J|A|C|S
ARTICLES
Published on Web 01/30/2007

Four-Stranded DNA Structures Can Be Stabilized by Two Different Types of Minor Groove G:C:G:C Tetrads

Núria Escaja,[‡] Irene Gómez-Pinto,[†] Enrique Pedroso,*[,‡] and Carlos González*[,†]

Contribution from the Instituto de Química Física "Rocasolano", CSIC, C/, Serrano 119, 28006 Madrid, Spain, and Departament de Química Orgánica, Universitat de Barcelona, C/, Martí i Franquès 1-11, 08028 Barcelona, Spain

Received August 25, 2006; E-mail: cgonzalez@iqfr.csic.es; epedroso@ub.edu

Abstract: Four-stranded nucleic acid structures are central to many processes in biology and in supramolecular chemistry. It has been shown recently that four-stranded DNA structures are not only limited to the classical guanine quadruplex but also can be formed by tetrads resulting from the association of Watson−Crick base pairs. Such an association may occur through the minor or the major groove side of the base pairs. Structures stabilized by minor groove tetrads present distinctive features, clearly different from the canonical guanine quadruplex, making these quadruplexes a unique structural motif. Within our efforts to study the sequence requirements for the formation of this unusual DNA motif, we have determined the solution structure of the cyclic oligonucleotide d⟨pCCGTCCGT⟩ by two-dimensional NMR spectroscopy and restrained molecular dynamics. This molecule self-associates, forming a symmetric dimer stabilized by two G:C:G:C tetrads with intermolecular G−C base pairs. Interestingly, although the overall three-dimensional structure is similar to that found in other cyclic and linear oligonucleotides of related sequences, the tetrads that stabilize the structure of d⟨pCCGTCCGT⟩ are different to other minor groove G:C:G:C tetrads found earlier. Whereas in previous cases the G−C base pairs aligned directly, in this new tetrad the relative position of the two base pairs is slipped along the axis defined by the base pairs. This is the first time that a quadruplex structure entirely stabilized by slipped minor groove G:C:G:C tetrads is observed in solution or in the solid state. However, an analogous arrangement of G−C base pairs occurs between the terminal residues of contiguous duplexes in some DNA crystals. This structural polymorphism between minor groove GC tetrads may be important in stabilization of higher order DNA structures.

Introduction

Among the different noncanonical DNA motifs, quadruplex DNA structures are probably the most extensively studied. These structures have attracted considerable attention from many research areas, ranging from molecular and structural biology to supramolecular chemistry and nanotechnology.[1] Four-stranded DNA may play a role in several biological processes, such as telomere integrity, genetic recombination, transcription, or replication (for reviews see refs 1−6). In addition, this DNA motif is an attractive target for drug design, especially in cancer chemotherapy.[7] Clear evidence of quadruplex formation in vivo has been found recently by electron microscopy[8] and by generation of quadruplex specific antibodies.[9,10]

Moreover, the number of potential applications of quadruplex DNA in supramolecular chemistry and material science is enormous, ranging from formation of biosensors and nanomachines to the assembly of different nanostructures like G-wires,[11] or frayed wires,[12] that could provide a basis for developing advanced biomaterials. Recently it has been described its potential use as a transmembrane transporter,[13] for nanoparticle assembly,[14] or as an electrochemical molecular beacon aptasensor.[15]

The fundamental unit of the G-quadruplex is the G-tetrad, where four guanines are paired through their Watson−Crick and Hoogsteen sides. In some cases, guanine quadruplexes contain unusual tetrads. For example, tetrads consisting only of adenine,[16] thymine,[17,18] and cytosine[19] have been found within

[†] Instituto de Química Física "Rocasolano".
[‡] Universitat de Barcelona.

(1) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668−98.
(2) Arthanari, H.; Bolton, P. H. *Chem. Biol.* **2001**, *8*, 221−30.
(3) Suhnel, J. *Biopolymers* **2001**, *61*, 32−51.
(4) Keniry, M. A. *Biopolymers* **2000**, *56*, 123−46.
(5) Phan, A. T.; Kuryavyi, V.; Patel, D. J. *Curr. Opin. Struct. Biol.* **2006**, *16*, 288−98.
(6) Patel, D. J.; Bouaziz, S.; Kettani, A.; Wang, Y. In *Oxford Handbook of Nucleic Acid Sructures*; Neidle, S., Ed.; Oxford University Press: New York, 1999; pp 389−453.
(7) Hurley, L. H. *Nat. Rev. Cancer* **2002**, *2*, 188−200.
(8) Duquette, M. L.; Handa, P.; Vincent, J. A.; Taylor, A. F.; Maizels, N. *Genes Dev.* **2004**, *18*, 1618−29.
(9) Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H. J.; Pluckthun, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8572−7.
(10) Paeschke, K.; Simonsson, T.; Postberg, J.; Rhodes, D.; Lipps, H. J. *Nat. Struct. Mol. Biol.* **2005**, *12*, 847−54.
(11) Marsh, T. C.; Vesenka, J.; Henderson, E. *Nucleic Acids Res.* **1995**, *23*, 696−700.
(12) Protozanova, E.; Macgregor, R. B., Jr. *Biochemistry* **1996**, *35*, 16638−45.
(13) Kaucher, M. S.; Harrell, W. A., Jr.; Davis, J. T. *J. Am. Chem. Soc.* **2006**, *128*, 38−9.
(14) Li, Z.; Mirkin, C. A. *J. Am. Chem. Soc.* **2005**, *127*, 11568−9.
(15) Radi, A. E.; Acero Sanchez, J. L.; Baldrich, E.; O'Sullivan, C. K. *J. Am. Chem. Soc.* **2006**, *128*, 117−24.
(16) Patel, P. K.; Koti, A. S.; Hosur, R. V. *Nucleic Acids Res.* **1999**, *27*, 3836−43.

10.1021/ja066172z CCC: $37.00

the scaffold of parallel guanine quadruplexes. Also, tetrads containing Watson−Crick base pairs have been found in parallel[20,21] and fold-back quadruplex structures.[22−24] In all these cases, the Watson−Crick base pairs associate through their major groove side. In the particular case of the G−G−G−C repeats of the adeno-associated viral DNA, two conformations of the major groove G:C:G:C tetrad are observed, depending on the cation present in the sample.[25] In the sodium form the two G−C pairs align directly opposite to each other, whereas in the presence of potassium the alignment of the two base pairs is slipped, with the two guanines coordinating the K^+ cation. These two arrangements have been also observed in major groove A:T:A:T tetrads. The slipped alignment was found in the dimeric solution structure of the octamer d(GACCAGGT),[24] whereas the direct A:T:A:T alignment has been observed between the external loops of the propeller-like quadruplex structure of d($AG_3(T_2AG_3)_3$), containing four human telomeric repeats.[26]

On the other hand, tetrads formed by minor groove alignment of Watson−Crick base pairs were first found in the crystallographic structure of d(GCATGCT)[27] and, more recently, in the crystallographic and solution structures of several cyclic oligonucleotides. In the case of d⟨pCATTCATT⟩ the dimer is stabilized by two minor groove A:T:A:T tetrads,[28,29] and in the case of d⟨pTGCTCGCT⟩ and d⟨pCGCTCATT⟩, the structures are stabilized by minor groove G:C:G:C[29] and G:C:A:T[30] tetrads, respectively. Whereas major groove tetrads (like the G-tetrads) are mainly planar, in the minor groove association the two base pairs have a mutual inclination of around 30°−40°. It is interesting to notice that major groove tetrads have only been observed in quadruplexes containing other pure guanine tetrads. In contrast, minor groove tetrads have been found in structures where no other kind of tetrad is present. Most probably the inclination between the two base pairs provokes that these tetrads are not compatible with pure guanine tetrads, which are planar, giving rise to a distinctive DNA fold.

We are interested in exploring other possible base-pair arrangements susceptible of forming minor groove tetrads as well as the sequence requirements for their formation. Although this motif has been observed in several crystallographic structures of the linear heptamer d(GCATGCT)[27,31] and, more recently, in hybrid complexes between linear and cyclic oligonucleotides,[32] self-associating cyclic oligonucleotides appear to be the best models to study these noncanonical DNA structures.

In this paper we report on the three-dimensional solution structure of d⟨pCCGTCCGT⟩ as obtained by restrained molecular dynamics calculations based on NMR derived experimental constraints. Whereas in diluted conditions all experimental evidence indicates that this oligonucleotide does not adopt a defined structure, at higher oligonucleotide concentration this molecule dimerizes by forming intermolecular G−C Watson−Crick base pairs. The two cyclic oligonucleotides are arranged in an antiparallel way. The core of the dimer consists in two tetrads resulting from the association of two G−C base pairs through their minor groove sides. The structure of this oligonucleotide is discussed in comparison with the dimeric structures of d⟨pTGCTCGCT⟩[29] and d(GCATGCT),[27] which are also stabilized by minor groove G:C:G:C tetrads. Most interestingly, although the overall fold of these three molecules is very similar, a single change in the order of the core residues (GC to CG) causes the formation of a different G:C:G:C tetrad. Whereas, in previous cases, the interaction between the base pairs occurs through two bifurcated H21(G)−O2(C) hydrogen bonds, in the structure of d⟨pCCGTCCGT⟩ the relative position of the two base pairs is slipped along the axis defined by the base pairs, and the interaction occurs through two H21(G)−N3(G) hydrogen bonds.

Methods

Experimental Details. The cyclic octamers were synthesized as reported by Alazzouzi et al.[33] Samples were suspended (in Na^+ salt form) in either D_2O or 9:1 H_2O/D_2O (25 mM sodium phosphate buffer, pH = 7). All NMR spectra were acquired in Bruker spectrometers operating at 600 and 800 MHz and processed with the XWIN-NMR software. In the experiments in D_2O, presaturation was used to suppress the residual H_2O signal. A jump-and-return pulse sequence[34] was employed to observe the rapidly exchanging protons in 1D H_2O experiments. NOESY[35] spectra in D_2O were acquired with mixing times of 100, 200, and 300 ms. TOCSY[36] spectra were recorded with the standard MLEV-17 spin-lock sequence and a mixing time of 80 ms. In 2D experiments in H_2O, water suppression was achieved by including a WATERGATE[37] module in the pulse sequence prior to acquisition. ^{31}P resonances were assigned from proton-detected heteronuclear correlation spectra.[38] ^{13}C resonances were assigned from natural abundance $^{13}C-^{1}H$ HSQC spectra. The spectral analysis program SPARKY[39] was used for semiautomatic assignment of the NOESY cross-peaks and quantitative evaluation of the NOE intensities.

Circular dichroism spectra at different temperatures were collected on a Jasco J-810 spectropolarimeter fitted with a thermostated cell holder. CD spectra were recorded at a 5 μM oligonucleotide concentration in H_2O or in salt conditions (10 mM Na_2PIPES buffer, pH 7, 100 mM NaCl, and 10 mM $MgCl_2$). For melting experiments, the samples were initially heated at 90 °C for 5 min and slowly cooled to room temperature and stored at 4 °C until use.

NMR Constraints. The intrinsic ambiguity between inter- and intramolecular distances in dimeric structures could be overcome by

(17) Patel, P. K.; Hosur, R. V. *Nucleic Acids Res.* **1999**, *27*, 2457−64.
(18) Caceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A. *Nucleic Acids Res.* **2004**, *32*, 1097−102.
(19) Patel, P. K.; Bhavesh, N. S.; Hosur, R. V. *Biochem. Biophys. Res. Commun.* **2000**, *270*, 967−71.
(20) Webba da Silva, M. *Biochemistry* **2003**, *42*, 14356−65.
(21) Webba da Silva, M. *Biochemistry* **2005**, *44*, 3754−64.
(22) Kettani, A.; Kumar, R. A.; Patel, D. J. *J. Mol. Biol.* **1995**, *254*, 638−56.
(23) Kettani, A.; Bouaziz, S.; Gorin, A.; Zhao, H.; Jones, R. A.; Patel, D. J. *J. Mol. Biol.* **1998**, *282*, 619−36.
(24) Zhang, N.; Gorin, A.; Majumdar, A.; Kettani, A.; Chernichenko, N.; Skripkin, E.; Patel, D. J. *J. Mol. Biol.* **2001**, *312*, 1073−88.
(25) Bouaziz, S.; Kettani, A.; Patel, D. J. *J. Mol. Biol.* **1998**, *282*, 637−52.
(26) Parkinson, G. N.; Lee, M. P.; Neidle, S. *Nature* **2002**, *417*, 876−80.
(27) Leonard, G. A.; Zhang, S.; Peterson, M. R.; Harrop, S. J.; Helliwell, J. R.; Cruse, W. B.; d'Estaintot, B. L.; Kennard, O.; Brown, T.; Hunter, W. N. *Structure* **1995**, *3*, 335−40.
(28) Salisbury, S. A.; Wilson, S. E.; Powell, H. R.; Kennard, O.; Lubini, P.; Sheldrick, G. M.; Escaja, N.; Alazzouzi, E.; Grandas, A.; Pedroso, E. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5515−8.
(29) Escaja, N.; Pedroso, E.; Rico, M.; González, C. *J. Am. Chem. Soc.* **2000**, *122*, 12732−12742.
(30) Escaja, N.; Gelpi, J. L.; Orozco, M.; Rico, M.; Pedroso, E.; González, C. *J. Am. Chem. Soc.* **2003**, *125*, 5654−52.
(31) Thorpe, J. H.; Teixeira, S. C.; Gale, B. C.; Cardin, C. J. *Nucleic Acids Res.* **2003**, *31*, 844−9.
(32) Escaja, N.; Gómez-Pinto, I.; Viladoms, J.; Rico, M.; Pedroso, E.; González, C. *Chemistry* **2006**, *12*, 4035−4042.
(33) Alazzouzi, E.; Escaja, N.; Grandas, A.; Pedroso, E. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 1506−1508.
(34) Plateau, P.; Güeron, M. *J. Am. Chem. Soc.* **1982**, *104*, 7310−7311.
(35) Kumar, A.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1−6.
(36) Bax, A.; Davies, D. J. *J. Magn. Reson.* **1985**, *65*, 355−360.
(37) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661−5.
(38) Sklenar, V.; Miyashiro, H.; Zon, G.; Miles, H. T.; Bax, A. *FEBS Lett* **1986**, *208*, 94−8.
(39) Goddard, D. T.; Kneller, G. *SPARKY*, 3; University of California: San Francisco, CA.

using similar procedures as were employed in previously calculated structures of the same family.[29,30] Most of the cross-peaks could be directly assigned on the basis of a rough molecular model and taking into account the similarities with the dimeric structure of d⟨pTGCTCGCT⟩.[29] Trial assignments were made for the few remaining cross-peaks and used in preliminary structure calculations. When a distance constraint was consistently violated in all the resulting structures, an alternative assignment was considered. After several cycles of assignment and structure calculation, a consistent set of constraints was obtained. Those cross-peaks that were still ambiguous were not included in the subsequent calculations. These initial calculations were carried out with qualitative distance constraints (classified as 3, 4, or 5 Å).

Refined structures were calculated employing more accurate distance constraints, obtained from NOE cross-peak intensities by using a complete relaxation matrix analysis with the program MARDIGRAS.[40] No solvent exchange effects were taken into account in the analysis of NOE intensities in H_2O, and therefore, only upper limits were used in the distance constraints involving labile protons. Error bounds in the interprotonic distances were estimated by carrying out several MARDIGRAS calculations with different initial models, mixing times, and correlation times. Three initial models were chosen from the structures resulting from the "low-resolution" DYANA calculation (vide infra). Correlation times of 1.0, 2.0, and 4.0 ns were employed, assuming, in all cases, a single correlation time for the whole molecule (isotropic motion). Experimental intensities were recorded at three different mixing times (100, 200, and 300 ms) for nonexchangeable protons and at a single mixing time (200 ms) for labile protons. Final constraints were obtained by averaging the upper and lower distance bounds in all the MARDIGRAS runs. A lower limit of 1.8 Å was set in those distances where no quantitative analysis could be carried out, such as very weak intensities or cross-peaks involving labile protons. In addition to these experimentally derived constraints, Watson−Crick hydrogen bond restraints were used. Target values for distances and angles related to hydrogen bonds were set as described from crystallographic data.

Torsion angle constraints for the sugar moieties were derived form the analysis of *J*-coupling data obtained from DQF-COSY experiments. Since only the sums of coupling constants were estimated, rather loose values were set for the dihedral angle of the deoxyriboses (δ angle between 110° and 170°, ν_1 between 5° and 65°, and ν_2 between −65° and −50°). Additional constraints for the γ angles of the backbone were used in those cases where stereoespecific assignments of H5′/H5″ resonances could be made.

Structure Determination. Structures were calculated with the program DYANA 1.4[41] and further refined with the SANDER module of the molecular dynamics package AMBER 5.0.[42] Initial DYANA calculations were carried out on the basis of qualitative distance constraints. The resulting structures were used as initial models in the complete relaxation matrix calculations to obtain accurate distance constraints, as described in the previous paragraph. These structures were taken as starting points for the AMBER refinement, consisting of an annealing protocol in vacuo, followed by long trajectories where explicit solvent molecules are included. In vacuo calculations were performed with hexahydrated Na^+ counterions, which were place around the exterior of the molecule and near the phosphate groups. The electrostatic term was calculated using a distance dependent dielectric constant, and the cutoff value for nonbounded interactions was 10 Å. The temperature and the relative weights of the experimental constraints were varied during the simulations according to standard annealing protocols used in our group.[29,43] The resulting structures were refined including explicit solvent and using the Particle-Mesh-Ewald method to evaluate long-range electrostatic interactions. The vacuum structures were placed in the center of a water box with around 3000 water molecules and 16 sodium counterions. The structures were first equilibrated for 160 ps using standard equilibration process.[44] The final equilibrated structures were then subjected to 10 runs of 500 ps MD simulations in water using NMR restraints, periodic boundary conditions, and the Particle-Mesh-Ewald method.[45] Simulations in water were performed in the isothermal−isobaric ensemble ($P = 1$ atm, $T = 298$ K) using SHAKE in all atoms and a 2 fs time step for integration of Newton's equations. MD-averaged structures were obtained by averaging the last 250 ps of individual trajectories and further relaxation of the structure. The same procedure was used to obtain a global MD-averaged structure, resulting from averaging 5 ns of independent MD-trajectories for each isomer. The AMBER-98 force field[46] was used to describe the DNA, and the TIP3P model was used to describe water molecules.[47] Analysis of the representative structures as well as the MD trajectories was carried out with the programs Curves V5.1[48] and MOLMOL.[49]

Results

Monomer−Dimer Equilibrium and Melting Behavior. Like in other cyclic oligonucleotides previously studied in our laboratory, the NMR spectra of d⟨pCCGTCCGT⟩ depend strongly on temperature and oligonucleotide concentration. At low temperature and salt concentration, an interconverting equilibrium between two species is observed (see Figure 1). This equilibrium is slow on the NMR time scale, so that resonances from both species can be observed simultaneously. Since the populations of both species depend on the oligonucleotide concentration, it can be concluded that this is an association equilibrium between a dimeric structure and a monomeric form. Under these circumstances, the equilibrium constants at different temperatures can be determined from the ratio of the peak areas. CD and NMR data show that the major conformer at low concentration (the monomeric form) lacks a well-defined structure.[50]

The melting behavior of the dimeric form of d⟨pCCGTCCGT⟩ is similar to that of the related cyclic oligonucleotide d⟨pTGCTCGCT⟩.[51] Exchangeable protons are observed even at temperatures close to the melting transition, suggesting slow dissociation kinetics (see Figure 1). Thermodynamic parameters for this dimer−random coil equilibrium were estimated from a van't Hoff analysis of the equilibrium constants at different temperatures, determined from the ratio of the peak areas in each species. At a 25 mM Na^+ concentration, the estimated $\Delta G°_{298}$ value is −35 kJ/mol.

The melting transition was also followed by circular dichroism. The CD spectra at low temperature exhibit a large positive band around 270 nm and a negative band around 245 nm. Both bands disappear progressively as the temperature is increased,

(40) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1990**, *87*, 475−487.
(41) Güntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, *273*, 283−98.
(42) Case, D. A. et al. *AMBER*, 5; University of California: San Francisco, CA, 1997.
(43) Soliva, R.; Monaco, V.; Gomez-Pinto, I.; Meeuwenoord, N. J.; Marel, G. A.; Boom, J. H.; González, C.; Orozco, M. *Nucleic Acids Res.* **2001**, *29*, 2973−85.
(44) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7463−7469.
(45) Darden, T. E.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(46) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(47) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(48) Lavery, R.; Sklenar, H. *CURVES, helical analysis of irregular nucleic acids.*, 3.0; Laboratory of Theoretical Biochemistry CNRS: Paris, 1990.
(49) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 29−32.
(50) Escaja, N.; Gómez-Pinto, I.; Rico, M.; Pedroso, E.; González, C. *ChemBioChem* **2003**, *4*, 623−32.
(51) González, C.; Escaja, N.; Rico, M.; Pedroso, E. *J. Am. Chem. Soc.* **1998**, *120*, 2176−2177.

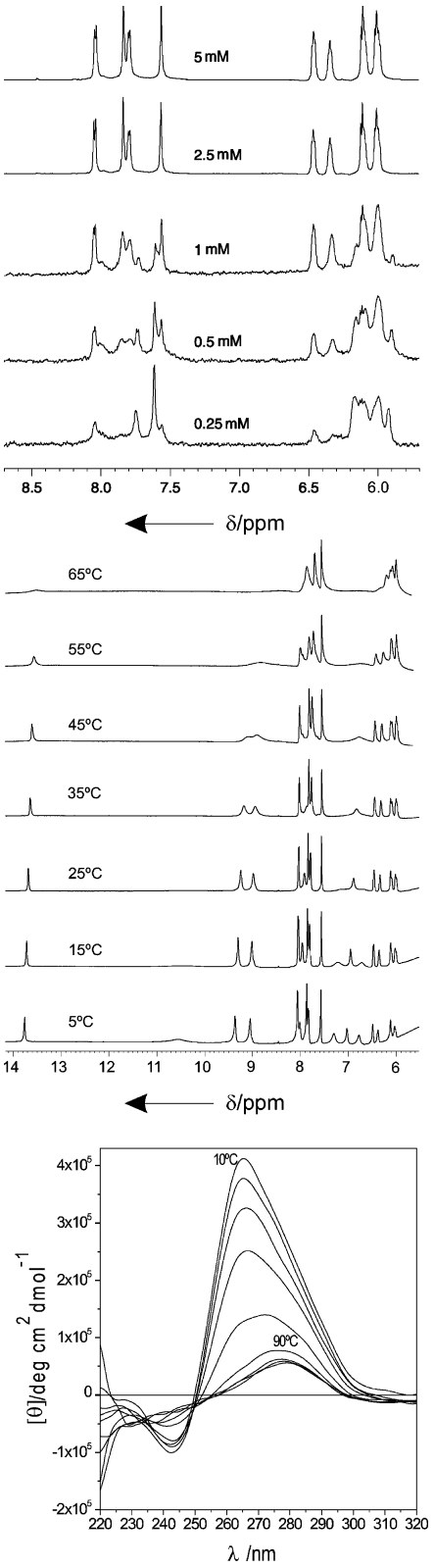

Figure 1. (Top) One-dimensional NMR spectra of d⟨pCCGTCCGT⟩ in D_2O at 5 °C and different oligonucleotide concentrations (25 mM sodium phosphate buffer, pH = 7). (Middle) NMR spectra of d⟨pCCGTCCGT⟩ in H_2O (5 mM oligonucleotide concentration, 100 mM NaCl, pH = 7) at different temperatures. (Bottom) CD spectra at different temperatures ranging from 10 to 90 °C (5 μM oligonucleotide concentration, 10 mM Na_2PIPES pH = 7, 100 mM NaCl, 10 mM $MgCl_2$).

with the maximum of the positive band shifting to larger wavelengths.

NMR Assignment. Sequential assignments of exchangeable and nonexchangeable protons and ^{13}C and ^{31}P resonances were conducted following standard methods. Some regions of the two-dimensional NOESY spectrum of d⟨pCCGTCCGT⟩ in H_2O are shown in Figure 2. Many spectral features are common to other structures of this family studied by NMR. The number of signals in the spectra indicates that the dimer is symmetric. All intra-nucleotide H1′-base NOEs are medium or weak, indicating that the glycosidic angle in all the nucleotides is in an *anti* conformation. Strong sugar-base sequential connections were observed between residues 2 → 3, and therefore 6 → 7. Sequential NOEs between residues 3 → 4 (7 → 8) are clearly observed in the H1′-base region but are very weak in the H2′/H2″-base region (see Figure 2). Almost all resonances were identified, including the stereoespecific assignment of some H5′/H5″ protons (see Table S1). Interestingly, in the NMR spectra of d⟨pTGCTCGCT⟩ (and in other molecules of the same family studied by NMR), the H4′/H5′/H5″ protons of the nucleotides in the first position of the loops present rather unusual chemical shifts. This effect is not observed in the spectra of d⟨pCCGTCCGT⟩.

The exchangeable proton spectra are particularly informative (see Figure 2). Amino protons of all cytosines were assigned from their NOE cross-peaks with the H5 protons. In the case of C1(5), these resonances are broad and completely degenerated, but in the case of C2(6), the two amino protons present strong cross-peaks with the imino signal observed at 13.76 ppm. This signal was assigned to G3(7), and its amino resonances could be also identified. This NOE pattern is characteristic of GC Watson−Crick base pairs. It is interesting to note that both guanine amino protons are shifted downfield (9.37 and 8.02 ppm). One more imino proton is observed in the 10.51 ppm region, indicating that the thymines in the loops are not base paired.

Assignment of ^{31}P and ^{13}C resonances was straightforward thanks to the good quality of the heteronuclear experiments (see Figure 3). In particular, the good dispersion of ^{31}P signals is a clear indication of a compact and well-structured molecule.

Experimental Constraints and Structure Calculations. In all dimeric molecules there is an intrinsic ambiguity between inter- and intramolecular cross-peaks. In the present case, these ambiguities were resolved easily thanks to the similarities with other structures of the same family previously studied in our group.[29,30] In the few cases where several distances might correspond with an NOE cross-peak, the final assignment could be done by carrying out several structure calculations with trial assignments (see Methods for details). After the complete relaxation matrix calculation, a total of 338 distance constraints where obtained. A summary of these constraints is shown in Table 1 and in Figure 4.

In addition to the NOE-derived information, an analysis of the *J*-coupling constants obtained from DQF-COSY spectra was carried out. The population of the major S conformer, estimated from the sum of $J_{1'2'}$ and $J_{1'2''}$,[52] was in all cases greater than 90%. Vicinal *J*-coupling constants between H4′ and H5′ and H5″ were small in all cases except in residue 1 (and 5). These *J*-values together with the pattern of intraresidual NOEs between H3′/H4′ and H5′/H5″ obtained using a short mixing time

(52) Rinkel, L. J.; Altona, C. *J. Biomol. Struct. Dyn.* **1987**, *4*, 621−49.

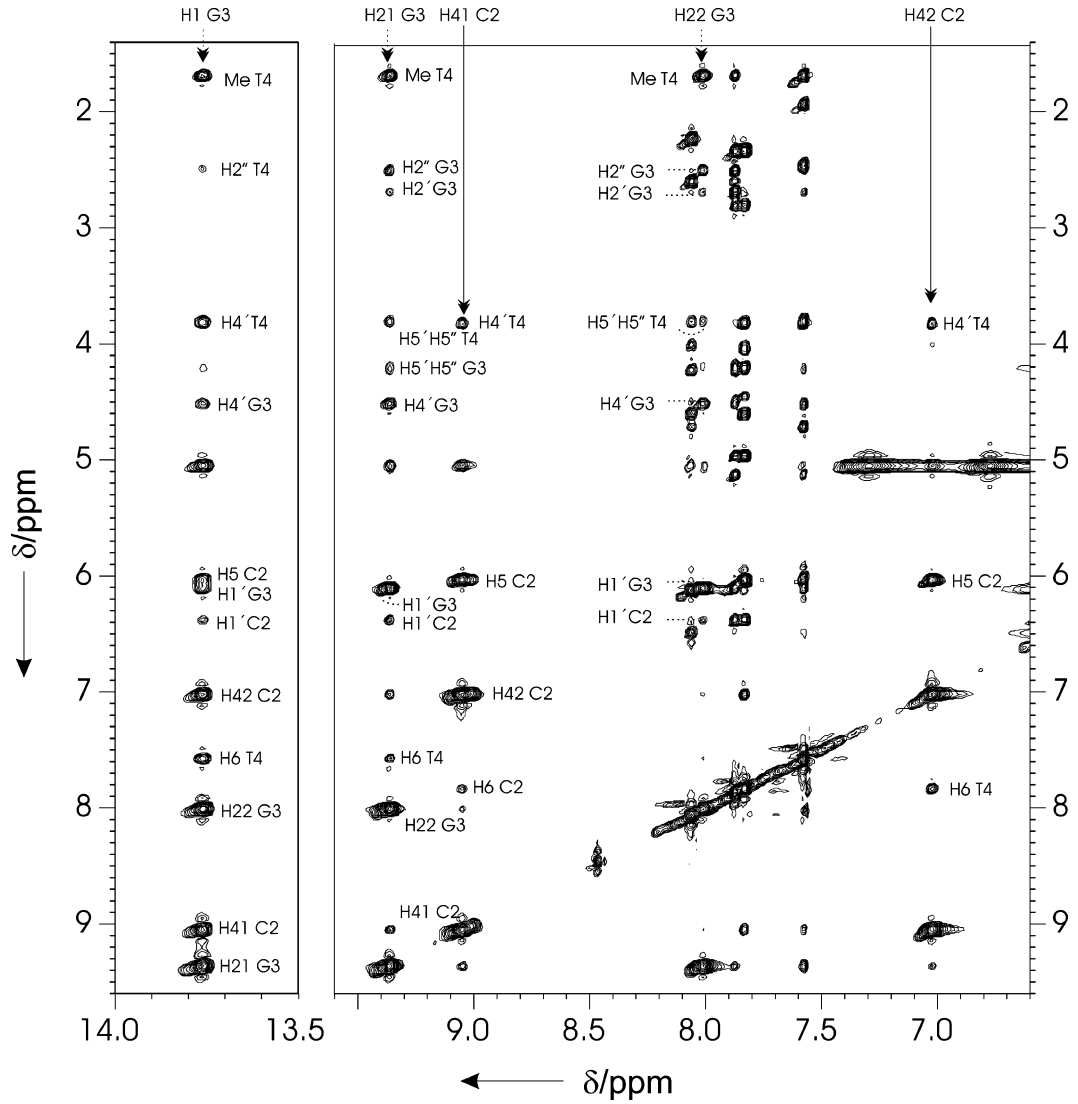

Figure 2. Regions of the NOESY spectrum (150 ms mixing time) of d⟨pCCGTCCGT⟩ in H_2O (5 mM oligonucleotide concentration, 100 mM NaCl, T = 5 °C, pH = 7). Watson−Crick base pairing can be established from the H1G2−H42C3, H1G2−H41C3, and H1G2(6)−H5C3(7) cross-peaks for the GC pair. Sequential assignment pathways for nonexchangeable protons are shown.

permitted the stereospecific assignment of some of the H5′/H5″[53] and indicates that the backbone angle γ in these nucleotides is in a single *gauche*$^+$ conformation. In addition, the good quality of $^{31}P-^1H$ correlation spectra (shown in Figure 3) permitted us to obtain dihedral angular constraints for the angles β and ϵ of the backbone. All the $^{31}P-^1H_{5'/5''}$ coupling constants are small, indicating the β torsion angle is mainly in *trans*. Also, none of the $^{31}P-^1H_{3'}$ coupling constants is larger than 10 Hz, and consequently, the ϵ angles are not in a *gauche* $^+$ conformation.[53]

All these experimental constraints were used to calculate the structure by using restrained molecular dynamics methods. Initial structures were calculated with the program DYANA and then refined with the AMBER package following two steps: first an annealing protocol in vacuo and then a long restrained molecular dynamics calculation including the solvent explicitly. Except for the C1(5) and the corresponding ones in the symmetry related subunit, all residues in the final structures are very well-defined, with an rmsd of 0.5 Å (see Figure 5). These values are even lower when only atoms in the bases are considered (see Table 1). The final AMBER energies and NOE terms are reasonably low in all the structures, with no distance constraint violation greater than 0.5 Å. Due to the large number of experimental constraints this oligonucleotide structure is particularly well-defined. Only the cytosines in the second position of the loops (residues 1, 5, 9, and 13) are less defined than the rest of the molecule, as shown by the order parameters of some of their torsion angles (see Table S3).

Description of the Structures. The resulting structure is a dimer consisting of two molecules of d⟨pCCGTCCGT⟩ arranged in an antiparallel way. As expected from the complete degeneration of the signals in the NMR spectra, the dimer is symmetric. This symmetry is reflected in the geometrical parameters (see Table S3). The two octamers associate with each other by forming four intermolecular Watson−Crick base pairs (see Figure 5). These base pairs form two G:C:G:C tetrads by aligning their minor groove sides. All glycosidic angles are *anti*, with values ranging from −110° to −152°. In addition to the six Watson−Crick hydrogen bonds, each tetrad is stabilized by two additional intermolecular hydrogen bonds between one of

(53) Wijmenga, S. *Progr. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287−387.

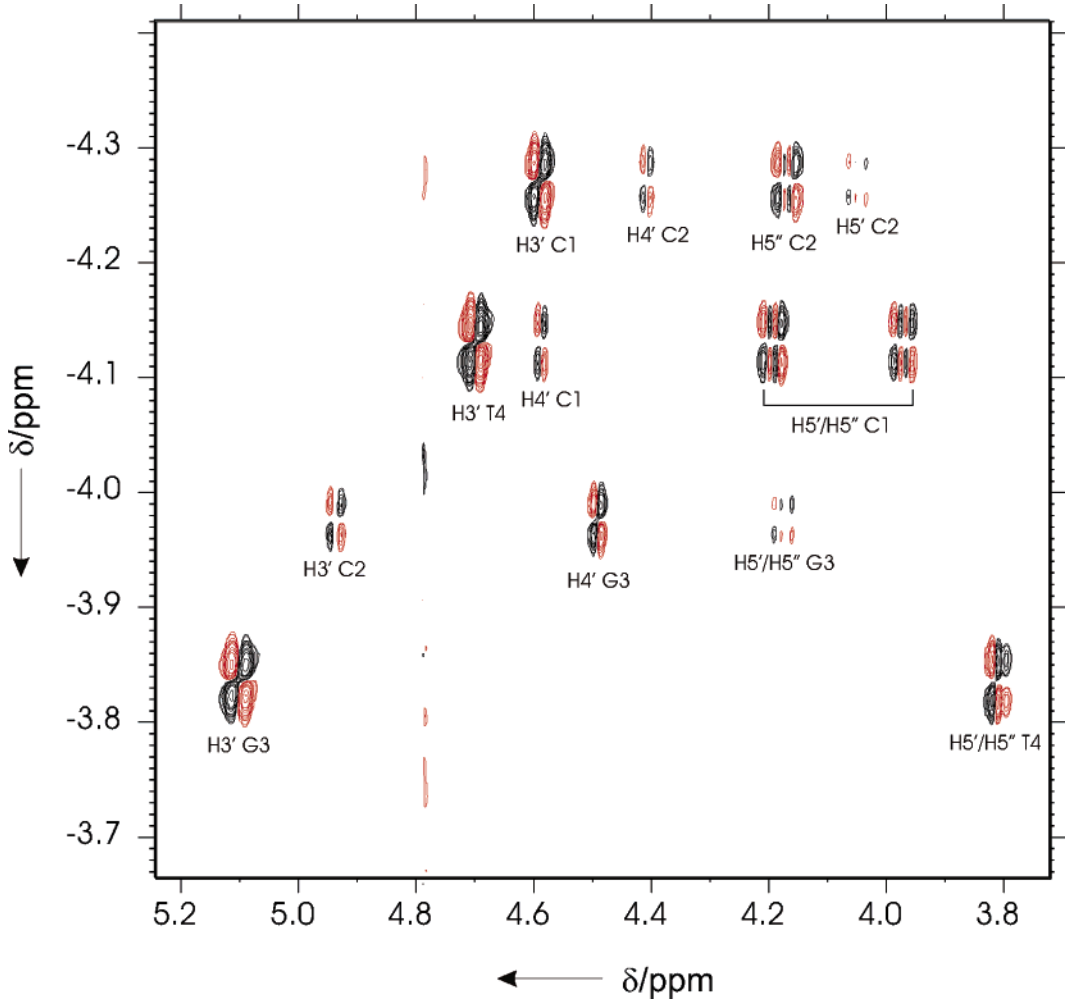

Figure 3. Heteronuclear $^{31}P-^{1}H$ correlation spectrum of d⟨pCCGTCCGT⟩ in D_2O (5 mM oligonucleotide concentration, 100 mM NaCl, $T = 5$ °C, pH = 7). H3′(i)−P(i + 1) and P(i)−H5′/5″(i) cross-peaks are indicated.

Table 1. Experimental Constraints and Calculation Statistics of d⟨pCCGTCCGT⟩

experimental distance constraints	
total number	338
intraresidue	113
sequential	74
range > 1	151
intramolecular	226
intermolecular	112

rmsd (Å)	
all well-defined bases[a]	0.30 ± 0.07 Å
all well-defined heavy atoms[a]	0.20 ± 0.02 Å
backbone	0.3 ± 0.07 Å
all heavy atoms	0.5 ± 0.1 Å

residual violations	average	range
sum of violation (Å)	6.9	6.80−7.08
max violation (Å)	0.3	0.27−0.31
NOE energy (kcal/mol)	40	38−42

[a] All except residues 1, 5, 9, and 13.

the amino protons of G3 and the N3 of the guanine in the adjacent base pair. This hydrogen bond is responsible for the unusual chemical shifts observed for the guanine amino protons. The amino proton resonating at 9.37 ppm is forming the Watson−Crick hydrogen bond, while the other proton (8.02 ppm) is involved in the hydrogen bond with N3. Both amino protons are very close to the sugar moiety of the opposite guanine. Unusually strong NOEs can be observed between these amino protons and the guanine H1′ and H4′ (see Figure 2). Like in other minor groove tetrads, the two base pairs are not in the same plane (see Figure 5) but have a mutual inclination of around 40°. The central C−G step exhibits a rise of 2.8 Å and a twist value of 42°. Some other local helical parameters are shown in Table 2.

The structure is very well packed, with the thymines in the first position of the loops forming two caps at both ends of the stacks of G−C base pairs and interacting with each other through hydrophobic contacts between their methyl groups. The glycosidic angles of these residues present values around −150°.

All deoxyribose rings are in the general S-domain, adopting in most cases a C2′-endo conformation. Only the deoxyriboses of nucleotides in position 3 (and 7) adopt an O4′-endo conformation, with pseudorotation phase angles of around 90°. A list of backbone angles is provided in the Supporting Information (Table S3). Most of the backbone angles are also well-defined and, in general, present values usually found in right-handed double-stranded DNA. The main exception is the ζ angle between residues 4 and 5 (8 and 1), which is in a *gauche*+ conformation. This angle, together with the α of C1 (and 5), is responsible for the chain reversal. The resulting tight turn causes several phosphate groups to be very close to each

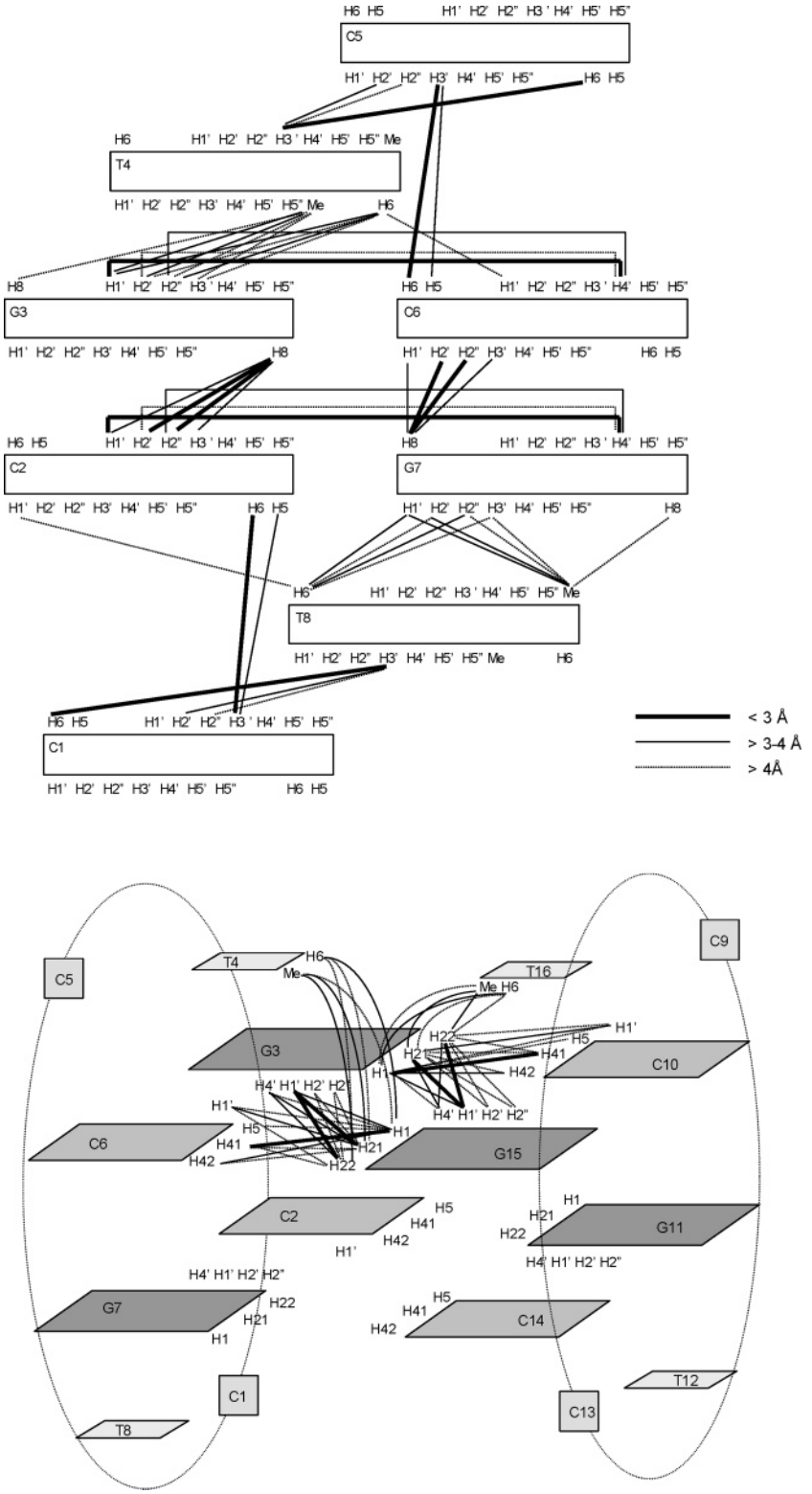

Figure 4. Schematic representation of (top) intra- and (bottom) intermolecular distance constraints. Constraints are classified in three categories according to their upper distance limit.

other; i.e., the distances between phosphorus atoms located in residues 3−4 and 7−8 are especially short. This close proximity between two strands gives rise to two extremely narrow grooves with a width of around 1 Å (considering P−P distance 7 Å and a sum of van der Waals radii of 5.8 Å). The width of the other two grooves is similar to that of the major groove in a B-DNA duplex (see Figure S1 in the Supporting Information).

Discussion

A Very Well-Defined Solution Structure. The dimeric solution structure of d⟨pCCGTCCGT⟩ is very well-defined. Such a high definition is a direct consequence of the number and quality of the distance constraints (around 40 per nucleotide). The distribution of these constraints is even more important. A

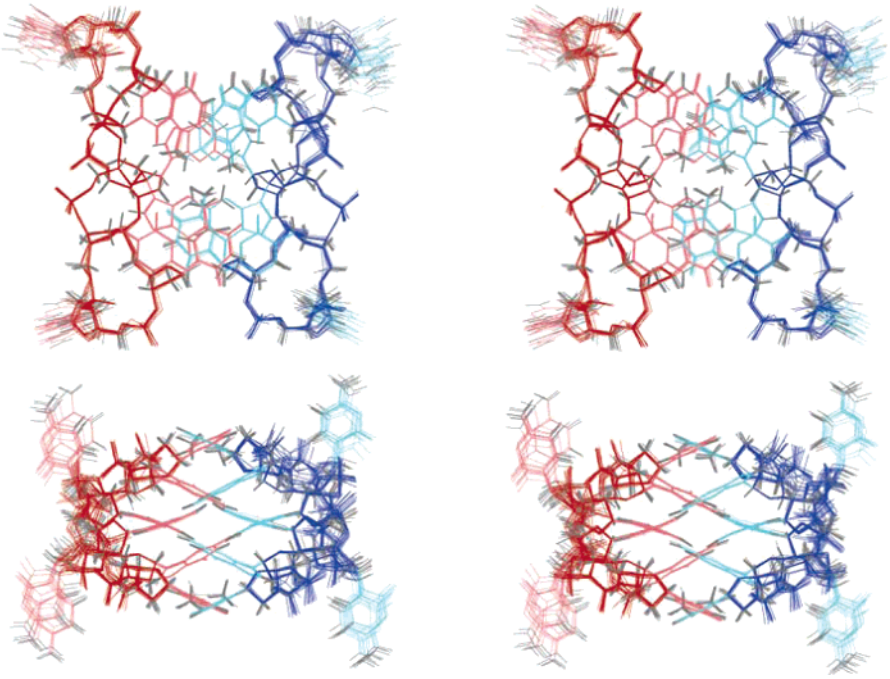

Figure 5. (Top) Stereoscopic views of the superposition of the 10 refined structures of d⟨pCCGTCCGT⟩. The axis of the molecule runs perpendicular to the view plane. (Bottom) Stereoscopic view of the average structure (lateral view). Red and blue indicate different molecules. The sugar−phosphate backbone is indicated in darker colors.

Table 2. Local Helical Parameters for the Dimeric Structure of d⟨pCCGTCCGT⟩

	shift Dx (Å)		slide Dy (Å)		rise Dz (Å)		tilt τ (deg)		roll ρ (deg)		twist (deg)	
step	average	SD	average	SD	average	SD	average	SD	average	SD	average	SD
C2/G3	0.9	0.1	−0.2	0.1	2.8	0.02	10.3	0.7	−11.5	1.1	41.8	0.4
G3/T4	−8.3	0.2	−1.5	0.1	7.4	0.2	−48.1	1.4	10.7	0.8	−24.2	1.7
C6/G7	0.9	0.1	−0.2	0.1	2.8	0.03	10.7	0.7	−12.8	1.3	41.7	0.6
G7/T8	−8.3	0.1	−1.6	0.1	7.4	0.2	−49.5	1.2	10.4	1.5	−24.5	1.4
T12/G11	8.3	0.04	−1.9	0.1	6.9	0.2	50.3	1.2	13.1	0.7	−26.7	0.6
G11/C10	−0.7	0.1	−0.3	0.1	2.8	0.04	−10.6	0.7	−12.5	1.5	42.4	0.5
T16/T15	8.3	0.1	−2.0	0.1	6.7	0.2	49.8	1.2	15.3	1.7	−26.5	1.2
G15/C14	−0.7	0.04	−0.3	0.1	2.8	0.04	−10.6	0.6	−11.4	0.8	42.6	0.6

Table 3. Atomic rmsd's (Å) between the Structures of Several Oligonucleotides Stabilized by Minor Groove Tetrads[a]

	d<pTGCTCGCT>	d<pCATTCATT>	d<pCGCTCATT>	d<pCCGTCCGT>
d<pTGCTCGCT>		1.2	2.4	1.2
d<pCATTCATT>	0.8		2.6	1.8
d<pCGCTCATT>	1.1	0.6		1.7
d<pCCGTCCGT>	0.8	0.9	1.0	
d(GCATGCT)	0.7	0.9	1.0	0.9

[a] Upper: rmsd between all C1′. Lower: rmsd between the C1′ of the nucleotides involved in the tetrads.

total of 151 distance constraints are neither intraresidual nor sequential (112 of them are intermolecular). This is almost half of the total number (see Table 1). Such a distribution of distance constraints is more typical of a globular protein than of a DNA fragment, where most of the constraints are intranucleotide or sequential. In general, these intraresidual and sequential constraints are less informative than long-range ones. Also, the repetitive sequence reduces the complexity of the NMR spectra, facilitating the extraction of many distance and angular constraints from spectral regions that are normally very crowded. For example, the H4′/H5′/5″ region could be analyzed exhaustively, including the stereospecific assignment of some H5′/5″ resonances and the estimation of some $J_{4',5'/5''}$ coupling constants.

Comparison with Other Structures of the Family. The structure of d⟨pCCGTCCGT⟩ belongs to the family of quadruplex structures stabilized by minor groove aligned tetrads. The strong similarity between the structures of the family is apparent from the rmsds between the sugar−phosphate backbone of several members of the family shown in Table 3. In particular, the backbone rmsd between d⟨pCCGTCCGT⟩ and d⟨pTGCTCGCT⟩ is 1.8 Å. This strong similarity confirms that minor groove aligned tetrads impose important restrictions on the three-dimensional structure of the quadruplex.

One of the main differences in the structure of d⟨pCCGTCCGT⟩ compared to the previous ones is the pronounced twist between the two consecutive G−C pairs (Figure 6). The stacking between the two central G−C base pairs and also between these pairs with the capping thymines is not as favorable in this structure as in that of d⟨pTGCTCGCT⟩ or d(GCATGCT).

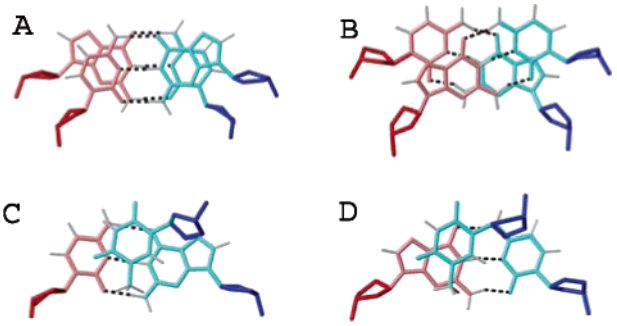

Figure 6. Details of the stacking interaction between G−C base pairs in the structure of d⟨pTGCTCGCT⟩ (A) and d⟨pCCGTCCGT⟩ (B). Details of the interaction with the capping thymines are shown in panels (C) and (D).

Unusual values for the chemical shifts of the H4′ and H5′/5″ protons of residues in the first position of the loops (T4 and T8 in all cases) are observed in the NMR spectra of d⟨pTGCTCGCT⟩, d⟨pCATTCATT⟩, and d⟨pCGCTCATT⟩. These unusual downfield shifts are due to the proximity of the purine bases in position 6 and 2, respectively, in the three-dimensional structure, and they are not observed in the NMR spectra of d⟨pCCGTCCGT⟩. In this case, the residues in the equivalent position are pyrimidines, and the shift induced by their ring current effects is smaller than that in purines.

Structural Polymorphism of Minor Groove Tetrads. In spite of the similar global fold, the nature of the minor groove tetrad and the type of interactions that stabilize them is very different in each case. In A:T:A:T and G:C:A:T tetrads, the stabilization comes from the presence of a cation in the center of the molecule, probably coordinating the O2 atoms of the surrounding thymines.[28,30] In G:C:G:C tetrads, this stabilization is due to extra hydrogen bonds between the adjacent G−C base pairs. Comparison of the structure of d⟨pCCGTCCGT⟩ with the dimeric structures of the cyclic octamer d⟨pTGCTCGCT⟩ and the linear heptamer d(GCATGCT) shows that there are two possible arrangements of G−C base pairs through their minor grooves. Whereas in the two latter cases the two G−C pairs align directly opposite each other, in the structure of d⟨pCCGTCCGT⟩ the relative position of the two base pairs is strongly shifted along the axis defined by the Watson−Crick base pairs. Such a shift has a strong influence in the pattern of hydrogen bonds that stabilizes the tetrads. While in the direct G:C:G:C tetrad the interaction between the base pairs occurs through two bifurcated H21(G)−O2(C) hydrogen bonds, in the "slipped" G:C:G:C tetrad the interaction occurs through two H21(G)−N3(G) hydrogen bonds.

Two different alignments between G−C base pairs have been also observed when the association occurs through their major groove. In the direct alignment, the tetrad is stabilized by two H6C−O6G and H6C−N7G bifurcated hydrogen bonds. In the slipped alignment, however, the stabilization comes from cation coordination,[25] as shown in Figure 7.

Thermal Stability. It is interesting to compare the stability of the different types of minor groove tetrads. Experimental ΔG values have been estimated for the dimeric structures of d⟨pTGCTCGCT⟩,[29] d⟨pCATTCATT⟩,[29] and d⟨pCGCTCATT⟩.[30] These values are −32, −13, and −9 kJ/mol, respectively. The relative stability between the structures derives from different interactions within each tetrad, since the structures of all the dimers are very similar. G:C:G:C, with direct hydrogen bonds between the base pairs, is more stable than A:T:A:T and G:C:A:T, which are stabilized by coordinating a Na^+ ion. The experimental ΔG value for the dimeric structure of d⟨pCCGTCCGT⟩ obtained in this work (−35 kJ/mol) is very similar to that of d⟨pTGCTCGCT⟩, indicating a similar stability of the two minor groove G:C:G:C tetrads.

This similar stability observed in the dimeric structures of d⟨pCCGTCCGT⟩ and d⟨pTGCTCGCT⟩ contrasts with the dramatic differences observed in their monomeric forms. In a previous study we found that, at low oligonucleotide concentration, d⟨pTGCTCGCT⟩ forms a dumbbell-like structure that melts with a T_m of 51 °C.[50] However, the NMR and CD spectra of the monomeric form of d⟨pCCGTCCGT⟩ clearly indicate that this molecule has no detectable structure. This difference in stability reflects the tendency, observed in other DNA hairpins[54,55] and dumbbells,[56] for the sequence 5′-C-XX-G-3′ to form two-residue loops, whereas the sequence 5′-G-XX-C-3′ prefers to adopt loops of four residues.

The dimeric structures of d⟨pCCGTCCGT⟩ and d⟨pTGCTCGCT⟩ can be envisioned as examples of hairpin−hairpin association, like the dimeric structure of the linear heptamer d(GCATGCT). The comparison between the relative stabilities of the dimeric and monomeric forms of d⟨pCCGTCCGT⟩ and d⟨pTGCTCGCT⟩ clearly shows that the rules governing the stability of monomeric DNA hairpins do not necessarily apply to the structures formed by multiple hairpin association. Conclusions extracted from thermodynamic studies on single DNA hairpins should not be directly extrapolated to biological processes where hairpin−hairpin interaction might be involved, like DNA repeat expansions.[57,58] Our finding that sequences that do not tend to form monomeric hairpins can adopt stable hairpinlike structures through dimerization should be taken into account when analyzing the propensities of different sequences to form hairpins or slipped structures.

Minor Groove Interactions in Higher-Order DNA Structures. To our knowledge, this is the first time that such a shifted association through the minor groove side of G−C base pairs has been found in solution. However, similar minor groove tetrads have been found between symmetry-related molecules in crystallographic structures of several DNA and DNA−RNA hybrid duplexes.[59,60] Hydrogen bonding between HN2 and N3 of terminal guanines in neighboring molecules is a common interaction that contributes to stabilize the crystal packing.[61,62] In many cases, these terminal guanines are not base paired and invade the minor groove of an adjacent duplex, giving rise to a base triplet.[63,64] In other cases, the guanine HN2 and N3

(54) Hilbers, C.; Heus, H.; van Dongen, M.; Wijmenga, S. The hairpin elements of nucleic acid structure: DNA and RNA folding. In *Nucleic Acids and Molecular Biology*; Eckstein, F., Lilley, D. M. J., Eds.; Spinger-Verlag: Berlin, 1994; pp 56−104.
(55) Blommers, M. J.; Walters, J. A.; Haasnoot, C. A.; Aelen, J. M.; van der Marel, G. A.; van Boom, J. H.; Hilbers, C. W. *Biochemistry* **1989**, *28*, 7491−8.
(56) Ippel, J. H.; Lanzotti, V.; Galeone, A.; Mayol, L.; van den Boogaart, J. E.; Pikkemaat, J. A.; Altona, C. *J. Biomol. NMR* **1995**, *6*, 403−22.
(57) Mirkin, S. M. *Curr. Opin. Struct. Biol.* **2006**, *16*, 351−8.
(58) Wells, R. D.; Dere, R.; Hebert, M. L.; Napierala, M.; Son, L. S. *Nucleic Acids Res.* **2005**, *33*, 3785−98.
(59) Wahl, M. C.; Sundaralingam, M. A-DNA duplexes in the crystal. In *Oxford Handbook of Nucleic Acid Structures*; Neidle, S., Ed.; Oxford University Press: New York, 1999; pp 389−453.
(60) Subirana, J. A.; Abrescia, N. G. *Biophys. Chem.* **2000**, *86*, 179−89.
(61) DiGabriele, A. D.; Steitz, T. A. *J. Mol. Biol.* **1993**, *231*, 1024−39.
(62) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *287*, 755−8.
(63) Ramakrishnan, B.; Sundaralingam, M. *J. Biomol. Struct. Dyn.* **1993**, *11*, 11−26.

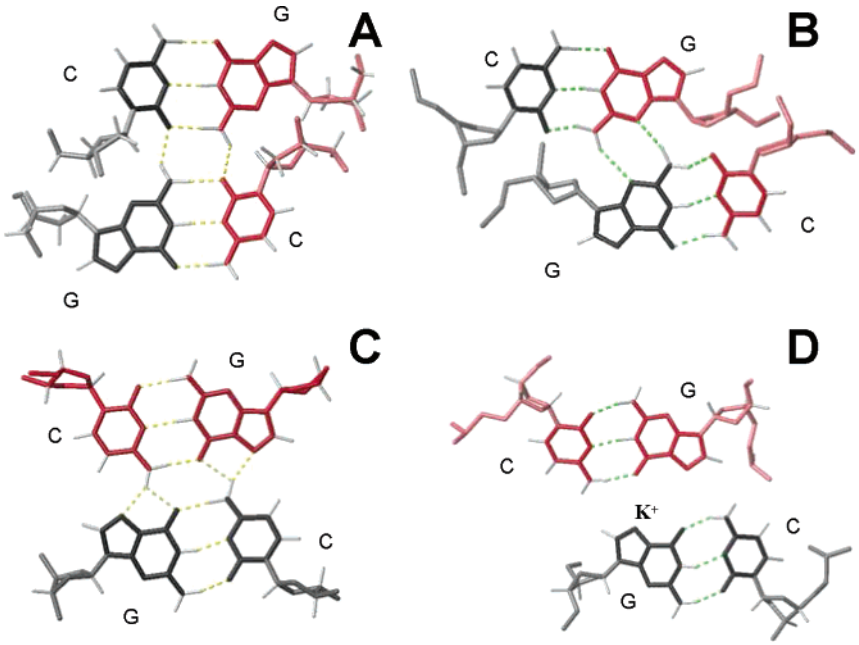

Figure 7. The four G:C:G:C tetrads known: (A) Direct minor groove G:C:G:C tetrad observed in the dimeric solution structures of d⟨pTGCTCGCT⟩; (B) slipped minor groove tetrad found in the solution structure of d⟨pCCGTCCGT⟩; (C) and (D) direct and slipped major groove G:C:G:C tetrads observed in the quadruplex structure of d(GGGCTTTTGGGC).[25]

interaction occurs without disruption of the G−C base pairs, forming G:C:G:C tetrads with a similar structure to the slipped minor groove tetrad reported here.[65,66]

The slipped minor groove tetrads found in the context of a crystallographic network are more common between DNA duplexes that crystallize in the B-form. However, direct minor groove tetrads have been also found in A-form DNA structures.[59] In A-form DNA the minor groove is wider than that in B-DNA, and this might facilitate the interaction.

Interestingly, in the case of a DNA−RNA chimeric duplex of sequence d(CCGGC)r(G)d(CCGG) containing a single 2′-hydroxyl group per strand, the only ribonucleotide is involved in the tetrad.[67] In this case, the minor groove G:C:G:C tetrad is direct (like that in Figure 7), and it is located in the middle of the duplex. This result indicates that minor groove tetrads can be formed with at least one of the four nucleotides being ribo-. However, our studies of association between cyclic and linear oligonucleotides suggest that minor groove tetrads containing ribonucleotides must be much less stable than those formed by four dexoyribonucleotides.[32]

Minor groove G:C:G:C tetrads have been also found involving terminal G and C residues that do not form base pairs with their own duplex but with symmetry related ones, forming junction-like quadruplexes. Such structures have been found in pure DNA crystals and, more recently, in several bisintercalative complexes involving acridine derivatives.[68−72] In these complexes, the acridine derivatives do not intercalate in the usual way but interact with the terminal nucleotides of four DNA duplexes forming a large intercalation platform between two G:C:G:C tetrads. The overall shape of these platforms resemble many features of the dimeric structure of d⟨pCCGTCCGT⟩, such as the relative angle between base pairs which is in most cases around 30°−40°. The structure of theses complexes have attracted significant interest because some of the intercalating drugs involved in them are potent topoisomerase inhibitors. Interestingly, in most cases these structures have been found in crystals containing Co^{2+} ions.[68,70] Such counterions might be relevant to increasing the stability of minor groove tetrads, since cobalt hexamine residues have also been found relevant to the stabilization of the crystal lattice in one of the crystallographic structures of the dimeric linear heptamer d(GCATGCT).[31]

It must be also mentioned that GCGC tetrads have been recently observed in ordered nanostructures based on GC pairing. High-resolution scanning tunneling microscopy studies in adlayers formed by coadsorption of guanine and cytosine at a graphite surface have revealed the formation of well-ordered periodic structures in the solid/liquid interface.[73] Such structures have been attributed to alternate arrangements of GC base pairs though their major and minor groove sides.

Implications in Biology. Tetrads formed by association of Watson−Crick base pairs have been invoked in processes where homologous DNA recognition is required, such as genetic recombination.[74] Since long ago, the association of DNA duplexes through minor groove contacts has been proposed as an initial step in recombination prior to strand exchange.[75] Minor

(64) Tippin, D. B.; Sundaralingam, M. *Acta Crystallogr., Sect. D* **1996**, *52*, 997−1003.
(65) Dickerson, R. E.; Goodsell, D. S.; Kopka, M. L.; Pjura, P. E. *J. Biomol. Struct. Dyn.* **1987**, *5*, 557−79.
(66) Tereshko, V.; Subirana, J. A. *Acta Crystallogr., Sect. D* **1999**, *55*, 810−9.
(67) Ban, C.; Ramakrishnan, B.; Sundaralingam, M. *J. Mol. Biol.* **1994**, *236*, 275−85.
(68) Adams, A.; Guss, J. M.; Denny, W. A.; Wakelin, L. P. *Acta Crystallogr., Sect. D* **2004**, *60*, 823−8.
(69) Adams, A.; Guss, J. M.; Collyer, C. A.; Denny, W. A.; Wakelin, L. P. *Nucleic Acids Res.* **2000**, *28*, 4244−53.
(70) Thorpe, J. H.; Hobbs, J. R.; Todd, A. K.; Denny, W. A.; Charlton, P.; Cardin, C. J. *Biochemistry* **2000**, *39*, 15055−61.
(71) Teixeira, S. C.; Thorpe, J. H.; Todd, A. K.; Powell, H. R.; Adams, A.; Wakelin, L. P.; Denny, W. A.; Cardin, C. J. *J. Mol. Biol.* **2002**, *323*, 167−71.
(72) Yang, X. L.; Robinson, H.; Gao, Y. G.; Wang, A. H. *Biochemistry* **2000**, *39*, 10950−7.
(73) Xu, S.; Dong, M.; Rauls, E.; Otero, R.; Linderoth, T. R.; Besenbacher, F. *Nano Lett* **2006**, *6*, 1434−8.
(74) McGavin, S. *J. Mol. Biol.* **1971**, *55*, 293−8.
(75) Wilson, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 3641−5.

groove tetrads may be responsible for the four-stranded structures observed in poly(CA)·poly(TG) fragments[76] and for some unusual fold-back structures recently observed in some palindromic repeats.[77] On the basis of biochemical data, it has been suggested that such fold-back structures induce RecA-independent homologous recombination.[78] Electrophoretic analysis of these repeats indicates the presence of quadruplex structures, but several biophysical data suggest that these quadruplexes are different than the usual G-quadruplex. The dimethyl sulfate (DMS) binding experiments are especially interesting, as the lack of methylation in the guanine N7 is an indication of the presence of G-tetrads, since the N7 is not accessible in these structures. The observation by Shukla et al. of fold-back quadruplex structures in which all the guanines are sensitive to DMS modification[77] may indicate the presence of minor groove G:C:G:C tetrads. It should be noticed that, in major groove G:C:G:C tetrads, either direct or slipped, the N7 position is buried (see Figure 7), and therefore, these tetrads are not consistent with the lack of DMS protection of the guanine N7.

Homologous recombination is not the only biological process where structures stabilized by minor groove tetrads may be involved. Quadruplex formation in DNA sequences containing short nucleotide repeats has been proposed to lead to errors in replication and cause the expansion of the repeat sequence.[79,80] However, the sequences of many repeats are not guanine-rich and cannot form classical quadruplexes. DNA hairpins and slipped structures that occur in sequences containing repeats have been also connected with expansions and deletions of the repeating track.[81] Systematic thermodynamical studies have been carried out to determine the sequence requirements for stable DNA hairpins formation, but their relative stability does not explain the expansion tendency of the different repeats.[57]

We have shown here and in previous studies that a quadruplex formed by minor groove association of Watson−Crick base pairs is a robust DNA motif that does not require guanine-rich sequences. These minor groove tetrads may stabilize the association of DNA hairpins or slipped structures in cellular DNA. The occurrence of such structures in sequences containing repeats within the cell may disturb the DNA metabolism and be responsible of the genetic instabilities associated with DNA repeats and the subsequent disease.[58]

Coordinates. Atomic coordinates have been deposited in the Protein Data Bank (Accession Number PDB: 2HK4; RCSB038436). The complete assignment list has been deposited at the BMRB.

Acknowledgment. This work was supported by the DGICYT Grants CTQ2004-08275-C02-01/02 and Generalitat de Catalunya (2005SGR693 and Centre de Referència en Biotecnologia). We gratefully acknowledge Dr. Doug Laurents and Mrs. Júlia Viladoms for careful reading of the manuscript and their useful comments.

Supporting Information Available: Two tables with the assignment list and average values and order parameters of the dihedral torsion angles, one additional figure, and complete ref 42.

JA066172Z

(76) Gaillard, C.; Strauss, F. *Science* **1994**, *264*, 433−6.
(77) Shukla, A. K.; Roy, K. B. *J. Biochem. (Tokyo)* **2006**, *139*, 35−9.
(78) Shukla, A. K.; Roy, K. B. *Biol. Chem.* **2006**, *387*, 251−6.
(79) Fry, M.; Loeb, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4950−4.
(80) Usdin, K. *Nucleic Acids Res.* **1998**, *26*, 4078−85.
(81) Gacy, A. M.; Goellner, G.; Juranic, N.; Macura, S.; McMurray, C. T. *Cell* **1995**, *81*, 533−40.