NMR studies on truncated sequences of human telomeric DNA: observation of a novel A-tetrad

Prasanta K. Patel, A. S. R. Koti and R. V. Hosur*

Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India

Received June 14, 1999; Revised August 10, 1999; Accepted August 20, 1999

ABSTRACT

The structure of the telomeric DNA has been a subject of extensive investigation in recent years due to the realization that it has important functional roles to play *in vivo* and the observations that truncated telomeric sequences exhibit a great variety of 3D structures in aqueous solutions. In this context, we describe here NMR structural studies on two truncated human telomeric DNA sequences, d-AG_3T and d-TAG_3T in solutions containing K^+ ions. The G_3 stretches in both the oligonucleotides were seen to form parallel-stranded quadruplexes. However, the AG_3 segment as a whole, had different structural characteristics. The structure of d-AG_3T revealed the formation of a novel A-tetrad, which was not seen in d-TAG_3T. The A's in the tetrad had *syn* glycosidic conformation as opposed to the *anti* conformation of the G's in the G-tetrads. The A-tetrad stacked well over the adjacent G-tetrad and the twist angle at this step was smaller in d-AG_3T than in d-TAG_3T. These observations are expected to be significant from the point of view of structural diversity and recognition in telomeres.

INTRODUCTION

Telomeres are specialised nucleo-protein structures found at the ends of linear eukaryotic chromosomes, which are reported to be crucial for various biological functions such as aging, chromosomal organization and stability, etc. (reviewed in 1–4). They also provide a mechanism for complete replication of the 3' termini without ablation (5,6; reviewed in 3,4). Recently it has been shown that telomeres may function as regulators for gene expression (7). The telomeric DNA contains a long stretch of duplex DNA (hundreds to thousands of base pairs) with evolutionarily conserved tandem repeats of short G-rich sequences (6–10 nucleotides in each repeat) in one strand (in the 5' to 3' direction) and their complementary sequences in the other strand (reviewed in 1–4). The G-rich strand contains about two to three repeats of a single-strand overhang at the 3' end except in the case of telomeres in human somatic cells, which have 20–30 repeats. A unique feature of this 3' overhang is its exceptional resistance to nuclease degradation (8). It has been reported that the length of the telomeric DNA is related to aging and diseases like cancers (9–11; reviewed in 12).

X-ray crystallographic and solution NMR studies (13–32, reviewed in 33–35) on different synthetic oligonucleotides related to above telomere repeats show that they can form G-quadruplex structures by association of four G-stretch sequences. Such associations could be either intra-molecular or inter-molecular in type. In intra-molecular structures, a single nucleotide chain having four stretches of G-nucleotides separated by $T_{2-4}A_{0-1}$ stretches folds into a unimolecular quadruplex (16–20). Here the chain folds three times to bring the four stretches of G-nucleotides in proper alignment in a plane to form tetrads where as the $T_{2-4}A_{0-1}$ stretches form the loops. An inter-molecular quadruplex is formed by association of two hairpin-shaped strands having two stretches of Guanine sequences each (21–25) or by linear association of four different strands forming either a parallel or an antiparallel quadruplex depending on the orientations of the four strands (26–32). The telomeric sequences such as T_2AG_3, T_2G_4, TG_4T, T_2G_4T, TG_4T, TG_3T and T_4G_4 were seen to form parallel-stranded G-quadruplexes with all the nucleotides having *anti* glycosidic torsion angles (26–32). In dimeric hairpin and unimolecular quadruplexes, the alignments of the four strands are necessarily of antiparallel type and the G-nucleotides have combinations of *syn* and *anti* glycosidic conformations (16–25). The *syn–anti* distribution of the G residues around the individual G-tetrads and along the individual strands is defined by the overall topology of the molecule, which, in turn, is dependent on the loop connectivities. For example, the sequence d-$AG_3(T_2AG_3)_3$ formed an antiparallel quadruplex in which each strand had one parallel and one antiparallel neighbour (16). The glycosidic torsion angles for the G's altered between *syn* and *anti* in each of the strands and in each tetrad, G(*syn*)-G(*syn*)-G(*anti*)-G(*anti*) pattern was observed. The *Oxytricha* sequence $G_4(T_4G_4)_3$ also formed a similar quadruplex with T4 loops in solution (17,18) but in the crystalline state, the G-tetrad had G(*syn*)-G(*anti*)-G(*syn*)-G(*anti*) alignment and also different loop orientations (13). Whether a linear or a folded structure is formed is dependent on the exact sequence and number of repeats per strand, and cations have a strong influence on their relative stabilities (35–38). It has been found that potassium ions preferably stabilize linear, four-stranded parallel quadruplexes whereas sodium ions stabilize folded forms (38).

We report here the formation and characterization of a novel K^+-stabilized AG_3 quadruplex in the sequence d-AG_3T. This is a parallel-stranded quadruplex in which an A-tetrad is formed at the 5' end. Such a feature is not seen in the d-TAG_3T sequence where the AG_3 stretch is flanked by T nucleotides at the 5' end. We observed that the G_3 parts had overall similar

*To whom correspondence should be addressed. Tel: +91 22 215 2971; Fax: +91 22 215 2110; Email: hosur@tifr.res.in

characteristics in both the molecules, but in d-AG_3T, the A's had *syn* conformations and the A-tetrad stacked well over the adjacent G-tetrad. The data also indicated a dynamism in the alignments of A's in the A-tetrad.

MATERIALS AND METHODS

DNA synthesis

The oligonucleotides were synthesized on an Applied Biosystems 392 automated DNA-synthesizer on 10 μM scale using solid phase β-cyanoethyl phosphoramidite chemistry, cleaved from the support and purified by standard procedures (39,40). NMR samples of the oligonucleotides were prepared at a concentration range of 0.5–4 mM strand concentration in 0.6 ml (90% H_2O/10% D_2O) buffer solutions having 2 mM potassium phosphate, 0.2 mM EDTA, pH 7.3, 2–100 mM KCl. For D_2O experiments, the samples were converted into D_2O samples by repeated lyophilization and redissolution in D_2O alone.

NMR data collection and processing

NMR data were obtained either on a VARIAN UNITY-plus 600 or a BRUKER AMX 500 spectrometer. Temperature-dependent 1D spectra (0–70°C) in H_2O were recorded using jump-and-return pulse sequence (41) for water suppression. Phase-sensitive NOESY spectra (42) in H_2O were recorded with mixing times of 60, 100, 180, 250 and 350 ms in the temperature range of 7–20°C. Here again water suppression was achieved by jump-and-return pulse sequence. Phase-sensitive NOESY spectra in D_2O were recorded with different mixing times (40, 80, 120, 200, 300 and 400 ms for d-AG_3T; 50, 100, 200, 300 and 400 ms for d-TAG_3T) at 20°C. TOCSY spectra (43) with mixing times of 20, 80 and 120 ms for both the molecules were recorded with D_2O solutions. ECOSY (44) spectra of the molecules were recorded with D_2O sample for ^1H–^1H coupling constant estimation in the sugar rings. The experiments on the VARIAN Unity plus spectrometer used States–Haberkorn method (45) for quadrature detection, while those on BRUKER used TPPI procedure (46) for the same. In all the 2D experiments, the time domain data consisted of 2048 complex points in t_2 and 400–600 fids (free induction decay signals) in t_1 dimension. The relaxation delay was kept at 1.5 s for all the experiments. The VARIAN data were processed using Felix-230 on an IRIS workstation. The data were apodized by shifted (60–90°) sine bell functions prior to 2D Fourier transformations.

Assignments

Sequence-specific resonance assignments for d-AG_3T and d-TAG_3T were obtained by using NOESY and TOCSY spectra following standard procedures (47,48). We could trace the self as well as sequential NOE connectivities from H8 to H1' and H2'/H2" protons for all the nucleotides and, in addition, from H8 to H3' protons for d-AG_3T. The connectivities followed the standard patterns as for B-DNA duplexes and complete assignments of base and sugar protons for both the nucleotide sequences were obtained.

Model building

Initial models of the quadruplex structures for the two sequences d-AG_3T and d-TAG_3T were generated by using the INSIGHT II program, version 3.1 (Biosym Technologies Inc.) on an IRIS work station. First, a G-tetrad was made by aligning four bases symmetrically in a plane to form appropriate H-bonds with known bonding scheme. The G_3-quadruplex was made by stacking three such G-tetrads on each other assuming B-DNA geometries with a plane separation of ~3.4 Å and a twist angle of 36°. The 5' and 3' ends of the different tetrads were joined. To this model, A and T nucleotides were joined at the 5' and 3' ends respectively for building the d-AG_3T quadruplex. Here the A1 nucleotides were kept in *syn* glycosidic angle. Similarly, the d-TAG_3T quadruplex was built by joining extra T nucleotides at the 5' ends of the d-AG_3T model. However, here the A nucleotide was kept in *anti* glycosidic conformation. The 5' ends of both the quadruplex models were dephosphorylated and all ends were capped by hydrogens. These molecules were then subjected to energy minimization by steepest descent followed by conjugate gradient methods by DISCOVER Program, version 3.1 using AMBER force field, for few steps to obtain the proper geometries. During the energy minimization, H-bond constraints involving G-tetrads were used with a force constant of 100 kcal mol^{-1} $Å^{-2}$.

NOE intensities and experimental constraints

Intensities of the NOE cross peaks contain the information about interproton distances. All the cross peaks in NOESY spectra recorded with different mixing times for d-AG_3T and d-TAG_3T were integrated several times and averaged to avoid manual integration errors. On the basis of the intensities in the short mixing time NOESY spectra (80 and 100 ms for d-AG_3T and d-TAG_3T respectively), the interproton distances were classified into three groups; 1.8–3.0, 2.5–4.0 and 3.5–5.0 Å. There were a total of 444 constraints (111 per strand) for d-AG_3T and 332 constraints (83 per strand) for d-TAG_3T. We used 48 H-bond distance constraints for the G-tetrads for each molecule. The other interproton distances involving exchangeable protons were constrained to the range of 2.5–5.5 Å. There were 20 and 12 such constraints respectively for d-AG_3T and d-TAG_3T. These constraints were used during restrained energy minimizations and structure calculations by simulated annealing as described below.

Structure calculations

Structures of d-AG_3T and d-TAG_3T quadruplexes were calculated by simulated annealing-restrained molecular dynamics protocol of DISCOVER Program, version 3.1, using the AMBER force field. The initial models were heated to 1000 K and 200 different structures were saved at regular intervals during an equilibration dynamics run of 100 ps. These were seen to be quite different from each other as judged from their pairwise r.m.s.d.s. Each of them was then slowly cooled to 293 K in 50 K steps with force constant of 20 kcal mol^{-1} $Å^{-2}$ for all the interproton distances except for H-bonds for which 100 kcal mol^{-1} $Å^{-2}$ was used. During cooling, at each step, 150 fs dynamics was done for each structure. At 293 K, dynamics was run for 4 ps and an average structure was taken from the last 2 ps dynamics; this was energy minimized by steepest descent and conjugate gradient methods until a predefined convergence limit (r.m.s. derivative <0.001) was reached. The structures were finally selected on the basis of constraint satisfaction, symmetry and low energies.

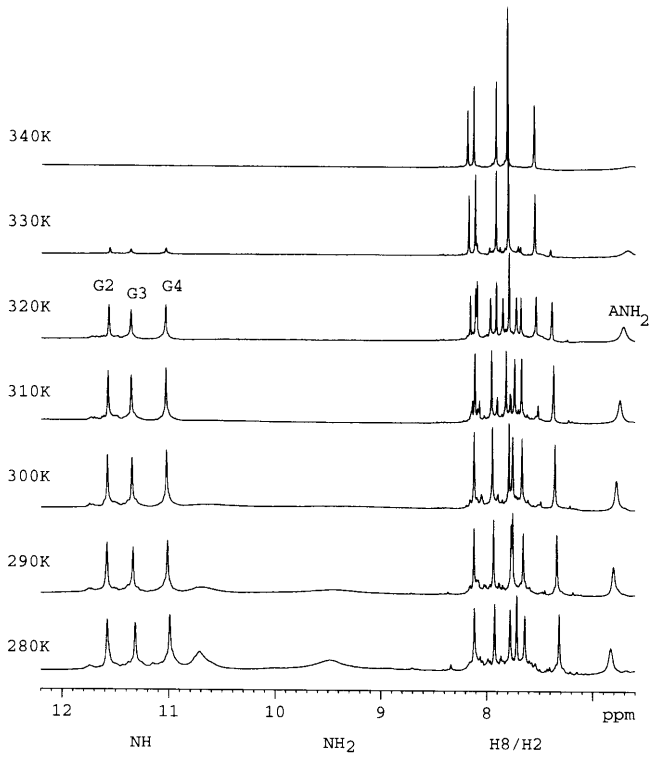

Figure 1. Imino, amino and base proton spectra of d-AG_3T in H_2O solutions containing 5 mM K^+ ions at pH 7.3 as a function of temperature. The G-imino and A-amino proton assignments have been marked in one of the spectra.

RESULTS AND DISCUSSION

NMR spectral features

Figure 1 shows temperature dependence of the exchangeable imino, amino and non-exchangeable base proton NMR spectra of d-AG_3T in H_2O solutions. At low temperature, one set of major resonances belonging to a major conformation is seen. As the temperature is increased, duplication of base proton peaks is observed. The intensities of the new peaks increase with increasing temperature whereas it is the reverse for the major set of peaks seen at low temperature. This indicates melting of the structure seen at low temperature into single strands. The three distinct exchangeable peaks in the region of 10.0–11.5 p.p.m., remain up to a very high temperature. Hydrogen–deuterium exchange studies indicated that two (outer ones) of these three imino protons were more easily accessible to the solvent while the central peak was well protected. The above features, namely imino peaks resonating around 10.0–11.5 p.p.m., high melting temperature for the imino resonances and low solvent accessibility of the imino peaks, are characteristics of G-quadruplex structures, as has been well documented in the literature (16–32). The peaks resonating at around 10.0–11.5 p.p.m. correspond to G-imino protons in the G-tetrads. The peak count in the base region as well as of the G-iminos for the quadruplex structure indicate that the structure formed in d-AG_3T is highly symmetrical and the four strands are equivalent. Similar features were also seen for d-TAG_3T.

Resonance assignments

The 2D NOESY spectra of the two molecules showed well dispersed cross peaks, with good intensities over a whole range of mixing times (40–400 ms). Assignments of all the exchangeable and non-exchangeable protons in the oligonucleotides were obtained by standard procedures based on NOESY and TOCSY spectra (47,48). Illustrative H8-H1' connectivities are shown for d-AG_3T and d-TAG_3T in Figure 2. The intra-residue H2' and H2" were distinguished on the basis of strong H1' to H2" and weak H1' to H2' intra-nucleotide cross peaks in low mixing time NOESY spectra. Sequential NOE connectivities were also observed from G-imino to GH8 protons of the 5'-flanking nucleotide (Fig. 3) and also from each G-imino to

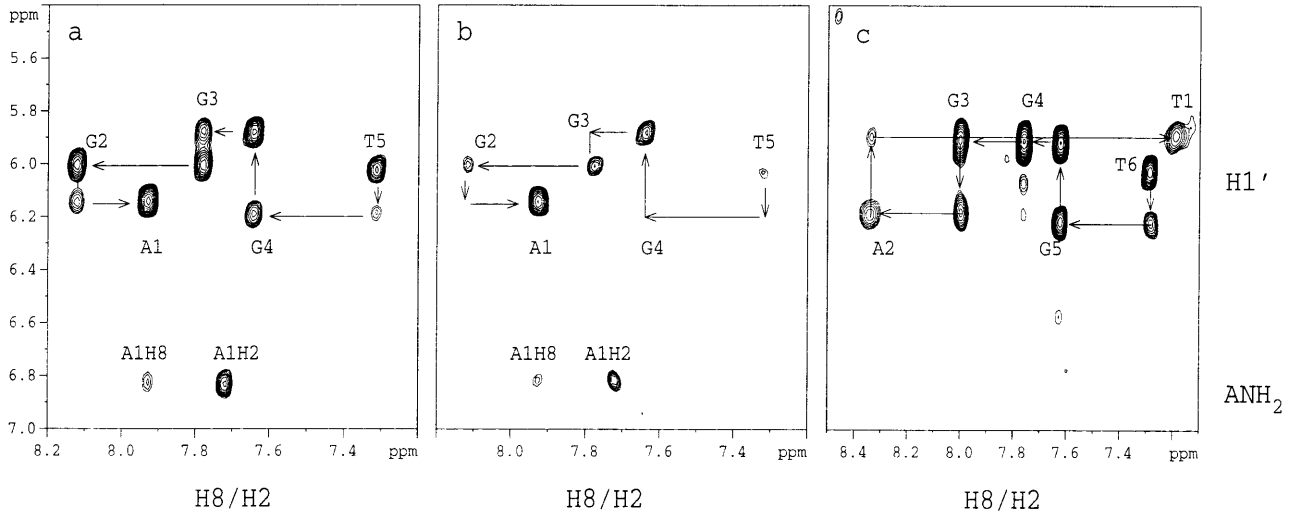

Figure 2. Selected spectral regions from NOESY spectra at 7°C in H_2O, pH 7.3 of (**a**) d-AG_3T (180 ms mixing time), (**b**) d-AG_3T (60 ms mixing time) and (**c**) d-TAG_3T (250 ms mixing time). In (a) and (b), cross peaks from A1NH_2 to A1H8 and A1H2 protons are also seen. In d-TAG_3T the corresponding region is empty indicating the absence of the A-tetrad type features; in fact, the A-amino resonances were not seen in this molecule.

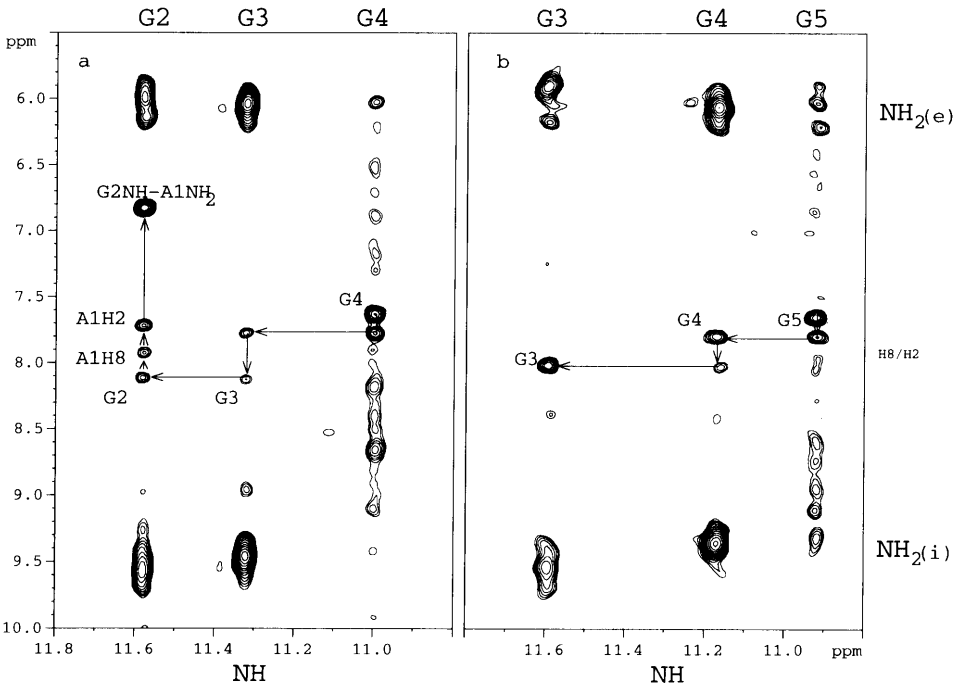

Figure 3. Self and sequential NOE cross peaks between imino and base (H8/H2) protons in d-AG_3T (**a**) and d-TAG_3T (**b**). The spectral regions have been taken from the same NOESY spectra as in Figure 2a and c. The sequential connectivities have been drawn in. In d-AG_3T the connectivities extend up to A1 while in d-TAG_3T they stop at the adjacent G3 nucleotide itself. Also strong cross peaks are seen from G2NH to A1H2 and $A1NH_2$ protons in d-AG_3T molecule. The spectra also show good cross peaks from imino to amino protons for two of the three G-nucleotides in both the molecules. The G-aminos adjacent to the terminal T are too broad as is seen by the long tail-like structures in the two spectra.

Table 1. Chemical shift[a] values of d-AG_3T and d-TAG_3T quadruplex at 20°C

d-AG_3T	H8/H6	H1'	H2'/2"	H3'	H4'	H5'/5"	H2/Me	NH_2	NH	d-TAG_3T
	7.24	5.94	2.92,1.81	4.62	4.00	3.54	1.60			T1
A1	7.98	6.20	2.66	4.93	4.25	3.76,3.72	7.81	6.80		
	8.37	6.25	2.91,2.86	5.06	4.41	4.10,4.00	8.03			A2
G2	8.16	6.08	3.06,2.83	5.08	4.52	4.20		9.61,6.10	11.64	
	8.07	6.00	2.97,2.71	5.06	4.51	4.25		9.51,5.96	11.66	G3
G3	7.80	5.97	2.69	5.05	4.52	4.33		9.53,6.09	11.38	
	7.81	5.99	2.70	5.06	4.50	4.31		9.38,6.10	11.25	G4
G4	7.70	6.27	2.70,2.56	4.91	4.52	4.29			11.05	
	7.71	6.28	2.72,2.58	4.92	4.53	4.29			11.00	G5
T5	7.38	6.09	2.19	4.50	4.07	4.23,4.10	1.65			
	7.38	6.09	2.20	4.50	4.08	4.10,4.23	1.63			T6

[a]With respect to sodium 3-trimethyl silyl-(2,2,3,3-$^2H_4$) propionate.

adjacent G-imino protons. The peak at 6.8 p.p.m. (Fig. 1) was identified to be an exchangeable proton peak on the basis of deuterium exchange. This peak was assigned to ANH_2 proton on the basis of its NOE connectivities (Figs 2 and 3). Complete assignments of the peaks are listed in Table 1.

Quadruplexes and strand orientations

The observation of only self $G_iNH_{2\ (int)}$ to G_iH8 and not any of the inter-strand inter-nucleotide cross peaks of the type: $G2NH_{2(int)}$-G4H8, $G4NH_{2(int)}$-G2H8 or G4NH to G2H8, G2NH to G4H8 in d-AG_3T, for instance, indicated that strand orientations were parallel in the quadruplex; the latter peaks must be expected for an antiparallel quadruplex structure. Furthermore, for an antiparallel tetraplex formed by linear association of four strands, combinations of both *syn* and *anti* glycosidic conformations for the G-nucleotides are expected as a geometrical requirement. In all the sequences studied here, the G-nucleotides were found to be in *anti* glycosidic conformation,

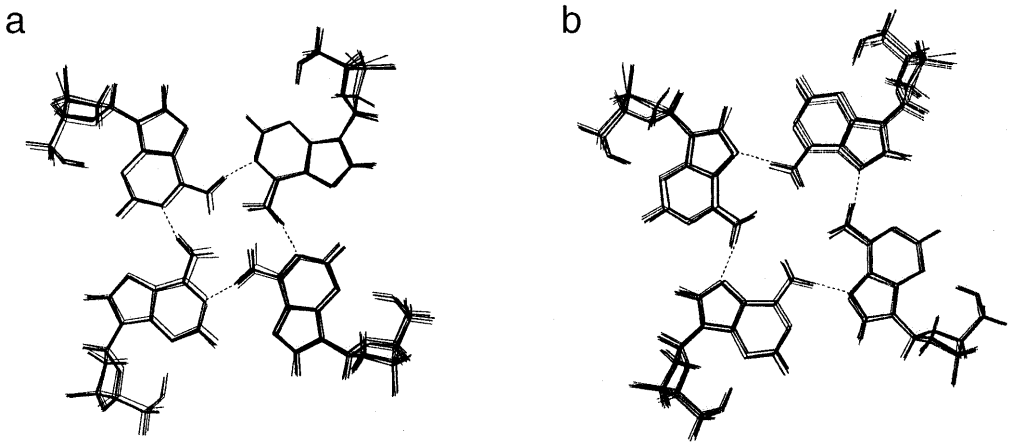

Figure 4. Different alignments of A nucleotides seen after the structure calculation. They have been labeled as N61 (**a**) and N67 (**b**) depending upon the nitrogen atom on A accepting the amino hydrogen in each of the H-bonds. The H-bonds are indicated by dotted lines.

which rules out the formation of antiparallel quadruplex structure.

The sequential NOE connectivity patterns between base and sugar protons indicate that each strand has a right-handed helical geometry. 2D TOCSY and ECOSY spectral patterns (data not shown) indicated that H1'-H2' coupling constants were reasonably large (>6 Hz), good cross peaks were seen even at very short TOCSY mixing time (20 ms). This observation fixes the sugar geometries to be predominantly S-type (48) and consequently the strand structures may be taken to belong to the B-DNA family.

Distinctive features between d-AG_3T and d-TAG_3T

Glycosidic conformation of A. At low mixing time (60 ms), self H8-H1' NOE cross peak was seen only for A in d-AG_3T (Fig. 2b). It is also seen in Figure 2a that the A1H8-A1H1' cross peak is highly intense compared to all the other peaks. This indicates that A in d-AG_3T has a *syn* glycosidic conformation. On the other hand, A2 in d-TAG_3T has an *anti* glycosidic conformation (Fig. 2c).

Interstrand interactions of A. The NOESY spectra of d-AG_3T under neutral pH conditions showed intense cross peaks from $A1NH_2$ to A1H8, A1H2 protons (Fig. 2a). In an A-base, the shorter of the two distances from NH_2 to H8 and also from NH_2 to H2 is >4.5 Å and hence one would not expect to observe these NOE cross peaks. In the present case, we observed these peaks even at a mixing time of 60 ms (Fig. 2b). From rough quantification, the intensities of the two cross peaks are comparable to some of the intra-nucleotide GH8-GH1' cross peaks. In fact the $A1NH_2$-A1H2 cross peak is more intense than some of the self H8-H1' peaks. Thus, it appears that the above peaks arise due to an inter-strand interaction in a structure where the A-bases come close and arrange in a symmetrical fashion. Furthermore, the above cross peaks persisted even after 3–4-fold dilution of the samples and their chemical shift values also did not change. This ruled out end-to-end stacking of two quadruplexes. In this context, supporting evidence for A–A interaction comes by noting that the ANH_2 resonance is highly stable with respect to temperature and disappears almost synchronously with the GNH resonances (Fig. 1).

Quadruplex structures in d-AG_3T and d-TAG_3T. The quadruplex structures for the two molecules were calculated by simulated annealing-restrained molecular dynamics as described in Materials and Methods. Although the NMR spectra showed indications of interstrand A–A interactions in d-AG_3T, we did not impose any of these constraints so as to avoid any bias in the nature of the interactions. The structure convergence parameters for the average structures are listed in Table 2. At the end of the above calculations for d-AG_3T quadruplex, the final structures were seen to be distributed between two distinct patterns of A alignment (Fig. 4). These we termed as A-tetrads, N61 and N67 in analogy with the G-tetrad, since we noticed that the A-amino hydrogens were in H-bonding distance range from the receptor nitrogens N1 and N7 respectively and also were appropriately oriented for H-bond formation. This enabled the identification of the nature of A–A interactions mentioned above and also provided a basis for the high temperature stability of ANH_2 resonance and $A1NH_2$-A1H8 (N67 model) and $A1NH_2$-A1H2 (N61 model) NOEs described above. Simultaneous observation of these NOEs also suggests a dynamism (discussed below) in the A-tetrad. Therefore, we did not attempt to refine the above structures by relaxation matrix calculations, as there is no way to estimate the contribution of the two models to the NOE intensities at the A1–G2 step.

Figure 5a and 5b shows superposition of eight structures having low energy values for the two AG_3 quadruplexes having N61 and N67 A-tetrad alignments respectively. Figure 5c shows the structure of the AG_3 quadruplex in d-TAG_3T molecule. It is seen that the structures are better defined in d-AG_3T than in d-TAG_3T, which is a consequence of stronger and larger number of NOEs in the former molecule. Nonetheless, the G_3 stretch in d-TAG_3T is well defined to enable comparisons. The structural differences at the A–G steps of the two molecules also reflected in the different patterns of H8 chemical shifts of A1 and G2 nucleotide units in d-AG_3T and A2 and G3 nucleotide

Table 2. NMR restraints and convergence statistics

	d-AG_3T	d-TAG_3T
1. Total NOEs		
Non-exchangeable	121	90
Inter-residue	33	18
Intra-residue	88	72
Exchangeable	26	18
Inter-residue	1	07
Intra-residue	15	11
2. Distance constraints used for structure calculations		
Exchangeable (for four strands including H-bonds)	68	60
Non-Exchangeable (per strand)		
Intra-residue	81	64
Inter-residue	30	19
3. Chosen convergent structures	08	08
r.m.s.d.s from the average structure[a]		
N61	0.28–0.40	
N67	0.28–0.40	
d-TAG_3T	0.50–1.12	
4. Violations (in AG_3 segment)		
Distance violation >0.05 Å	Nil	Nil

[a]Eight structures were used for r.m.s.d. analysis.

units in d-TAG_3T spectra (Fig. 2). The average structures of the two molecules are analyzed below for some details.

The pitch values between adjacent tetrads are ~3.2 Å for both d-AG_3T and d-TAG_3T quadruplexes. The twists angles at successive G–G, G–G and G–T steps in the GGGT-3' stretches of the two molecules are similar, being 22°, 30° and 40°. This results in different extents of stacking of the adjacent tetrads. For A1–G2 steps, the twist angles are 11° and 18° for N61 and N67 pairing schemes respectively. In d-TAG_3T, the corresponding value is 22°.

The structures of the G_3 segments in the quadruplexes d-AG_3T and d-TAG_3T are largely similar. Base stacking at all the steps in the two molecules is mostly intrastrand, like in B-DNA. At the G2–G3 step, the imino and amino protons of G2 nucleotides stack on the six-membered ring of G3. But in the next, G3–G4 step, the base stacking is rather poor. At the G4–T5 step there is again a good overlap of the bases.

Figure 6a and 6b shows stacking of the A-tetrads over the adjacent G-tetrads in the two AG_3 quadruplexes. In the N61 pairing scheme, there is a good stacking between the five-membered rings of A1 and G2 bases within the same strand. Similar kinds of stacking between adjacent G-tetrads were reported earlier in antiparallel tetraplexes, which also had *syn* conformations for some of the G's (13; reviewed in 35). The H-bonded ANH_2 stacks on top of the G2NH from the adjacent strand. In the N67 pairing scheme, the rings themselves do not stack over each other but the A1H8 proton stacks over the five-membered ring of the G2-nucleotide of the same strand.

In both the final structures of the d-AG_3T quadruplex, we noticed that the 5' hydroxyl of A1 forms a H-bond to its own N3 atom. This could be a driving force for the A's to take a *syn* conformation and form the A-tetrad. In d-TAG_3T, there is no free hydroxyl proton to form such a H-bond.

Dynamism in the A-tetrad. As mentioned earlier, simultaneous observation of $A1NH_2$-A1H8 and $A1NH_2$-A1H2 NOEs and also the fact that only one resonance is seen for each proton in d-AG_3T despite the possibility of two modes of A-tetrad formation having different stacking patterns indicates the existence of dynamism in the structure. A rapid equilibrium between the two modes accounts for the chemical shift of 6.8 p.p.m., for both the $A1NH_2$ protons, since both of them would be exchanging between a H-bonded state (8.0–9.5 p.p.m.) and a free state (5.0–6.5 p.p.m.). The exchange rate has to be fast compared to the chemical shift differences between the two models, but slower than the NOE time scales. Since NOE time scales are in a nanosecond range and chemical shift averaging is in a millisecond time scale, this dynamism is easily conceivable. We may point out that a transition from the N61 to the N67 model and vice-versa, can be easily brought about by few concerted rotations at the A1–G2 step.

CONCLUSIONS

We have described in this paper NMR investigations in K^+-containing solutions, on two DNA sequences, d-AG_3T and

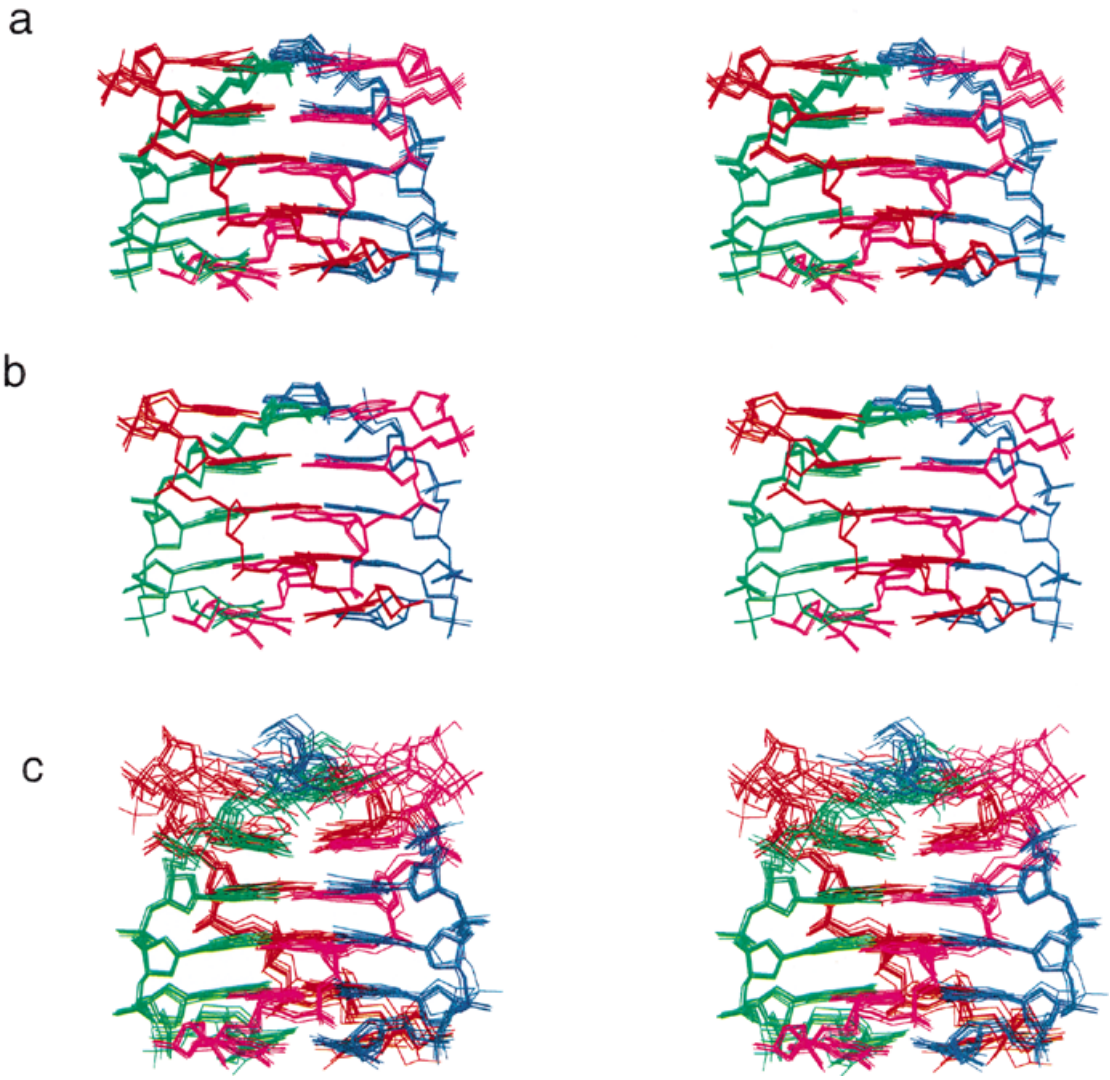

Figure 5. Stereo views for the superposition of eight structures each of d-AG_3T [(**a**) and (**b**) are for N61 and N67 models respectively] and d-TAG_3T (**c**) molecules.

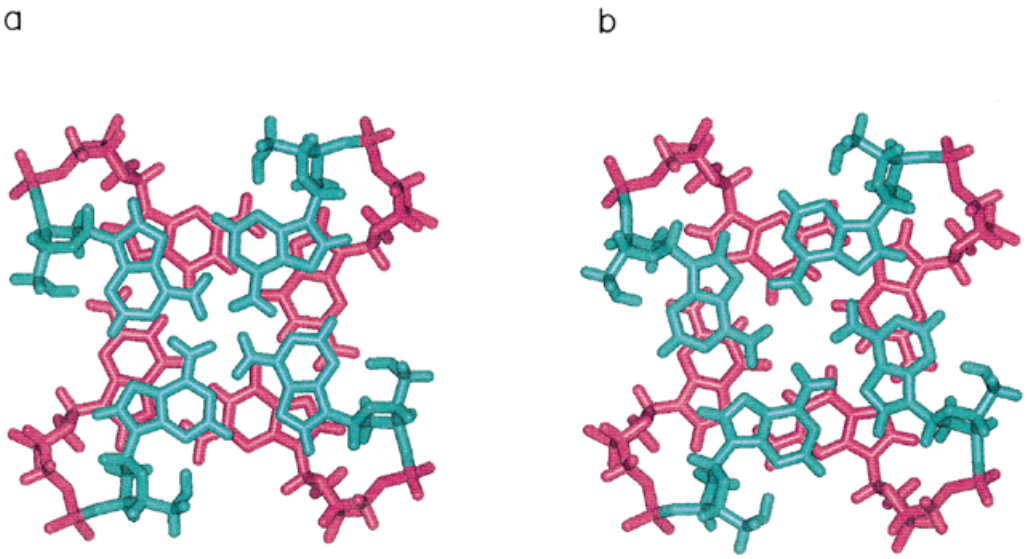

Figure 6. Stacking of A1-tetrad (cyan) over G2-tetrad (pink) in N61 (**a**) and N67 (**b**) models in the final average structures of the two quadruplexes.

d-TAG$_3$T, corresponding to different truncations of the human telomeric sequence repeats $(T_2AG_3)_n$. This is of significance since the length of the repeat overhang at the 3' end of the chromosome is variable and may consist of fractional repeats as well. The study has led to some important conclusions. The truncation position was seen to have a significant effect on the structure of the DNA. Even a single base difference made a remarkable difference in the behaviors of the molecules in aqueous solutions. In d-AG$_3$T, the 5' terminal A formed a novel A-tetrad, in which four A's were held together by H-bonds in a plane and each of the A's was in a *syn* glycosidic conformation. In contrast, in d-TAG$_3$T, such a structure was not observed.

Our observation of an A-tetrad in the quadruplex structure of d-AG$_3$T is unprecedented and indicates that, like the G-nucleotide, the A-nucleotide too has the potential of generating a variety of structures. In contrast to the earlier observations that all the nucleotides adopt *anti* glycosidic angle in a parallel-stranded tetraplex, we see a combination of *syn* (for A1) and *anti* (for G's) glycosidic conformations in the d-AG$_3$T quadruplex. The A-tetrad adds another member to the list of new base platforms such as T-tetrad (49), G:C:G:C-tetrad (50–52) and U-tetrad (53). Observation of A-tetrad is significant since A-rich sequences are known to be widely present in the DNA of both prokaryotic and eukaryotic chromosomes and thus their structural diversity would be expected to contribute significantly to different functions inside the cell.

ACKNOWLEDGEMENT

The facilities provided by the National Facility for High Field NMR at Tata Institute Fundamental Research are gratefully acknowledged.

REFERENCES

1. Zakian,V.A. (1995) *Science*, **270**, 1601–1604.
2. Zakian,V.A. (1989) *Annu. Rev. Genet.*, **23**, 579–604.
3. Blackburn,E.H. (1990) *J. Biol. Chem.*, **265**, 5919–5921.
4. Blackburn,E.H. and Greider C.W. (eds), (1995) *Telomeres.* Cold Spring Harbor Laboratory, Plainview, NY.
5. Greider,C.W. and Blackburn,E.H. (1989) *Nature*, **337**, 331–337.
6. Shipen-Lentz,D. and Blackburn,E.H. (1990) *Science*, **247**, 546–552.
7. Raghuraman,M.K., Brewer,B.J. and Fragman,B.J. (1997) *Science*, **276**, 806–812.
8. Blackburn,E.H. (1986) In Bogorad,L. (ed.), *Molecular Developmental Biology*. Alan R. Liss, Inc., New York, pp. 69–82.
9. Cooke,H.J. and Smith,B.A. (1986) *Cold Spring Harbor Symp. Quant. Biol*., **51**, 213–219.
10. Harley,C.B., Futcher,A.B. and Greider,C.W. (1990) *Nature*, **345**, 458–460.
11. Hastie,N.D. *et al.* (1990) *Nature*, **346**, 866–868.
12. Blackburn,E.H. (1991) *Nature*, **350**, 569–573.
13. Kang,C.H., Zhang,X., Ratcliff,R., Moyzis,R. and Rich,A. (1992) *Nature*, **356**, 126–131.
14. Laughlan,G., Murchie,A.I.H., Norman,G., Moore,M.H., Moody,P.L.E., Lilley,D.M.J. and Luisi,B. (1994) *Science*, **265**, 520–523.
15. Phillips,K., Dauter,Z., Murchie,A.I., Lilley,D.M. and Luisi,B. (1997) *J. Mol. Biol.*, **273**, 171–182.
16. Wang,Y. and Patel,D.J. (1993) *Structure*, **1**, 263–282.
17. Wang,Y. and Patel,D.J. (1995) *J. Mol. Biol.*, **251**, 76–94.
18. Smith,F.W., Schultze,P. and Feigon,J. (1995) *Structure*, **3**, 997–108.
19. Wang,K.Y., Swaminathan,S. and Bolton,P.H. (1994) *Biochemistry*, **33**, 7517–7527.
20. Smith,F.W. and Feigon,J. (1992) *Nature*, **356**, 164–168.
21. Schultze,P., Smith,F.W. and Feigon,J. (1994) *Structure*, **2**, 221–233.
22. Smith,F.W., Francis,W.L. and Feigon,J. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 10546–10550.
23. Strahan,G.D., Shafer,R.H. and Keniry,M.A. (1994) *Nucleic Acids Res.*, **22**, 5447–5455.
24. Scaria,P.V., Shire,S.J. and Shafer,R.H. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 10336–10340.
25. Smith F.W. and Feigon,J. (1993) *Biochemistry*, **32**, 8682–8692.
26. Aboul-ela,F., Murchie,A.I.H. and Lilley,D.M.J. (1992) *Nature*, **360**, 280–282.
27. Jin,R.Z., Gaffney,B.L., Wang,C., Jones,R.A. and Breslaurer,K.I. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 8832–8836.
28. Aboul-ela,F., Murchie,A.I.H., Norman,D.J. and Lilley,D.M.J. (1994) *J. Mol. Biol.*, **243**, 458–471.
29. Wang,Y. and Patel,D.J. (1992) *Biochemistry*, **31**, 8112–8119.
30. Gupta,G., Garcia,A.E., Guo,Q., Lu,M. and Kallenbach,N.R. (1993) *Biochemistry*, **32**, 7098–7103.
31. Wang,Y. and Patel,D.J. (1993) *J. Mol. Biol.*, **234**, 1171–1183.
32. Sarma,M.H., Luo,J., Umemoto,K., Yuan,R. and Sarma,R. (1992) *J. Biomol. Struct. Dyn.*, **9**, 1131–1153.
33. Feigon,J., Koshlap,K.M. and Smith,F.W. (1995) In James,T.L. (ed.), *Methods Enzymol.* Academic Press, San Diego, CA, pp. 225–255.
34. Shafer,R.H. (1998) *Prog. Nucleic Acids Res. Mol. Biol.*, **50**, 55–93.
35. Patel,D.J., Bouaziz,S., Kettani,A. and Wang,Y. (1998) In Neidle,S. (ed.), *Oxford Handbook of Nucleic Acid Structure*. Oxford University Press, Oxford, pp. 389–453.
36. Sen,D. and Gilbert,W. (1990) *Nature*, **344**, 410–414.
37. Ross,W.S. and Hardin,C.C. (1994) *J. Am. Chem. Soc.*, **116**, 6070–6080.
38. Strahan,G.D., Keniry,M.A. and Shafer,R.H. (1998) *Biophysical J.*, **75**, 968–981.
39. Beaucage,S.L. and Caruthers,M.H. (1981) *Tetrahedron Lett.*, **22**, 1859–1862.
40. Sinha,N.D., Biernat,J., McMamis,J. and Koster,H. (1984) *Nucleic Acids Res.*, **12**, 4539–4557.
41. Plateau,P. and Gueron,M. (1982) *J. Am. Chem. Soc.*, **104**, 7310–7311.
42. Jeener,J., Meier,B.H., Bachmann,P. and Ernst,R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.
43. Braunschwiler,L. and Ernst,R.R. (1983) *J. Magn. Reson.*, **53**, 521–528.
44. Griesinger,C., Sorenson,O.W. and Ernst,R.R. (1987) *J. Magn. Reson.*, **75**, 474–492.
45. States,D.J., Haberkorn,R.A. and Ruben,D.J. (1982) *J. Magn. Reson.*, **48**, 286–292.
46. Marion,D. and Wuthrich,K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
47. Wuthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley and Sons, New York, NY.
48. Hosur,R.V., Govil,G. and Miles,H.T. (1988) *Magn. Reson. Chem.*, **26**, 927–944.
49. Patel,P.K. and Hosur,R.V. (1999) *Nucleic Acids Res.*, **27**, 2457–2464.
50. Kettani,A., Kumar,A. and Patel,D.J. (1995) *J. Mol. Biol.*, **254**, 638–656.
51. Bouaziz,S., Kettani,A. and Patel,D.J. (1998) *J. Mol. Biol.*, **282**, 619–636.
52. Kettani,A., Bouaziz,S., Gorin,A., Jones,R.A. and Patel,D.J. (1998) *J. Mol. Biol.*, **282**, 637–652.
53. Cheong,C. and Moore,P.B. (1992) *Biochemistry*, **31**, 8406–8414.