431, 931–945.
19. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature **2001**, 409, 860–921.
20. Patel, D.J.; Bouaziz, A.; Kettani, A.; Wang, Y. Structures of guanine-rich and cytosine-rich quadruplexes formed in vitro by telomeric, centromeric, and triplet repeat disease DNA sequences in *Nucleic Acid Structure*, ed. S. Neidle, 389–453, Oxford University Press, Oxford, UK, 1999.
21. Inglehearn, C.F.; Cooke, H.J. A VNTR immediately adjacent to the human pseudoautosomal telomere. Nucleic Acids Research **1990**, 18, 471–476.
22. Kasai, K.; Nakamura, Y.; White, R. Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science. Journal of Forensic Sciences **1990**, 35, 1196–1200.
23. Armour, J.A.; Wong, Z.; Wilson, V.; Royle N.J.; Jeffreys, A.J. Sequences flanking the repeat arrays of human minisatellites: association with tandem and dispersed repeat elements. Nucleic Acids Research **1989**, 17, 4925–4935.
24. Abad, J.P.; Carmena, M.; Baars, S.; Saunders, R.D.; Glover, D.M.; Ludeña, P.; Sentis, C.; Tyler-Smith, C.; Villasante, A. Dodeca satellite: a conserved G+C-rich satellite from the centromeric heterochromatin of *Drosophila melanogaster*. Proceedings of the National Academy of Sciences USA **1992**, 89, 4663–4667.
25. Chou, S.H.; Chin, K.H. Quadruple intercalated G-6 stack: a possible motif in the fold-back structure of the *Drosophila* centromeric dodeca-satellite? Journal of Molecular Biology **2001**, 314, 139–152.