NMR Structure of Two Cyclic Oligonucleotides. A Monomer−Dimer Equilibrium between Dumbbell and Quadruplex Structures

Carlos González,*,† Núria Escaja,‡ Manuel Rico,† and Enrique Pedroso‡

Instituto de Estructura de la Materia, CSIC
C/. Serrano 119, 28006 Madrid, Spain
Departament de Química Orgànica
Universitat de Barcelona
C/. Martì i Franquès 1-11, 08028 Barcelona, Spain

Received October 24, 1997
Revised Manuscript Received December 9, 1997

Unusual DNA conformations, distinctive of the classical double helix, have attracted considerable attention. Over the past few years a number of motifs of biological relevance have been found, such as hairpins, triplexes, quadruplexes, etc.[1] Very recently, a new four-stranded motif, named "bi-loop",[2] has been observed for the first time in the crystal structure of two different oligonucleotides. In the first case reported,[3] the linear heptamer d(GCATGCT) folds into a loop structure which self-associates to form a dimer with four G·C Watson−Crick base pairs. An almost identical structure was found by Salisbury et al. in the cyclic octamer d⟨pCATTCATT⟩,[2] where the intermolecular base pairs are A·T instead of G·C. The fact that two apparently unrelated sequences adopt very similar structures suggests that they may be examples of a general motif, whose implications for recognition between DNA molecules have been previously discussed.[2] The occurrence of such structures in solution is of primary importance to assess their possible biological relevance. For this reason, we have undertaken the investigation of the "bi-loop" motif by using NMR techniques. In this paper, we report on the solution structures of two cyclic octamers of sequence d⟨pCATTCATT⟩ and d⟨pTGCTCGCT⟩. To our knowledge, this is the first time that the "bi-loop" motif has been found in solution. In both oligonucleotides, the dimeric "bi-loop" structure is observed in equilibrium with a monomeric dumbbell form.

The cyclic octamers were synthesized as previously reported.[4]

At standard oligonucleotide concentration for NMR studies (∼2 mM) the one-dimensional nonexchangeable proton spectra show different characteristics in both samples (Figure 1).[5] The number and the relative intensities of the signals observed in the aromatic region of d⟨pTGCTCGCT⟩ suggest the presence of two species in a slow exchanging equilibrium on the NMR time scale. These two species could be explained either by a conformational or by an association equilibrium. The latter was confirmed by the dramatic changes observed in spectra recorded at different oligonucleotide concentrations (Figure 1a). At low concentrations,

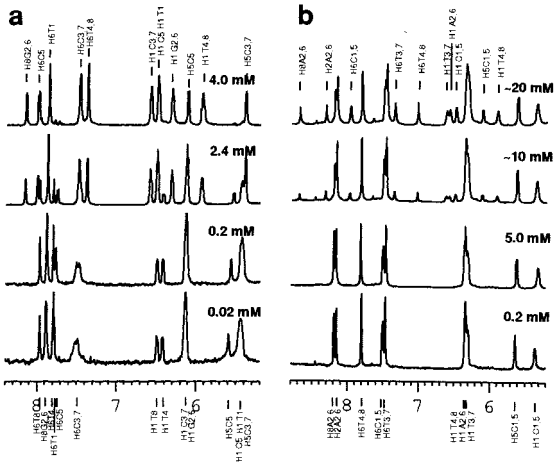

Figure 1. Nonexchangeable proton spectra of d⟨pTGCTCGCT⟩ (a) and d⟨pCATTCATT⟩ (b) at different oligonucleotide concentration (D_2O, T = 20 °C, pH 7.0). Oligonucleotide concentrations are indicated per monomer. Labels on the bottom indicate assignments for the monomer, and, on the top, assignments for the dimeric form.

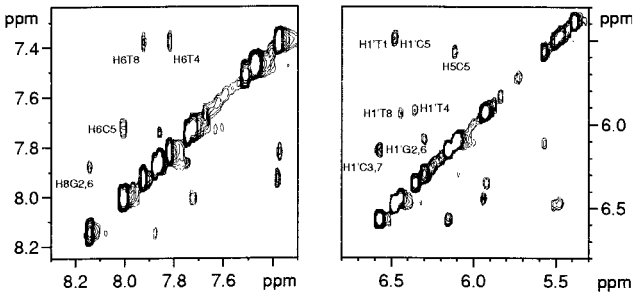

Figure 2. Aromatic and H1′ regions of the ROESY spectrum of d⟨pTGCTCGCT⟩ (D_2O, T = 20 °C, pH 7.0, oligonucleotide concentration of 2.4 mM (per monomer), 300 ms mixing time). All labeled peaks present the same sign as the diagonal and opposite to the cross-relaxation peaks (not shown in these regions).

only one set of signals remains, whereas the other set is the only one observed at high concentrations. The simplest interpretation of this observation is that the two sets of signals correspond to monomeric and dimeric forms of the molecule. For d⟨pCATTCATT⟩, the number of resonances in the aromatic region are those expected for one single species of a symmetric molecule when the concentration is lower than 10 mM (Figure 1b). Interestingly, a second set of signals arises when the concentration is increased, showing that a dimer is also formed, but at higher concentration than in the previous octamer.

The signals corresponding to the two exchanging species could be identified by recording ROESY experiments under conditions where the two forms coexist. Cross-relaxation and exchange cross-peaks were distinguished by their different sign. In Figure 2, two regions of the ROESY spectra for d⟨pTGCTCGCT⟩ are shown, indicating the exchange cross-peaks in the aromatic and H1′ region. Both for d⟨pTGCTCGCT⟩ and d⟨pCATTCATT⟩, cross-relaxation and exchange cross-peaks were readily identified, and a complete assignment of all nonexchangeable resonances in the two conformers could be achieved by using standard homonuclear 2D techniques.[6]

The number of signals observed in the spectra of the dimeric forms is clearly indicative of a symmetric structure. Only four spin systems are found in the dimer of d⟨pCATTCATT⟩ and five in d⟨pTGCTCGCT⟩, with all resonances of residues 1 to 4 and 5 to 8 in d⟨pCATTCATT⟩, and between 2 and 4 and 6 to 8 in

† CSIC.
‡ Universitat de Barcelona.

(1) For recent reviews, see: (a) Lebrun, A.; Lavery, R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 348−354. (b) Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703−730. (c) Rhodes, D.; Giraldo, R. *Curr. Opin. Struct. Biol.* **1995**, *5*, 311−322. (d) Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379−404. (e) Frank-Kamenetskii, M. D.; Mirkin, S. M. *Annu. Rev. Biochem.* **1995**, *64*, 65−95.

(2) Salisbury, S. A.; Wilson, S. E.; Powel, 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*, 5515−5518.

(3) Leonard, G. A.; Zhang, S.; Peterson, M. R.; Harrop, S. J.; Helliwell, J. R.; Cruse, W. B. T.; d'Estaintot, B. L.; Kennard, O.; Brown, T.; Hunter, W. N. *Structure* **1995**, *3*, 335−340.

(4) Alazzouzi, E.; Escaja, N.; Grandas, A.; Pedroso, E. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 1506−1508.

(5) NMR spectra were acquired in a Bruker AMX spectrometer operating at 600 MHz. Spectra at high oligonucleotide concentration were acquired in SIGHEMI tubes, which allow spectra to be recorded in small volume samples under good magnetic field homogeneity conditions. See the Supporting Information for experimental details.

S0002-7863(97)03691-3 CCC: $15.00
Published on Web 02/20/1998

d⟨pTGCTCGCT⟩, completely degenerated.[7] Equivalent residues are indicated between brackets in the following discussion. This degeneracy hinders the assignment of NOESY cross-peaks to specific distance constraints between pairs of protons in the molecule and introduces an additional difficulty in the structural calculations. However, some interesting structural features of the dimeric structure can be readily derived from the observed NOE cross-peaks between nonexchangeable protons. In the two oligonucleotides, intranucleotide H1′ base NOE cross-peaks are medium or weak, indicating that the glycosidic angle in all nucleotides is in the anti conformation. H1′ base sequential connections were observed between residues 2 → 3 → 4 (and the corresponding 6 → 7 → 8), but not between 1 → 2 and 4 → 5 (and the corresponding 5 → 6 and 8 →1). Stacking interaction between residues 2(6) and 3(7) is deduced from the strong sequential H2″ base cross-peaks, but no sequential NOE is observed between H2′/H2″ of residue 3(7) and the base protons of residue 4(8). However, these latter residues are also involved in stacking interactions with the rest of the molecules since strong NOEs are observed between residues 4(8) and residues 2(6) and 3(7). In the monomers, a pattern of NOE cross-peaks characteristic of stacking interactions between the two central bases (2 → 3 and 6 → 7) is clearly observed.

More insight into the structural features of the monomeric and dimeric forms of these molecules can be gained from the exchangeable protons spectra observed in H_2O experiments. Under conditions where the two conformers coexist, two sets of signals for these protons are also observed and can be assigned to each form from 1D spectra acquired at different oligonucleotide concentrations. Sequential assignment of the exchangeable protons in the dimer could be carried out from NOESY spectra in H_2O.[8] Both the monomeric and the dimeric forms present some imino protons at very low field, indicating that Watson−Crick base pairs are formed. In the dimeric structure, the presence of Watson−Crick base pairs is confirmed by the strong cross-peaks observed between the imino protons of the base-paired thymine (or guanine) and the amino protons of the corresponding adenine (or cytosine). An intense cross-peak between the imino proton of thymine and the H2 of adenine is also observed in the A·T base-pair. The arrangement of Watson−Crick base pairs, the pattern of NOE cross-peaks between nonexchangeable protons, and the symmetry of the dimeric form are consistent with an association of two cyclic octamers in an antiparallel manner as indicated in Figure 3a. The base-paired imino protons in the resulting four-stranded structure are well protected, since their signals are narrow and present low exchange rates with the solvent. No significant line broadening is observed, even at temperatures close to the melting transition of the dimer.

The temperature dependence of the labile signals indicates that the melting temperatures of the dimeric forms are higher than the loss of structure on the monomer in both oligonucleotides (Figure 3c and the Supporting Information). For d⟨pTGCTCGCT⟩, the imino signal of H1G2(6) in the dimer is observed up to 60 °C, while the equivalent signal in the monomeric form disappears

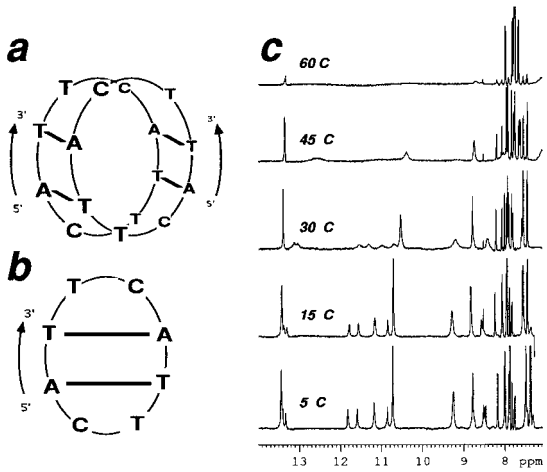

Figure 3. (a) Schematic diagram of the dimeric structure of d⟨pCATTCATT⟩. (b) Scheme of the monomeric structure. (c) 1D exchangeable protons spectra at different temperatures for the oligonucleotide d⟨pTGCTCGCT⟩ at a concentration of 2.4 mM (per monomer).

at 45 °C. In d⟨pCATTCATT⟩, the melting temperatures of the two forms are about 45 and 35 °C, respectively. Chemical shifts of the exchangeable and nonexchangeable signals of the dimer remain constant with temperature, while the corresponding signals of the monomer vary as the temperature increases. This indicates a different melting behavior of the two species. Coalescence between the signals of the two forms is not achieved before the melting of the two structures, indicating that the equilibrium between both species is slow on the NMR time scale over the entire range of temperatures where the two forms coexist.

In the two sequences studied, all available experimental information is consistent with a dimeric structure in slow equilibrium with a monomeric form. The monomer adopts a dumbbell-like conformation, with four of the bases involved in two intramolecular Watson−Crick base pairs and the other four forming two mini-hairpin loops of two bases each (Figure 3b). From this result, it appears that dumbbells with two-base hairpins can be formed from cyclic oligonucleotides of only eight residues, such hairpins have been more commonly reported in larger cyclic oligonucleotides.[9] The dimer is a four-stranded symmetric structure with four intermolecular Watson−Crick base pairs (Figure 3a). It can be concluded from the present data that the dimer in solution is very similar to the "bi-loop" found in solid state,[2,3] although the structural calculations, currently in progress, would provide a more detailed view of this structure. The fact that this new motif has been now observed in three different sequences, both in cyclic and linear oligonucleotides, strongly supports the hypothesis that we may be observing a motif of general occurrence in nature. The existence of such motif in solution encourages the investigation of these structures in living organisms.

Acknowledgment. This work was supported by the DGICYT grants PB96-0885, PB94-0844, and Generalitat de Catalunya (Centre de Referència de Biotecnologia). We gratefully acknowledge Dr. M. Bruix, Dr. S. Padmanabhan, Dr. A. Grandas, and Dr. S. A. Salisbury for many helpful comments and suggestions.

Supporting Information Available: Assignment tables of the dimeric and monomeric forms of d⟨pCATTCATT⟩ and d⟨pTGCTCGCT⟩, 1D exchangeable protons spectra at different temperatures for d⟨pCATTCATT⟩ at 20 mM oligonucleotide concentration (per monomer) (3 pages). See any current masthead page for ordering information and Web access instructions.

JA973691G

(6) Spin systems in the sugar moiety were identified in the TOCSY and COSY spectra. Almost all resonances, including H4′ and H5′ and H5″, were identified but no stereospecific assignment of the H5′/H5″ could be done. Base protons of cytosines and thymines were identified by the H5−H6 and Me−H6 cross-peaks in the TOCSY spectra and connected to their sugar spin system by the H6−H1′ cross-peak.

(7) This observation is especially surprising in the case of d⟨pTGCTCGCT⟩, where the sequence is not symmetric. Only the H3′ protons of T4 and T8 show slightly different chemical shifts.

(8) Only protons involved in base pairs gave rise to observable cross-peaks. In the dimeric form of d⟨pCATTCATT⟩, the amino protons of A2(6) were assigned by their cross-peaks with the H3 imino proton of T3(7). Similarly, in d⟨pTGCTCGCT⟩ the labile protons of the dimer could be assigned by following the connection from H5 of the C3(7) (H5C → HN4C → H1G → HN2G) between the G·C base pair.

(9) (a) Rinkel, L. J.; Tinoco, I. *Nucleic Acids. Res.* **1991**, *19*, 3695−3700. (b) Ippel, J. H.; Lanzotti, V.; Galeone, A.; Mayol, L.; Boogaart, J. E. v. d.; Pikkemaat, J. A.; Altona, C. *J. Biomol. NMR* **1995**, *6*, 403−422 and references cited therein.