Self-association of a DNA loop creates a quadruplex: crystal structure of d(GCATGCT) at 1.8 Å resolution

Gordon A Leonard[1], Shude Zhang[2], Mark R Peterson[1], Stephen J Harrop[1], John R Helliwell[1], William BT Cruse[2], Beatrice Langlois d'Estaintot[2†], Olga Kennard[2], Tom Brown[3] and William N Hunter[1]*

[1]Department of Chemistry, University of Manchester, Oxford Rd, Manchester M13 9PL, UK, [2]University Chemical Laboratory, Lensfield Rd, Cambridge, CB2 1EW, UK and [3]Department of Chemistry, University of Edinburgh, King's Buildings, West Mains Rd, Edinburgh EH9 3JJ, UK

Background: The flexibility of DNA enables it to adopt three interconvertible types of duplex termed the A-, B- and Z-forms. It can also produce hairpin loops, triplex structures and guanine-rich quadruplex structures. Conformational flexibility assists in the tight packaging of DNA, for example in chromosomes. This is important given the large quantity of genetic information that must be packaged efficiently. Moreover, the ability of DNA to specifically self-associate or interact with complementary sequences is fundamental to many biological processes. Structural studies provide information about DNA conformation and DNA–DNA interactions and suggest features that might be relevant to how the molecule performs its biological role.

Results: We have characterized the structure of a synthetic heptanucleotide that folds into a novel loop structure. The loop is stabilized by association with a cation, by intra-strand hydrogen bonds between guanine and cytosine that are distinct from the normal Watson–Crick hydrogen bonds, and by van der Waals interactions. Two loops associate through the formation of four G·C pairs that exhibit pronounced base-stacking interactions. The formation of a symmetric A·A base pair further stabilizes loop dimerization. Stacking of the A·A pair on a symmetry-related A·A pairing assists the formation of a four-stranded assembly. A T·T pairing is also observed between symmetry-related loops.

Conclusions: This analysis provides a rare example of an experimentally determined non-duplex DNA structure. It provides conformational detail relevant to the tight packaging or folding of a DNA strand and illustrates how a cation might modulate phosphate–phosphate repulsion in a tightly packed structure. The observation of base quartets involving G·C base pairs suggests a further structure to be considered in DNA–DNA interactions. The structure also provides detailed geometries for A·A and T·T base pairs.

Introduction

The recognition between nucleic acids in duplex DNA is determined by a precise hydrogen-bonded pairing of the bases. This principle ensures the unique hybridization of complementary strands of a duplex and contributes to the fidelity of replication. Alternative uses of the hydrogen-bonding capacity of polynucleotides offer an additional dimension to the modes of interaction between these molecules. For example, the hydrogen-bonding sites in the major groove of duplex DNA can allow a third strand to associate and so form a triplex, whereas hypoxanthine and guanine-rich polynucleotides are able to form quadruplex structures. The biological roles of such assemblies have been investigated and the importance of quadruplex structures is becoming apparent. Multiple repeats of guanines found in telomeres may assist the chromosomes, which they terminate, to join together during cell division. The looping of the individual chains can bring together guanine tetrads linked by inter-strand and intra-strand hydrogen bonds [1–7]. Numerous biophysical studies on guanine quartets and tetraplexes have resulted in an extensive literature on this subject. Indeed, the study of G-rich structures has come to dominate the field. Recent developments in this area are discussed by Wang and Patel [8].

It has, however, been suggested that similar quadruplex structures, where the composition of the base tetramers is not exclusively guanine, may also be important in processes where homologous DNA recognition is required [2,9–11]. In this context, we now report the crystal structure of the synthetic oligonucleotide d(GCATGCT). This sequence was designed to enable the characterization of dinucleotide steps in duplex DNA not previously studied by crystallography. To our surprise, we observe a loop that dimerizes to produce a novel DNA assembly. This structure presents a rare example of non-duplex DNA tertiary structure determined experimentally. Full conformational details are given.

Results and discussion

Structure solution and refinement

The DNA heptamer was synthesized, purified and crystallized as described in the Materials and methods section. The crystals are orthorhombic with space group

*Corresponding author. †Present address: Laboratoire de Cristallographie, Université de Bordeaux I, 351 cours de la Libération, 33405 Talence Cedex, France.

C222, unit cell dimensions a=22.52 Å, b=59.37 Å, c=24.35 Å. The structure was solved with the use of a single brominated derivative incorporating anomalous dispersion information derived by tuning the X-ray wavelength to the absorption edge of bromine. Refinement utilized simulated annealing methods in X-PLOR [12] and restrained least-squares techniques [13,14]. The final model consists of 140 non-hydrogen DNA atoms, 2 magnesium ions and 49 water molecules (each treated as an oxygen atom). Refinement converged at an R-factor of 21.4% for all data in the resolution range 7.0–1.8 Å (R-factor=$\Sigma |F_o-F_c| / |F_o|$, where F_o and F_c represent observed and calculated structure factors respectively). This represents 96% of the theoretical data. Residues are labelled G1 to T7 in the 5'$\rightarrow$3' direction, cations are Mg1 and Mg2, water molecules are numbered 1–49 with the label HOH. The average thermal parameters are as follows: 13.2 Å^2 for the single DNA strand; 16.3 Å^2 for phosphates; 12.9 Å^2 for furanose atoms; and 10.4 Å^2 for bases. Root mean square (rms) deviations from ideality for bond lengths are 0.013 Å and 0.022 Å for the bases and the sugar-phosphate backbone respectively. For those distances defining bond angles, the rms deviations from ideality are 0.018 Å and 0.020 Å. The rms deviation from planarity of the nucleotide bases is 0.015 Å, and for chiral volumes (which were not restrained) is 0.285 Å^3. Examples of the electron density are shown in Fig. 1.

A novel loop structure

The asymmetric unit comprises a single strand of d(GCATGCT) that folds back on itself to form a loop structure not previously observed for DNA (Fig. 2). This arrangement is not, as is generally proposed for hairpin loops, stabilized by intra-strand (Watson–Crick) base pairing but by two examples of hydrogen bonds between guanine N2 and cytosine O2 atoms. The stem of the loop thus consists of the bases in the two GpC steps and the 3'-terminal thymine that becomes an overhanging base. The two 'legs' of the stem are joined by a bend consisting of the bases in the ApT segment. All of the bases in the loop, with the exception of the terminal thymine, are in the anti(–) configuration. The hydrogen-bond donor and acceptor groups that are used in G·C Watson–Crick

base-pair formation are positioned side-by-side rather than head-to-head. They are thus available to participate in 12 intermolecular hydrogen bonds. The loop structure brings the phosphates of C2 and G6 close together but phosphate–phosphate repulsive effects are counteracted by the presence of a cation (see below).

The recent observation of an 11 nucleotide single-stranded DNA structure from canine parvovirus [15] is relevant to this discussion. In canine parvovirus, a very different DNA structure is observed, highlighting the diversity that can be accommodated by flexible nucleic acids. Although distinct, some structural themes remain conserved and comparisons will be made at appropriate points in the text that follows.

### Conformation of the phosphate-furanose backbone

Table 1 lists the torsion angles associated with the phosphate-furanose backbone. There are no unusual values for torsion angles α and β, which are gauche(–) and trans respectively, nor for γ, which, with the exception of G1, is gauche(+). Furanose conformations are all close to C2'-endo conformation. This is explicit for A3, T4 and G5. G1 displays a C3'-exo conformation, whereas C2, C6 and T7 are C1'-exo. There is a bimodal distribution of torsion angle δ depending on whether a purine or a pyrimidine is involved; the average δ is 151° for purines and 122° for pyrimidines. The furanose conformation in the heptamer is C2'-endo and, given the steric influence of the 2'-hydroxyl substituent in ribonucleic acid, it is unlikely that an RNA strand would adopt the exact loop structure determined in this study.

The anticorrelation principle (discussed by Saenger, [16]) linking torsion angles δ and χ, which is observed in many double-helical systems, does not apply to this loop structure. Excluding the terminal thymine, the average χ for purines and pyrimidines is 114°. This may simply be a consequence of the different environment in which each nucleotide is found. The trans conformations of torsion angles ϵ (at nucleotides G1, C2, A3, G5 and C6) and ζ (at C2, G5 and C6) are symptomatic of a non-helical structure [16]. The arrangement whereby the A3 furanose is positioned to stack on the base G5 to stabilize the turn is
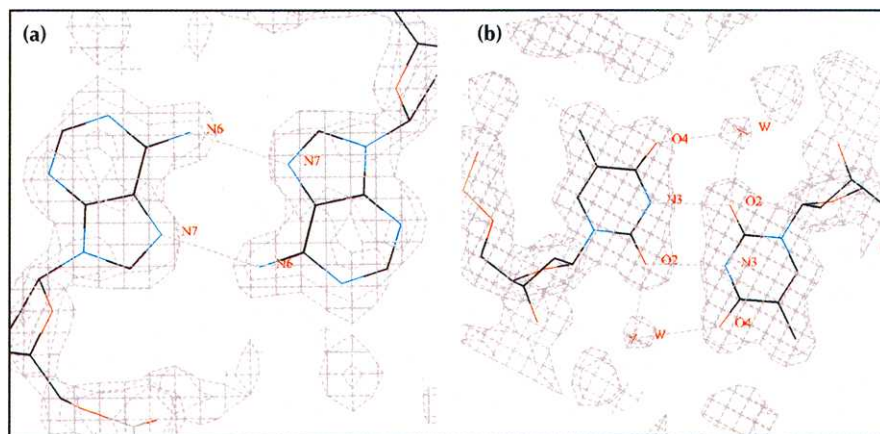
Fig. 1. **(a)** The inter-strand adenine-adenine base pair and associated electron density. The map was calculated with coefficients $2F_o-F_c$, α_{calc} and is represented at the 1.2σ level, depicted as lilac-coloured chicken wire. **(b)** The intermolecular thymine-thymine base pair in the electron-density map. In each diagram, the DNA is shown as a stick model, red indicates oxygen positions, cyan nitrogen and black carbon. W indicates a water position. (Figure drawn using 'O' [32].)

Fig. 2. Ball-and-stick stereo diagram of the loop showing the bound cation. Atom colours are as follows: carbon, white (except for the methyl group of thymines T4 and T7 which are grey); nitrogen, cyan; oxygen, red; phosphorus, yellow; magnesium, magenta. (Figure produced with MOLSCRIPT [33].)
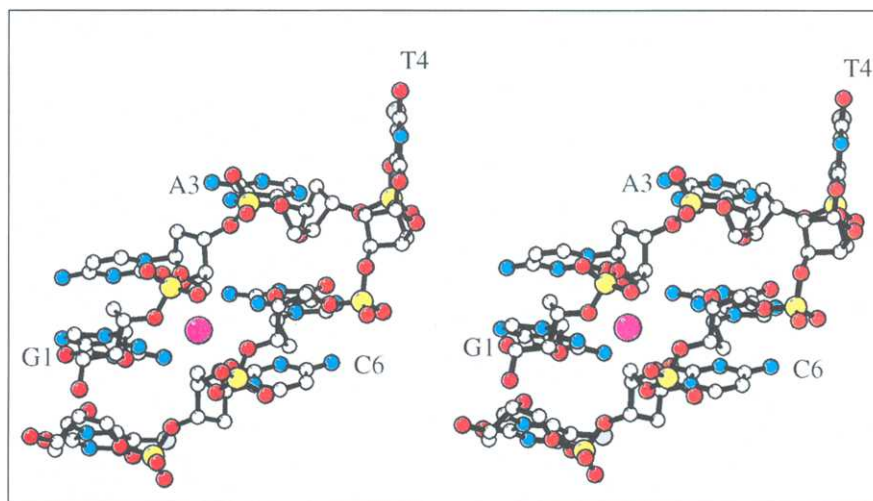
Table 1. Phosphate-furanose backbone and glycosyl torsion angles (°) for d(GCATGGCT).

Residue	α	β	γ	δ	ε	ζ	χ
G1			−152	167	−175	−94	−98
C2	−70	174	44	117	−122	153	−105
A3	−48	143	43	138	−143	70	−140
T4	−79	−166	67	146	−81	−77	−129
G5	−58	−168	60	147	−147	−131	−104
C6	−37	139	39	113	−131	−145	−110
T7	−74	169	67	113			−166

Main-chain torsion angles are defind by O3′-P-α-O5′-β-C5′-γ-C4′-δ-C3′-ε-O3′-ζ-P-O5′. The glycosyl torsion angle, χ, is defined by O4′-C1′-N1-C2 for pyrimidines and O4′-C1′-N9-C4 for purines.

reminiscent of sugar–base interactions observed in Z-DNA, nucleoside and nucleotide crystal structures [16].

In the single-stranded (ss) DNA structure of canine parvovirus [15], the torsion angles associated with the sugar-phosphate backbone bear greater similarity to values observed in tRNA structures. We conclude that, in the case of the viral ssDNA and in the present study, the pliability of the sugar-phosphate backbone allows the accommodation of strong stacking interactions, hydrogen bond formation and a geometrical arrangement of phosphates to coordinate cations effectively.

Quadruplex formation by loop dimers

In the crystal lattice, two d(GCATGCT) strands associate in an antiparallel manner related by a two-fold axis of symmetry (Fig. 3a, b). The bases at the GpC steps in the stems of different loops pair up in a Watson–Crick manner to form a quadruplex structure. A further interstrand base pair, involving the adenine bases in the bend regions of the loops, is also formed (Fig. 1a), as detailed below. Two of the loop dimers associate, one on top of the other, to form a compact four-stranded assembly (Fig. 3c). The T4 nucleotides have unstacked bases and are involved in intermolecular T·T base pairing that helps to stabilize the crystal structure (Fig. 1b).

The core of the quadruplex structure consists of two (C·G)·(G·C) quartets that exhibit almost perfect stacking of the pyrimidine bases on the six-membered ring of the guanines (Fig. 4). Each of the two inter-loop base pairs is stabilized by Watson–Crick hydrogen bonds and the quartets are completed by the two intra-strand minor groove hydrogen bonds discussed above. Thus, each C·G quartet is held together by the same number of hydrogen bonds (eight) as the G-quartet [1,2]. The base pairs are not coplanar with the partners in the quartet — instead they are tilted by ~30°. Unlike the structure of the guanine quartets in the structure of four-stranded *Oxytricha* telomeric DNA [6–8,17] the sugar-phosphate chains are not found at the corners of the quartets but are arranged in such a way that the individual loops are further stabilized by van der Waals interactions between the deoxyribose moieties. A consequence of this conformation is the close approach of the G2 and C6 phosphate groups. The phosphorus–phosphorus distance is 5.15 Å. The potential destabilization of the structure is alleviated by the presence of an ion, modelled during refinement as a magnesium ion (Fig. 2). The presence of this cation must be important in stabilizing this loop structure. Furthermore, the ion interacts with another two phosphate groups of a symmetry-related loop and displays a slightly distorted octahedral coordination completed with two solvent molecules. The cation is thus able to stabilize the close contact of four phosphates.

The d(GCATGCT) loop quadruplex adopts neither the 'Greek key' nor the 'Indian key' conformations [9] that have been observed in the structures of quadruplexes formed by *Oxytricha* telomeric DNA [6,7]. It is more similar to models initially proposed by Sundquist and Klug [2], and Williamson *et al.* [4]. The bend regions of the loops link bases on the same side of the quartets (Fig. 2) and are both above the stack of quartets. As detailed earlier, the bend regions then interact by the formation of an inter-loop A·A base pair (Figs 1a, 2) that helps to stabilize the four-stranded assembly through base-stacking interactions with the symmetry-related loop dimer.
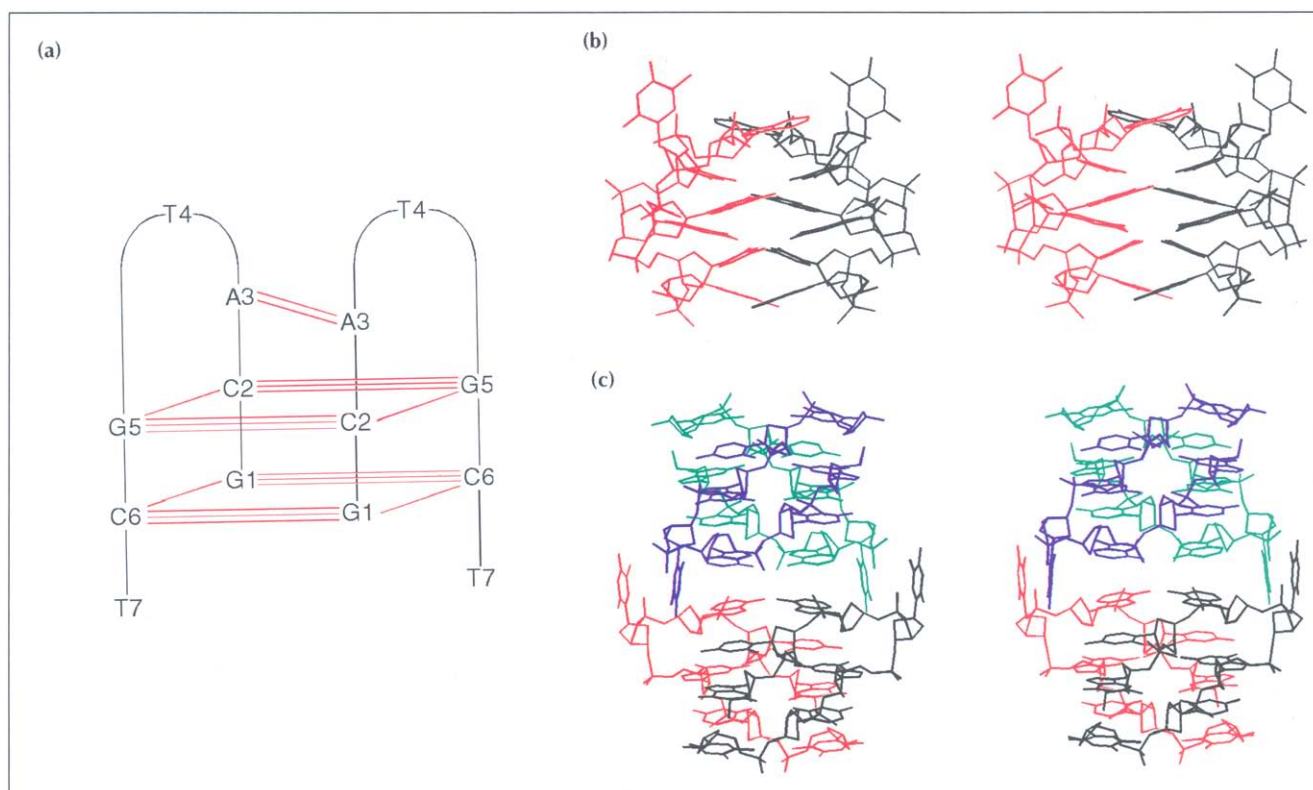
Fig. 3. (a) Schematic diagram of the d(GCATGCT) loop dimer. Nucleotides are numbered from G1 to T7 in the 5'→3' direction. Red lines identify hydrogen bonds. **(b)** Stereoview line drawings illustrating how two strands associate to form a quadruplex structure and **(c)** the four-stranded assembly. The asymmetric unit is in black. T7 caps the loop and T4 is unstacked. In (b), the two T4 residues of the loop dimer are at the top of the figure, T7 at the bottom. Both bases contribute to stabilizing the crystal structure through interactions with symmetry-related molecules. (Stereo figures produced with MOLSCRIPT [33].)
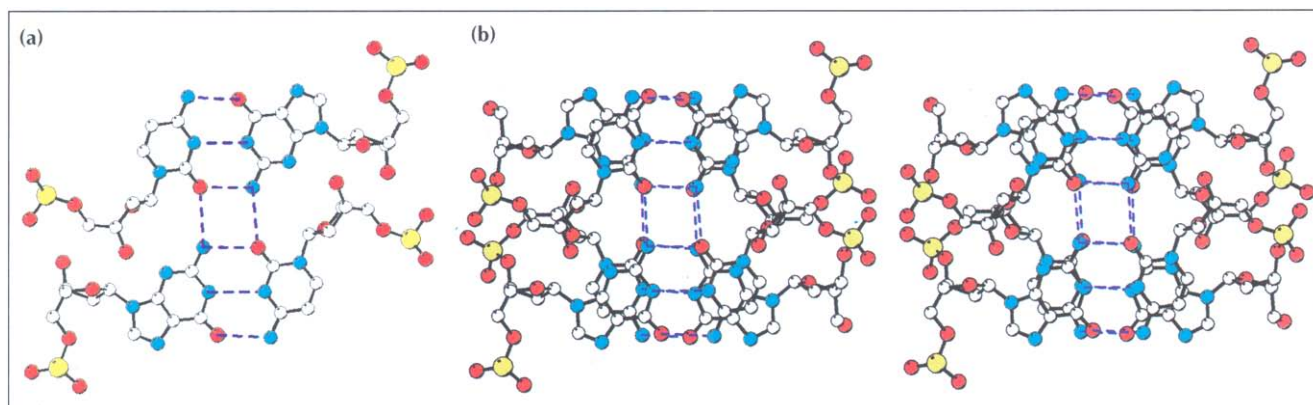
Fig. 4. Base–base hydrogen-bonding interactions. **(a)** One of the (C·G)·(G·C) quartets observed in the quadruplex structure adopted by d(GCATGCT). The atom colouring is the same as for Fig. 2. C2 is shown at the bottom-right, G5 at top-right and the symmetry-related C2 and G5 are top-left and bottom-left, respectively. The quartets are stabilized by an array of eight hydrogen bonds, shown as blue dashed lines, two of which are intra-strand and six of which join bases on different loops. **(b)** Stereoview to show the base-stacking pattern that stabilizes the dimer loop assembly. (Figure produced with MOLSCRIPT [33].)

Adenine·adenine and thymine·thymine pairs

The A·A base pair adopts a symmetric N7–N6–amino hydrogen-bonded conformation. The hydrogen-bonding distance is 3.0 Å. This base pair is similar to that observed for the A9/A23 association in the crystal structure of yeast tRNA^{Phe} [16,18] and in the A·A pair found in the CpA–proflavin complex [19]. Such a conformation for the A·A base pair does not involve hydrogen bonding to the Watson–Crick face of either of the bases. This feature

has previously been discussed with reference to another purine–purine base pair, namely the conformation observed for tandem G·A mismatches in the structure of DNA and RNA duplexes [20,21]. With such a conformation it is possible that other (external) bases could interact with this base pair and thus recognize it by using the functional groups usually reserved for Watson–Crick base pairing. Such a conclusion would also apply to the A·A base pair geometry observed in the present study.

Such a conformation could promote further DNA–DNA or DNA–protein associations.

The T·T base pairs result from a symmetric hydrogen bonding between N3 and O2 atoms of the extruding thymine T4 (Fig. 1b) with a symmetry-related partner. The hydrogen-bonding distance is 2.9 Å. A well-ordered solvent molecule (number 32) helps to stabilize this pairing by bridging the O4 and O2 functional groups. This water molecule is 3.2 Å from O4 and 2.8 Å from O2. T·T pairing has been reported in two other oligonucleotide crystal structures [22,23] but in both cases T·T pairing occurs in regions of local disorder in the crystals. Another example is provided by a protein–DNA complex where the DNA has an overhanging thymine that interacts with a symmetry-related thymine in the crystal lattice [24].

The crystal packing of the hairpin structure d(CGCG-CGTTTTCGCGCG) [22] produces a T·T pairing that involves O4–O4 and N3–N3 interactions. It is necessary to invoke the rare enol tautomer of one of the thymine bases and couple this with a statistical disorder of the bases in the crystal lattice in order for the pairing to be physically and chemically possible. In the second example, r(GCUUCGGC)d(^{Br}U) [23], the stacking of duplex units is facilitated by pyrimidine–pyrimidine pairs of the thymine analogue, bromouracil. Hydrogen bonds are formed between O4 and N3, and between O2 and N3 and mimic a geometry similar to that of a Watson–Crick base pair. However, disorder is evident and bromouracil can adopt two possible orientations. In the present study, as in the example of the trp repressor–operator complex [24], no disorder is observed and therefore there is no need to invoke rare tautomeric forms. The 'wobbling' of the base pair ensures that complementary hydrogen bonding groups of the two bases are aligned. The T·T base pair characterized in the present analysis is an accurately determined example of how this particular interaction, with both bases in the most stable amido tautomeric form, could contribute to sequence recognition in non-helical nucleic acid structures.

Biological implications
The vast majority of structural studies on DNA, either by itself or in complex with proteins, have concentrated on duplex structures [16,25] although more recent work has characterized G-rich quadruplexes [6–8]. These studies illustrate the conformational variability of DNA but given the range of biological roles that DNA performs, and the tight packaging it undergoes, it is evident that other DNA conformations are possible. The crystallographic analysis of d(GCATGCT) provides details of a DNA conformation not previously observed. The structure presents features that may be relevant to the tight packaging of nucleic acids in single-stranded DNA viral genomes such as that of parvovirus [15]. These features include

non-Watson–Crick base pairing and the use of cations, not only to allow negatively charged phosphate groups to associate and to stabilize a loop structure, but also to bring together four distant parts of a nucleic acid chain. This could complement the formation of stable intermolecular base-pairing and base-stacking arrangements between nucleic acids.

It is noteworthy that guanines are not the only nucleotides that can participate in quartet structures. This suggests that additional sequences should be considered when quartet formation is being invoked as relevant to biological processes such as replication, telomerase inhibition, chromosomal DNA organization, meiosis and mitosis. In addition, we have provided an experimentally derived model whose structural features indicate how "the recognition and association of identical, rather than complementary DNA sequences" [2] might occur.

Materials and methods
Synthesis, crystallization and data collection
d(GCATGCT) and several brominated derivatives were prepared on an Applied Biosystems automatic DNA synthesizer. Purification was by reversed-phase high-performance liquid chromatography. Crystals of the native compound and only one of the derivatives, d(G^{Br}-CATCGT), were grown by vapour diffusion from sitting drops containing the heptanucleotide (~5 mg ml^{-1}), 50 mM lithium cacodylate buffer at pH 6.8, 50 mM magnesium chloride, 3mM spermine tetrahydrochloride and 10% (v/v) hexane-1,6-diol (HD) equilibrated against an external reservoir containing 50% HD. The conditions were maintained between 18°C and 22°C and plate-shaped crystals appeared after 1–2 weeks. The crystals developed cracks after 2 months but could be regenerated by dissolution with the addition of water and replenishing the reservoir. Many of the crystals were twinned.

Crystals were first characterized by precession photography using an Enraf-Nonius camera coupled to a sealed tube X-ray source with a copper target (nickel-filtered CuKα radiation, λ=1.5418 Å). The crystals are orthorhombic, space group C222, and diffract to high resolution. Accurate unit cell dimensions of a=22.52 Å, b=59.37 Å, c=24.35 Å for the native sequence and a=22.44 Å, b=59.65 Å, c=24.27 Å for the derivative, were obtained from the least-squares fit of 15 centred reflections using Syntex and Rigaku diffractometer systems. In many cases the DNA polymorph can be identified from the precession photographs, but with d(GCATGCT) we were unable to ascertain any features characteristic of A-, B- or Z-form DNA. The symmetry and size of the unit cell indicated that the asymmetric unit volume of ~4000 Å^3 could contain either 3.5 base pairs or a single strand of DNA.

Intensity data for the native sequence were collected from one crystal of dimensions 0.35 mm×0.20 mm×0.20 mm at 6°C on a Syntex P2$_1$ diffractometer with CuKα radiation (graphite monochromator). A total of 4259 measurements to 1.8 Å resolution reduced to 1606 unique reflections with R$_{sym}$=2.5% [R$_{sym}$=(Σ|I$_i$–<I$_i$>|/ΣI$_i$)×100, where <I$_i$> is the average of I$_i$

over all symmetry equivalents]. An empirical absorption correction [26] was applied with minimum transmission of 0.57. The bromine derivative was first shown to be suitable from a 2.5 Å resolution data set collected at the Laboratory of Molecular Biology, Cambridge on an Enraf-Nonius FAST TV-detector/rotating anode system using CuKα radiation. Subsequently data were collected at 4°C using monochromatic synchrotron radiation, tuned to provide a wavelength that optimized anomalous differences due to resonance scattering from the bromine (λ=0.9102 Å), with a MAR image-plate detector on station PX 9.5 of the Synchrotron Radiation Source, Daresbury Laboratory [27]. Two data sets were collected from the same derivative crystal which had dimensions of 0.35 mm×0.25 mm×0.15 mm. The first data set yielded 6419 measurements to 2.0 Å resolution which reduced to 1078 unique reflections with R_{sym}=5.1%. For the second data set the incident beam intensity was attenuated using aluminium foil, to allow measurement of the very intense medium and low angle data which had previously been saturated. A total of 7830 measurements to 2.0 Å resolution reduced to 1175 reflections with R_{sym}=6.5%.

Structure solution and refinement

The structure was solved using a modified single isomorphous replacement with optimized anomalous scattering technique. It was advantageous to keep the two derivative data sets separate in the phase calculation [28,29]. A partial model consisting of four nucleotides was constructed from the first map and phase recombination of experimental and model based phases then produced a map of high quality (figure of merit=0.79) and completion of the model was straightforward. All calculations were carried out using the CCP4 suite of programs [30]. The structure was refined using simulated annealing techniques [12] followed by a restrained least-squares procedure [13,14] interspersed with graphics fitting with FRODO [31] of the model to the calculated electron density. Several rounds of refinement and electron-density fitting were carried out with the careful inclusion of solvent molecules and cations. The crystallographic R-factor converged at 21.4% for all 1579 reflections with F>0 in the resolution range 7.0–1.8 Å (96% complete) and 19.2% for those 1282 reflections in the resolution range with F>2σ(F).

Coordinates and structure factors have been deposited with the Brookhaven PDB (tracking codes T5194 and T5195, respectively) and the Cambridge Crystallographic Databank.

References

1. Sen, D. & Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364–366.
2. Sundquist, W.I. & Klug, A. (1989). Telomeric DNA dimerises by formation of guanine tetrads between hairpin loops. *Nature* **342**, 825–829.
3. Zahler, A.M., Williamson J.R., Cech, T.R. & Prescott, D.M. (1991). The inhibition of telomerase by G-quartet structures. *Nature* **350**, 718–720.
4. Williamson, J.R., Raghuraman, M.K. & Cech, T.R. (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* **59**, 871–880.
5. Wang, Y., Sin, R., Gaffney, B., Jones, R.A. & Breslauer, KJ. (1991). Characterisation by ¹H NMR of glycosidic conformations in the tetramolecular complex formed by d(GGTTTTGG). *Nucleic Acids Res.* **19**, 4619–4622.
6. Smith, F.W. & Feigon, J. (1992). Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature* **356**, 164–168.
7. Kang, C., Zhang, X., Ratliff, R., Moyzis, R. & Rich, A. (1992). Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature* **356**, 126–131.
8. Wang, Y. & Patel, D.J. (1993). Solution structure of the human telomeric repeat d[AG₃(T₂AG₃)₃] G-tetraplex. *Structure* **1**, 263–278.
9. Frank-Kamenetskii, M. (1989). The turn of the quadruplex? *Nature* **342**, 737.
10. Frank-Kamenetskii, M. (1992). The quadruplex and the vase. *Nature* **356**, 105.
11. Williamson, J.R. (1993). G-quartets in biology: reprise. *Proc. Natl. Acad. Sci USA* **90**, 3124.
12. Brünger, A.T. (1990). *X-PLOR. Manual Version 2.1.* Yale University, New Haven, CT.
13. Hendrickson, W.A. & Konnert, J.H. (1981). In *Biomolecular Structure, Conformation, Structure and Evolution*. Vol. **1**, (Srinivasan, R., ed), pp. 43–57, Pergammon Press, Oxford.
14. Westhof, E., Dumas, P. & Moras, D. (1985). Crystallographic refinement of yeast aspartic acid transfer RNA. *J. Mol. Biol.* **184**, 119–145.
15. Chapman, M.S. & Rossmann, M.G. (1995). Single-stranded DNA–protein interactions in canine parvovirus. *Structure* **3**, 151–162.
16. Saenger, W. (1984). *The Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
17. Laughlin, G., *et al.*, & Luisi, B. (1994). The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science* **265**, 520–524.
18. Tinoco, I. (1993). Structures of base pairs involving at least two hydrogen bonds. In *The RNA World.* (Gesteland, R.F. & Atkins, J.F., eds), pp. 603–607, Cold Spring Harbour Laboratory Press, NY.
19. Westhof, E., Rao, S.T. & Sundaralingam, M. (1980). Crystallographic studies of drug–nucleic acid interactions: proflavin intercalation between the non-complementary base pairs of cytidyl-3′,5′-adenosine. *J. Mol. Biol.* **142**, 331–361.
20. SantaLucia, J. & Turner, D.H. (1993). Structure of (rGGCGAGCC)₂ in solution from NMR and restrained molecular dynamics. *Biochemistry* **32**, 12612–12623.
21. Lane, A.N., Ebel, S. & Brown, T. (1994). Properties of multiple G·A mismatches in stable oligonucleotide duplexes. *Eur. J. Biochem.* **220**, 717–727.
22. Chattopadhyaya, R., Grzeskowiak, K. & Dickerson, R.E. (1990). Structure of a T4 hairpin loop on a Z-DNA stem and comparison with A-RNA and B-DNA loops. *J. Mol. Biol.* **211**, 189–210.
23. Cruse, W.B.T., Saludjian, P., Biala, E., Strazewski, P., Prangé, T. & Kennard, O. (1994). The structure of a mispaired RNA double helix at 1.6 Å resolution and implications for the prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. USA* **91**, 4160–4164.
24. Otwinowski, Z., *et al.*, & Sigler, P.B. (1988). The crystal structure of the trp repressor/operator complex at atomic resolution. *Nature* **335**, 321–329.
25. Kennard, O. & Hunter, W.N. (1991). Single-crystal X-ray diffraction studies of oligonucleotides and oligonucleotide–drug complexes. *Angew. Chemie* **30**, 1254–1277.
26. North, A.C.T., Phillips, D.C. & Mathews, F.S. (1968). A semi-empirical method of absorption correction. *Acta Crystallogr. A* **24**, 580–584.
27. Brammer, R.C., *et al.*, & Worthington, K. (1988). A new protein crystallography station on the SRS wiggler beamline for very rapid Laue and rapidly tunable monochromatic experiments: 1. design principles, ray tracing and heat calculations. *Nucl. Instrum. and Meth. A* **271**, 678–687.
28. Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. & Saraste, M. (1992). Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **359**, 851–855.
29. Jones, T.A., Bergfors, T., Sedzik, J. & Unge, T. (1988). The three-dimensional structure of P2 myelin protein. *EMBO J.* **7**, 1597–1604.
30. Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763.
31. Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–276.
32. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119.
33. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.