# Effect of Thymine Tract Length on the Structure and Stability of Model Telomeric Sequences†

Qiu Guo,‡ Min Lu,§ and Neville R. Kallenbach*,⊥

*Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, and Department of Chemistry, New York University, New York, New York 10003*

*Received November 6, 1992; Revised Manuscript Received January 19, 1993*

ABSTRACT: DNA from the telomeres at the ends of eukaryotic chromosomes contains a stretch of simple tandemly repeated sequences in which clusters of G residues alternate with clusters of T/A sequences along one DNA strand. Model telomeric G-clusters form four-stranded structures in $Na^+$ or $K^+$, stabilized by Hoogsteen pairing between G bases. DNA containing a single copy of the G-cluster can self-associate to form tetramers, with a parallel-stranded, right-handed helical structure. Two copies of the 3′-terminal G strand form a folded-back hairpin that dimerizes to create an antiparallel quadruplex structure. We show here that the tetrameric structure is strongly influenced by the T residue flanking either side of the G-cluster. The parallel tetraplex formed by single copies of the sequences $dT_nG_4$ is most stable for $n = 1$ and least stable for $n = 8$, the longest tract we have studied. At least two thymine residues are required to allow formation of antiparallel folded-back hairpin dimers from two-copy oligomers of sequence $d(T_nG_4)_2$ in $Na^+$; additional T's destabilize this structure. In $K^+$, the predominant structure formed is the four-stranded parallel tetramer in all cases. Kinetic analysis indicates that the quadruplex structure formed by *Oxytricha* telomeric DNA overhangs in the presence of $Na^+$ arises by dimerization of two Hoogsteen base-paired hairpins, with a relatively low energy barrier.

Telomeres are unique nucleoprotein structures at the ends of eukaryotic chromosomes. They serve to maintain the stability and integrity of chromosomes and allow the complete replication of the 5′ ends of the chromosomal DNA [for a review, see Blackburn (1991)]. Telomeric DNA, located at the chromosome termini, contains repeats of short tracts of G residues in one strand (reviewed in Blackburn and Szostak (1984) and Zakian (1989)]. In addition, the G-rich strand in telomeres overhangs the 3′ chromosomal terminus by two additional copies, 12–16 nucleotides in length, of the G-rich telomeric DNA sequence (Zakian, 1989). DNA with G-cluster sequences at the ends has unusual concentration-dependent conformational properties (Lipps, 1980; Oka & Thomas, 1987). The single-stranded overhangs in preparations of purified *Oxytricha* macronuclear DNAs are responsible for formation of higher molecular weight structures they are fully reversible by dilution (Oka & Thomas, 1987). The in vitro properties of oligonucleotide models containing single or multiple telomeric sequences have been investigated in order to determine the structural and energetic basis for this interaction. Synthetic DNA oligonucleotides containing one or more copies of G-rich telomeric sequences form defined structures in solution in the presence of $Na^+$ or $K^+$, stabilized by cyclical Hoogsteen hydrogen bonding between four G bases (Henderson et al., 1987; Williamson et al., 1990; Sen & Gilbert, 1988, 1990; Sundquist & Klug, 1989; Kang et al., 1992; Smith & Feigon, 1992; Gupta et al., 1992). Oligomers corresponding to the overhang structure in *Tetrahymena* $d(T_2G_4)_2$ form an intramolecular, double-stranded hairpin loop structure at low ionic strength (Henderson et al., 1987); addition of $Na^+$ or $K^+$ ions results in dimerization of the two hairpins to yield an antiparallel, folded-back quadruplex structure (Sundquist & Klug, 1989). Similarly, four copies of G-rich sequences such as $d(T_2G_4)_4$ and $d(T_4G_4)_4$ form an intramolecular, four-stranded G-quadruplex in the presence of $Na^+$ or $K^+$ (Williamson et al., 1989). A high-resolution X-ray structure of $dG_4T_4G_4$ in the presence of $K^+$ has recently been reported (Kang et al., 1992), in which two folded-back hairpins dimerize to form an antiparallel structure with the thymine loops lying along the edges (Figure 1B). On the other hand, an NMR study of the same sequence in the presence of $Na^+$ has been interpreted to indicate that this telomeric DNA oligomer dimerizes to form a symmetric antiparallel quadruplex with the thymine loop running diagonally across the G-tetrad (Figure 1C) (Smith & Feigon, 1992). In both cases, the sugar–phosphate backbone is found to have an alternating pattern of syn and anti glycosyl conformations along each strand in the G-quartet.

Oligonucleotides containing a single copy of a G-cluster can associate to form a stable parallel four-stranded complex in $Na^+$ (Figure 1A) (Sen & Gilbert, 1988, 1990; Lu et al., 1992; Wang & Patel, 1992). In the presence of $K^+$, in addition to tetramers, higher order products can form, provided the 3′ end of the strand is guanine (Sen & Gilbert, 1992; Lu et al., 1992). The tetrameric structure in $K^+$ formed by the *Oxytricha* telomeric unit sequence $dT_4G_4$ has recently been determined by NMR (Gupta et al., 1992); the stacked guanine tetrads form a parallel, right-handed helix, with all the glycosyl conformations anti (Gupta et al., 1992). We have previously investigated the structure and thermal stability of two short model oligomers, $dT_4G_4$ and $dT_4G_4T$, in $Na^+$ and $K^+$, and found that the additional T at the 3′ end not only inhibits formation of higher order polymeric complexes in $K^+$ but

† This work was supported by Grant CA-24101 from the National Institutes of Health (N.R.K.). Q.G. and M.L. are supported by National Research Service Awards of NIH (HL-08775 and GM-15405).

* Author to whom correspondence should be addressed.
‡ Massachusetts Institute of Technology.
§ Harvard University.
⊥ New York University.

0006-2960/93/0432-3596$04.00/0 

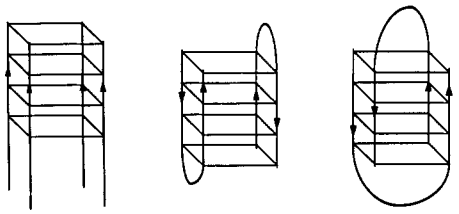


FIGURE 1: Schematic structures of the single and double repeats of *Oxytricha* telomeric DNA sequence. Panel A illustrates formation of the parallel-stranded tetrameric structure of $dT_4G_4$ in the presence of $Na^+$ or $K^+$ (Sen & Gilbert, 1988, 1990; Lu et al., 1992; Gupta et al., 1992; Wang & Patel, 1992). Schematic diagrams are shown of antiparallel folded-back dimeric structures of $dG_4T_4G_4$ in the presence of $K^+$ (B) (Kang et al., 1992) and $Na^+$ (C) (Smith & Feigon, 1992).

stabilizes the tetrameric structure over that in $dT_4G_4$ in $Na^+$ (Lu et al., 1992). The NMR study indicates that the T flanking 5′ guanine in the tetramer $d(T_4G_4)_4$ is helical, while the remaining T's are progressively more disordered (Gupta et al., 1993). Here we have synthesized the *Oxytricha* unit sequence, $dT_4G_4$, together with a series of variants, $dT_nG_4$, with $n$ = 1–3 and 8, in order to investigate the role of thymine residues with G-tetrads in parallel tetrameric structure. We demonstrate that the T flanking either side of the G-cluster interacts strongly with the cyclic guanine tetrads, stabilizing the tetrameric complex. Moreover, at low temperature the remaining T's in the tetramer are found to assume an orderly structure.

Naturally occurring telomeric sequences identified so far contain the consensus sequence $d(T/A)_{1-4}G_{1-8}$ (Blackburn & Szostak, 1984). Sundquist and Klug (1989) hypothesized that the role of thymines in the telomeric overhang is to provide "spacers" that allow hairpin formation by looping and thereby reduce steric constraints between the G-tetraplex and adjacent duplex regions. We are interested here in the question of how flanking thymines influence the formation of G-quadruplex structures in vitro. To address this question, we have synthesized oligonucleotides corresponding to the 3′-terminal overhang of the *Oxytricha* telomeric sequence $dT_4G_4T_4G_4$ and a series of variants, $dT_nG_4T_nG_4$, $n$ = 1–3 and 8, for analysis of the role of the T tract in the structure and stability of the G-quartet. Our results indicate that the formation of dimer or tetramer depends on both the length of the T tract and the presence of $Na^+$ or $K^+$. One T ($n$ = 1) is insufficient to allow formation of antiparallel folded-back dimeric structure in $Na^+$; for $n \geq 2$, CD spectroscopy indicates formation of folded-back dimers. Potassium ions favor parallel four-stranded tetramer formation. Thermal unfolding experiments indicate that the optimal length of the T tract in the two-copy sequence is 2. The second-order rate constant for forming G-quartet structure in $d(T_4G_4)_2$ in $Na^+$ is found to have a relatively low activation energy, 12 kcal/mol. Thus there is no indication in these complexes that the folded-back structures are not simple dimers.

## MATERIALS AND METHODS

*Synthesis and Purification of Oligonucleotides.* Oligonucleotides used in this study were synthesized on an ABI 391 DNA synthesizer and deprotected by routine phosphoramidite procedures (Caruthers, 1991). Strands were purified by preparative HPLC on a Du Pont Zorbax Bio Series oligonucleotide column, following the manufacturer's recommended elution protocol. Oligonucleotides were labeled at their 5′ termini using T4 polynucleotide kinase (Boehringer); the labeled strands were purified by polyacrylamide gel electrophoresis.

*Solution Preparation.* Concentration of DNA strands were determined spectrophotometrically at 260 nm and 80 °C as described (Cantor et al., 1970). Unless otherwise indicated, DNA solutions were prepared in a 10 mM phosphate buffer, with 0.1 mM EDTA and NaCl or KCl added, adjusted to pH 7. The DNA strands of $dT_nG_4$, $n$ = 1–4 and 8, were evaporated to dryness at 50 °C over 10 h, dissolved in 10 mM phosphate buffer (pH 7), and incubated at 4 °C for 48 h. The DNA solutions of $dT_nG_4T_nG_4$, $n$ = 1–4 and 8, were heated to 95 °C for 2 min, cooled slowly, and incubated at room temperature for 24 h and at 4 °C for 12 h.

*DMS Protection Assay.* Methylation protection experiments were performed as described by Maxam and Gilbert (1980). A total of 1 μL of a fresh 1:100 dilution of DMS in water was added to 10 μL of DNA samples in TE buffer without or with indicated salts for 30 min at 4 °C. The reaction were terminated by two sequential rapid ethanol precipitations, and the mixture was then lyophilized. The DNAs were cleaved at the methylation sites by treatment with 100 μL of 1 M piperidine at 90 °C for 30 min and lyophilized extensively.

*Circular Dichroism (CD) Spectroscopy.* CD spectra were recorded using an AVIV Model 60DS CD spectropolarimeter equipped with a programmable, thermoelectrically controlled cell holder. Each spectrum corresponds to an average of three scans from which the buffer background was subtracted. The total DNA strand concentration is $1 \times 10^{-4}$ M. Samples were preequilibrated at 4 °C for 1 h. The quantity $\epsilon_L - \epsilon_R$ were calculated for each species by dividing the apparent signal in degrees, $\theta$, by $3.3C$, where $C$ is defined in moles of residues. CD melting curves were recorded at 265 or 290 nm. Ellipticities at 4 °C were normalized to 1.0. Samples were heated at a rate of 0.5 °C/min. These melting curves allow us to measure the transition temperature ($T_m$) as well as the relevant thermodynamic parameters. These parameters were calculated by using standard procedures reported by Marky and Breslauer (1987) and correspond to a two-state approximation of the helix–coil transition of each molecule.

*Kinetic Experiments.* The $d(T_4G_4)_2$ strand was denatured by heating in boiling water for 4 min in TE buffer (10 mM Tris·HCl and 1 mM EDTA) (pH 7) and suspended in 10 mM phosphate buffer, pH 7, containing 0.1 mM EDTA and 50 mM NaCl, in a total volume of 10 μL. Each sample contained an equal quantity of labeled DNA with different concentrations of cold strands. Samples were incubated at 12.0 °C for 25 min, 26.0 °C for 10 min, and 38.5 °C for 5 min, respectively, and the products were assayed on a 20% nondenaturing gel at 4 °C and in 0.6× TBE with 50 mM NaCl. The ratios of counts in monomer and dimer bands were determined after cutting out the bands and scintillation counting.

*Gel Electrophoresis.* Native polyacrylamide gels were run on 20% native gels (19:1 monomer/bis ratio) at 4 °C for 20 h at 100 V (ca. 8 V/cm). The electrophoresis plates were jacketed and cooled with circulating water to provide a running temperature of 4 ± 1 °C in the gel throughout the electrophoresis. The buffer system contained 53.4 mM Tris borate–EDTA buffer, pH 8.3 (0.6× TBE), wtih added 50 mM NaCl or KCl. For denaturing polyacrylamide gels, the samples for electrophoresis were taken up in formamide loading buffer, heated briefly to 90 °C, cooled, and then run on a 20% denaturing polyacrylamide gel for 3 h at 2000 V (ca. 50 V/cm) and 40 °C. The gels were dried immediately on a vacuum-drying apparatus (Hoefer) and exposed at room temperature to X-ray film.

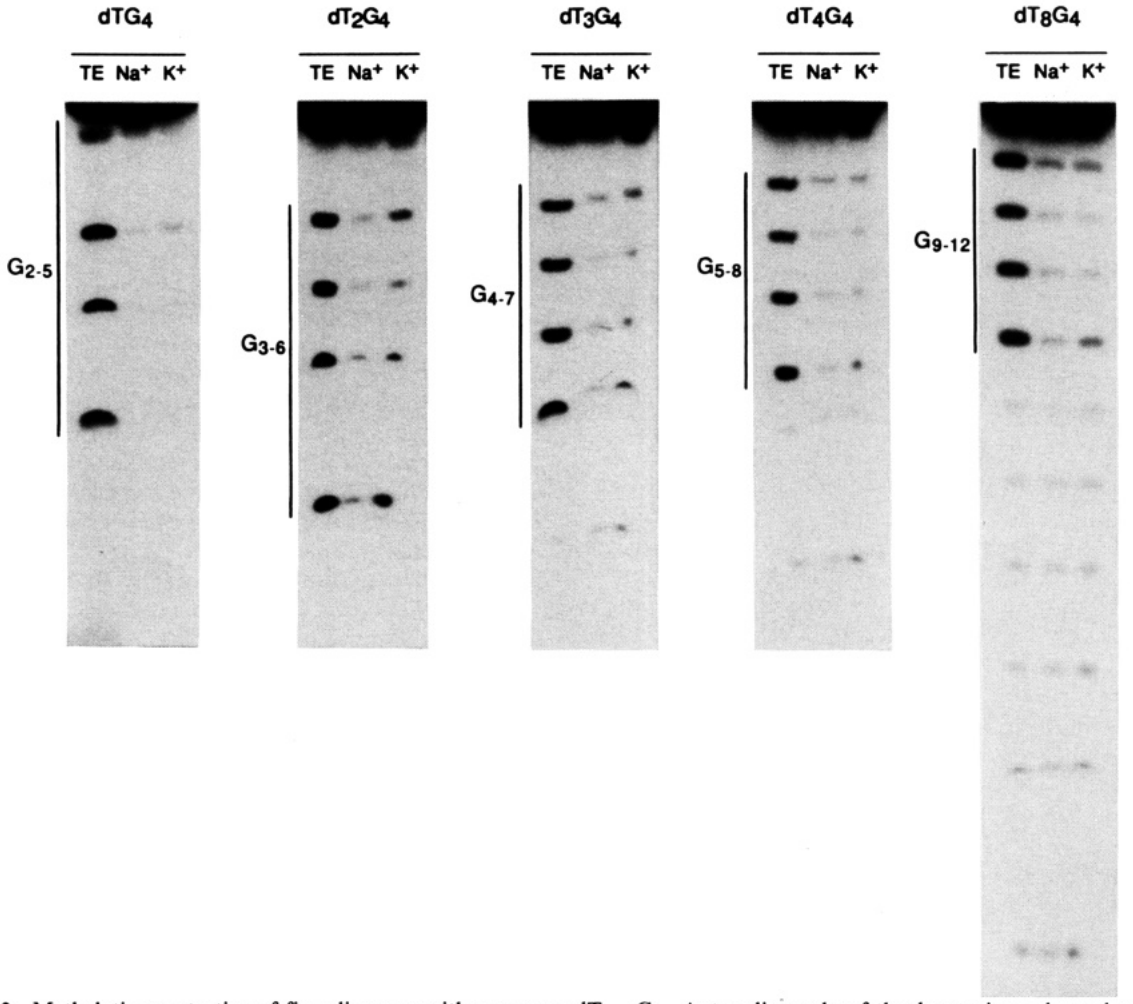


FIGURE 2: Methylation protection of five oligomers with sequences $dT_{1-4,8}G_4$. Autoradiographs of the denaturing polyacrylamide gels are shown. The DNA strands 5′-labeled with $^{32}P$ were methylated by DMS in 10 mM phosphate buffer (pH 7) and 0.1 mM EDTA, without or with 200 mM NaCl or KCl added, and then cleaved with piperidine.

## RESULTS

*Formation of Parallel Tetraplex Structures by Single Copies of G-rich Oligomers That Differ in T Tract Length.* Samples of five oligomers with sequences $dT_{1-4,8}G_4$ were prepared in TE buffer, without and with 200 mM NaCl or KCl. The mobility of the tetrameric DNA products was examined on nondenaturing polyacrylamide gels (data not shown). All five oligomers are found to form tetramers in the presence of $Na^+$ or $K^+$ by using an assay in which strands of different T length are mixed (Sen & Gilbert, 1988) (data not shown). Formation of Hoogsteen G·G base pairing in tetraplex DNA structure can be probed chemically with DMS, which preferentially methylates N7 of guanine (Sen & Gilbert, 1988). When the five oligomers are methylated by DMS in 200 mM NaCl or KCl, protection of N7 in all guanines of these sequences results (Figure 2). This indicates that these single G-rich oligomers are structured in $Na^+$ or $K^+$. We have previously shown that the parallel G-tetraplex structure has a characteristic CD spectral signature with a peak near 265 nm and a trough near 240 nm (Lu et al., 1992). The CD spectra of oligomers with sequences $dT_{1-4,8}G_4$ in $Na^+$ or $K^+$ are similar, showing these features (data not shown). Thus each oligomer containing one copy of G-cluster with different T lengths is able to form a four-stranded, parallel structure in $Na^+$ or $K^+$.

*Thermodynamic Profiles for Parallel G-Tetraplex Formation.* It is of interest to determine the energetics of the interactions that stabilize these DNA complexes and how these vary with sequence. We have evaluated the thermodynamics of forming tetrameric DNA structures from single-copy G-cluster oligomers with different lengths of the T tract in $Na^+$ by determining the midpoint thermal denaturation temperatures of these DNA molecules from CD data as a function of total DNA strand concentration. All CD melting profiles are monophasic (data not shown). Plots of $1/T_m$ versus ln $C_T$ for each DNA molecule in $Na^+$ are shown in Figure 3. The van't Hoff thermodynamic parameters derived from these plots are summarized in Table I. The results show that $dTG_4$ forms the most stable tetramer among the oligomers, while additional T's destabilize the tetrameric structure by nearly 10 °C over that in $dT_4G_4$ in the presence of 200 mM NaCl, the effect arising from concerted enthalpic and entropic contributions (see Table I). This provides further evidence to support the NMR data that in the tetramer $d(T_4G_4)_4$ the T adjacent to the G-cluster is tightly stacked over the guanine G-tetrads, while the succeeding T's are more disordered (Gupta et al., 1992).

UV cross-linking experiments have previously been described which show that the thymine tail in the parallel G-tetraplex structure is not completely flexible (Lu et al., 1992). This raises the question of whether the thymine tail stabilizes or destabilizes the tetrameric complex. Since the thermal melting behavior of these oligomers reflects their relative stability at the relatively elevated midpoint denaturation temperature, we would like to know whether or not the

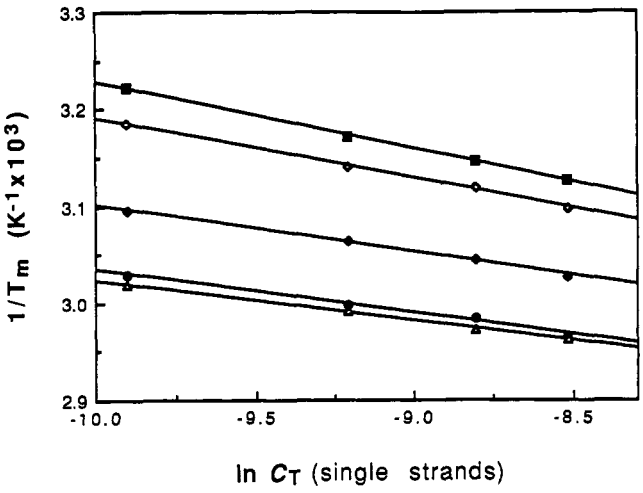


FIGURE 3: Dependence of the transition temperature of $dT_{1-4,8}G_4$ on strand concentrations in a 10 mM phosphate buffer containing 0.1 mM EDTA and 200 mM NaCl at pH 7: $dTG_4$ (▲), $dT_2G_4$ (●), $dT_3G_4$ (♦), $dT_4G_4$ (◇) and $dT_8G_4$ (■).

Table I: Thermodynamic Parameters for Complex Formations of One Copy of G-Cluster Oligonucleotides[a]

| strand | $T_m$[b] (°C) | $\Delta H°_{VH}$[c] (kcal/mol) | $\Delta H°_{VH}$[d] (kcal/mol) | $T\Delta S°_{VH}$[d] (kcal/mol) | $\Delta G°$[e] (kcal/mol) |
|---|---|---|---|---|---|
| $dTG_4$ | 61.2 | −140.1 | −142.3 | −106.8 | −35.5 |
| $dT_2G_4$ | 60.4 | −133.4 | −133.7 | −99.5 | −34.2 |
| $dT_3G_4$ | 53.2 | −124.7 | −122.3 | −95.3 | −27.0 |
| $dT_4G_4$ | 45.3 | −96.1 | −94.8 | −72.4 | −22.4 |
| $dT_8G_4$ | 42.1 | −88.2 | −86.8 | −65.8 | −21.0 |

[a] All melting experiments were done in 10 mM phosphate buffer (pH 7), 0.1 mM EDTA, and 200 mM NaCl. [b] Calculated for $10^{-4}$ M single-strand concentration. [c] Calculated from curve-fitted parameters with $n = 4$. [d] Calculated from $1/T_m$ versus ln $C_T$ parameters with the molecularity $n = 4$. [e] Calculated at 25 °C assuming $\Delta C_p = 0$.

T tail assumes a structure at low temperature that stabilizes the overall tetrameric complex. Accordingly, we have carried out an electrophoretic equilibrium assay to measure the relative distribution of five tetrameric bands formed when equal concentrations of pairs of different oligomers, for example, $dTG_4$ and $dT_4G_4$, are annealed together in $Na^+$. The reasoning is as follows. If the T tract in the tetramer is completely disordered, the tetramer $d(TG_4)_4$ should be the most stable species, resulting in the highest band intensity, and the intensities of the five bands should progressively weaken as the number of thymine residues in the tetramer increases. On the other hand, if the T tail has a partially ordered conformation, species other than $d(TG_4)_4$ can predominate and form the highest intensity band. Figure 4 shows the results of this experiment at 4 °C. The combination of two long and two short T tails, e.g., $d(TG_4)_2d(T_4G_4)_2$ (lane 7) and $d(TG_4)_2d(T_8G_4)_2$ (lane 8), shows higher intensity than any other (Figure 4). Thus the T tail in the tetrameric DNA complex appears to be structured at low temperature, presumably forming some kind of folded domain itself. To verify this, we measured the $T_m$ of each species by annealing two oligomers, $dTG_4$ and $dT_4G_4$ in different ratios in $Na^+$; the thermal melting temperatures are summarized in Table II. The higher the fraction of $dTG_4$ in the mixture, the more stable the tetrameric complex. This implies that the orderly T tail structure detected at low temperature in the gel experiment is unfolded near the $T_m$, where the interaction of the T flanking the 5′ end of the G-cluster with the terminal G-tetrad becomes the dominant factor stabilizing the tetramer.

We have previously shown that the additional 3′ T residue in the oligomer $dT_4G_4T$ stabilizes tetrameric structure by almost 20 °C over that in $dT_4G_4$ at low strand concentrations in $Na^+$ (Lu et al., 1992). One obvious possibility is that the stabilization is due to the thymine base stacking over the 3′ G-tetrad. In order to test this hypothesis, we performed a CD melting experiment by mixing two oligomers, $dT_4G_4$ and $dT_4G_4T$, in different ratios in $Na^+$. As summarized in Table III, the stability of the tetrameric structure is proportional to the fraction of $dT_4G_4T$ present. This provides evidence that the stabilization of a tetrameric complex is achieved by thymine base stacking and possibly pairing on either terminal G-tetrad in the complex.

*pH Dependence of Formation of Four-Stranded Parallel Structure in $Na^+$ or $K^+$.* It has been found that potassium ions have a profound effect on the stability and structure of G-cluster sequences (Sen & Gilbert, 1990; Guschlbauer et al., 1990; Lu et al., 1992). For example, while the tetramer $d(T_4G_4)_4$ in $Na^+$ normally unfolds on heating, in the presence of $K^+$ it generates polymeric complexes and forms a more asymmetric structure at high temperature (Lu et al., 1992). However, relatively little is known about the mechanism that accounts for the differential ion effect. In order to address how the ions affect the cyclic guanine H bonding in these DNA complexes, we have used CD spectroscopy to measure the tetrameric structure in $dT_4G_4$ in 200 mM $Na^+$ or $K^+$ at different pH values. The pH titration curves are shown in Figure 5, which shows that at room temperature the structure in $Na^+$ is different from that in $K^+$. The latter is more stable, as seen by its higher pH of unfolding at alkaline pH; the difference is greater than 1 pH unit at these concentration of salt. This is consistent with tighter binding of $K^+$ than $Na^+$ to these oligomers. The possibility that the stoichiometry differs between the $Na^+$ and $K^+$ complexes also exists.

*Factors Affecting Dimer vs Tetramer Formation.* In two-copy G-cluster oligomers, dimer vs tetramer formation is controlled by T tract length and metal ions. Electrophoretic mobilities of the series of two-copy oligonucleoties under native conditions are shown in Figure 6. In the absence of metal ions, each oligonucleotide has a mobility consistent with the monomeric species (Figure 6A). In the presence of $Na^+$, $d(TG_4)_2$ shows a dramatic decrease in mobility (Figure 6B), presumably forming the parallel-strand tetrameric structure described above. Whereas $d(T_2G_4)_2$ also has a clearly retarded mobility, $d(T_3G_4)_2$ and $d(T_4G_4)_2$ show only slightly slower mobility in $Na^+$ (Figure 6B). Following Williamson et al. (1989), we interpret these shifts as due to dimerization, forming so-called "G-quartet" structures. By contrast, $d(T_8G_4)_2$ shows enhanced mobility in $Na^+$, indicating formation of a more compact structure (Figure 6B). In $K^+$, the situation is more complicated still (Figure 6C): slowly migrating bands are observed, which we attribute to parallel-strand tetramers (Sen & Gilbert, 1988, 1990). In addition, more rapidly migrating bands appear, which migrate relatively like the dimers in $Na^+$. In the presence of 50 mM NaCl or KCl, DMS protection assays indicate that all the G residues are protected, hence involved in pairing (data not shown). These data support two hypotheses: (i) potassium ions preferentially stabilize a different structure from that in sodium, presumably a four-stranded parallel tetrameric DNA structure (Sen & Gilbert, 1990), and (ii) the single T in $d(TG_4)_2$ does not allow fold-back to create a loop in an antiparallel dimeric structure. Formation of G-quartet structure is evidently controlled by the T tract length as well as by the metal present. The

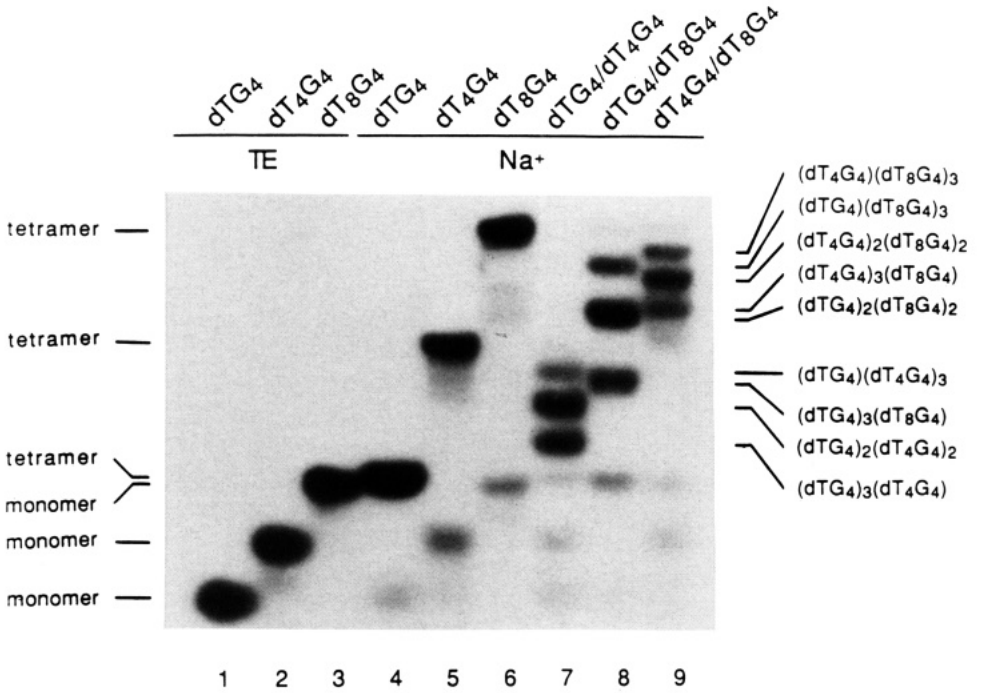


FIGURE 4: Electrophoretic equilibrium assay. An autoradiogram of the gel is shown of the equilibrium distributions of the labeled strands among five product bands arising from the four-stranded nature of the complex in $Na^+$. Lanes 1–3 are $dTG_4$, $dT_4G_4$, and $dT_8G_4$ in TE buffer (pH 7), respectively. Lanes 4–6 are $dTG_4$, $dT_4G_4$, and $dT_8G_4$ in TE buffer with 200 mM NaCl, respectively. Lanes 7–9 are $dTG_4/dT_4G_4$, $dTG_4/dT_8G_4$, and $dT_4G_4/dT_8G_4$ in TE buffer with 200 mM NaCl, respectively. Each sample contains an equal amount of denatured labeled oligomer(s) to give a final DNA concentration of $1 \times 10^{-4}$ M DNA.

Table II: Midpoint Transition Temperatures of Mixtures of $dTG_4$ and $dT_4G_4$[a]

| | oligomer | | | | |
|---|---|---|---|---|---|
| | $dTG_4$ | $dTG_4/dT_4G_4$ (3:1) | $dTG_4/dT_4G_4$ (1:1) | $dTG_4/dT_4G_4$ (1:3) | $dT_4G_4$ |
| $T_m$ (°C) | 61.2 | 56.9 | 53.1 | 50.4 | 45.3 |

[a] The thermal melting experiments were done in 10 mM sodium phosphate buffer, pH 7, 0.1 mM EDTA, and 200 mM NaCl, with a total strand concentration of $1 \times 10^{-4}$ M.

Table III: Midpoint Transition Temperatures of Mixtures of $dT_4G_4$ and $dT_4G_4T$[a]

| | oligomer | | | | |
|---|---|---|---|---|---|
| | $dT_4G_4$ | $dT_4G_4/dT_4G_4T$ (3:1) | $dT_4G_4/dT_4G_4T$ (1:1) | $dT_4G_4/dT_4G_4T$ (1:3) | $T_4G_4T$ |
| $T_m$ (°C) | 45.4 | 52.7 | 56.2 | 59.2 | 61.4 |

[a] The thermal melting experiments were done in 10 mM sodium phosphate buffer, pH 7, 0.1 mM EDTA, and 200 mM NaCl, with a total strand concentration of $1 \times 10^{-4}$ M.

structures in several cases are not unique, but represent a mixture of forms.

*Conformational Aspects of T Loop Size.* CD spectroscopy has proven to be a very sensitive technique for probing the overall conformation of telomeric models (Henderson et al., 1990; Hardin et al., 1991; Lu et al., 1992; Guo et al., 1992a,b; Balagurumoorthy et al., 1992). Here, we have used CD to determine how thymine residues in the loop affect the conformation of the quadruplex structures. The CD spectra of the series of oligomers containing two copies of G-clusters are shown in Figure 7. In 50 mM NaCl, the spectrum of $d(TG_4)_2$ is very similar to that assigned to parallel-stranded tetrameric DNA, with a trough at 240 nm and a peak at 265 nm (Figure 7A) (Lu et al., 1992). Taken together with the gel mobility data (Figure 6B), this is consistent with formation

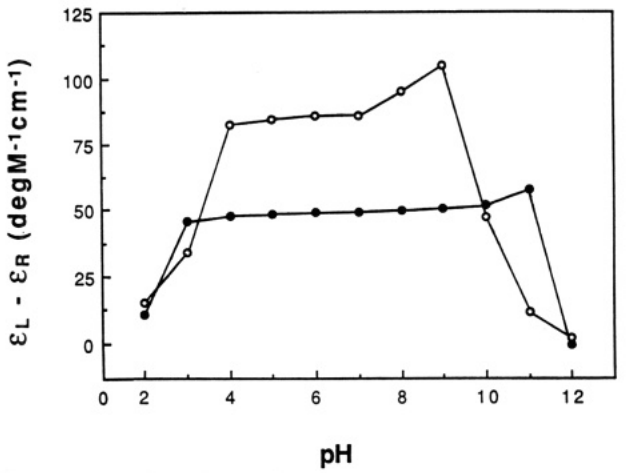


FIGURE 5: pH dependence of the formation of G4 DNA in 10 mM phosphate buffer, pH 7, 0.1 mM EDTA, and 200 mM NaCl (○) or KCl (●).

of a parallel four-stranded tetramer, as sketched in Figure 1A. By contrast, $d(T_2G_4)_2$, $d(T_3G_4)_2$, and $d(T_4G_4)_2$ have similar CD spectra to the four-copy sequence, $d(T_{G4})_4$ (Guo et al., 1992), showing a minimum at 260 nm and a positive band at 295 nm (Figure 7A). The latter forms a G-quartet structure (Williamson et al., 1989), leading us to conclude that each of these species forms an antiparallel folded-back dimer structure. By contrast, the CD spectrum of the oligomer $d(T_8G_4)_2$ is intermediate, having a minimum at 240 nm and a very broad peak centered at 270 nm (Figure 7A). We cannot deduce the predominant configuration of this molecule in $Na^+$.

In the presence of $K^+$, the CD spectrum of $d(TG_4)_2$ is the same as in $Na^+$, consistent with the parallel-strand tetrameric structure (Figure 7B) (Lu et al., 1992). The spectrum of $d(T_2G_4)_2$ is very similar, with a trace of new band appearing at 295 nm (Figure 7B). This band strengthens progressively relative to the band at 265 nm in the spectra of $d(T_3G_4)_2$ and $d(T_4G_4)_2$ (note the scale change in the successive spectra)

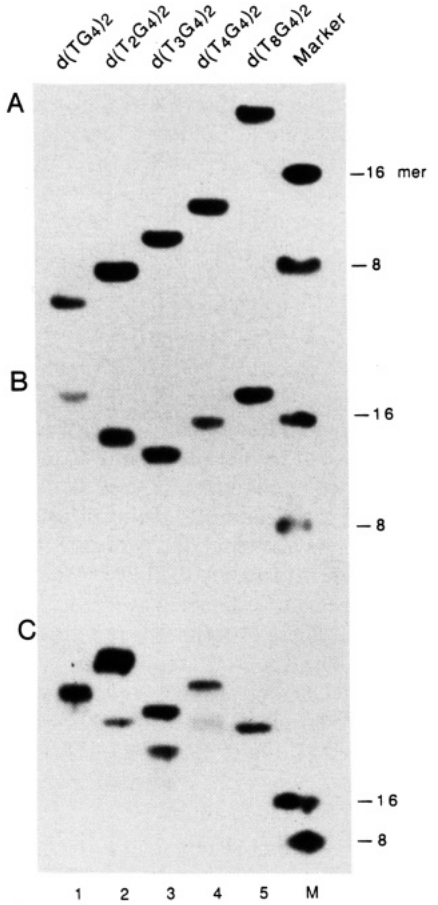


FIGURE 6: Electrophoresis of the two copies of telomeric oligonucleotides in nondenaturing gels: (A) 0.6× TBE; (B) 0.6× TBE + 50 mM NaCl; (C) 0.6× TBE + 50 mM KCl. The DNA samples were prepared as described under Materials and Methods. The gels were run at 4 °C until the dye, bromophenol blue, migrated 10 cm.

(Figure 7B). Again, the spectrum of $d(T_8G_4)_2$ is quite distinct, although the shape resembles that of the parallel tetramer. In combination with the gel mobility data, we conclude that $d(TG_4)_2$ can only form a parallel-strand tetramer, while the other chains form this structure preferentially in potassium, along with other variants. In $Na^+$, $d(T_2G_4)_2$, $d(T_3G_4)_2$, and $d(T_4G_4)_2$ show evidence of forming an antiparallel dimer of hairpins, as does the series $dG_4T_nG_4$ ($n$ = 1–4) studied by Sasisekharan's group (Balagurumoorthy et al., 1992) and $dG_4T_4G_4$ (Kang et al., 1992).

*Relative Stabilities of T Loop Structures.* Since $d(T_2G_4)_2$, $d(T_3G_4)_2$, and $d(T_4G_4)_2$ form similar structures in the presence of $Na^+$, we can ask how these differ in their stability. To address this, we used the CD signals to monitor the thermal unfolding of the structure in each molecule (Lu et al., 1992; Guo et al., 1992). The results are summarized in Table IV. At identical strand concentrations, the midpoint denaturing temperature of $d(T_2G_4)_2$ is higher than that of $d(T_3G_4)_2$, which in turn is higher than that of $d(T_4G_4)_2$. The shorter the T tract, the more stable the complex. This effect is not due to electrostatic repulsion from the pendant T's in the free tail of each chain, since the differences in $T_m$ persist in 500 mM NaCl.

## DISCUSSION

In terms of the different structural models proposed for the hairpin dimers in $dG_4T_4G_4$ from X-ray and NMR data, these results point to an interesting conclusion. If each member of the series forms the same type of structure, as indicated by the CD spectra, then the maximal stability of the repeat containing only two T's would seem to make it unlikely for this loop to span the G-tetrad diagonally (Figure 1C) rather than run along the shorter edge of the face (Figure 1B). The distances along these routes differ by 13 Å, according to the distances from the X-ray model (Kang et al., 1992). Possibly the $d(T_2G_4)_2$ structure differs in subtle ways from that in $d(T_4G_4)_2$. However, the mechanism of formation of the tetrad differs fundamentally in the two structures, as noted by Smith and Feigon (1992). This is, the structure proposed by Kang et al. (1992) represents a simple dimer formed from apposing two hairpins, while that of Smith and Feigon (1992) requires invasion of one hairpin by the second strand, as sketched in Figure 1. The latter would differ from the former in requiring a greater activation barrier, since the two dimers can interact readily in their folded states according to the first mechanism. We have measured the second-order rate of forming the tetrad in $d(T_4G_4)_2$, in order to deal with the same loop size as that in the NMR experiment. The results are summarized in Figure 8. Panel A illustrates the procedure used. A small constant quantity of labeled strand was mixed with increasing concentrations of unlabeled $d(T_4G_4)_2$, and the mixture was

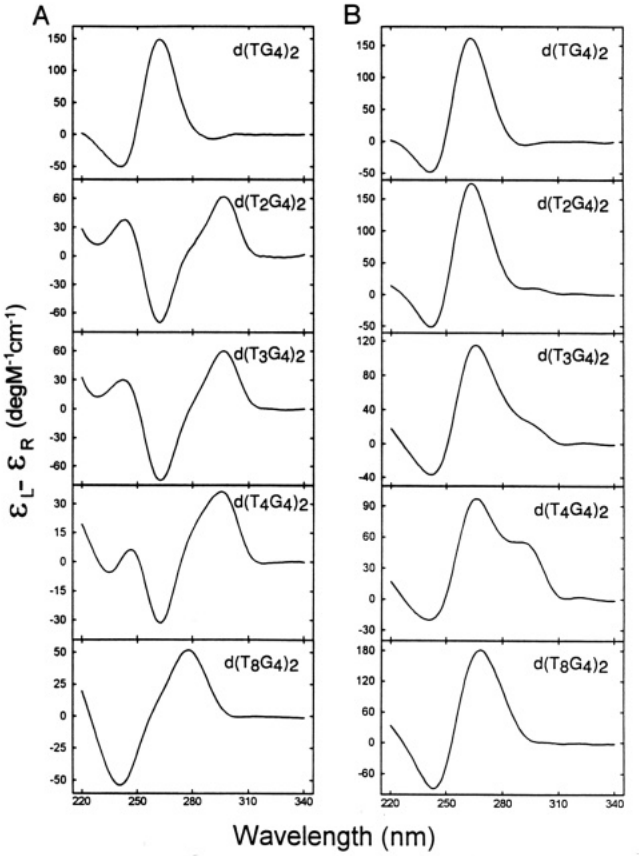


FIGURE 7: CD spectra of the two copies of telomeric oligonucleotides in the presence of 50 mM $Na^+$ (A) and $K^+$ (B). The total DNA strand concentration is $1 \times 10^{-4}$ M. Samples were preequilibrated at 4 °C for 1 h.

Table IV: Midpoint Denaturing Temperatures of Two-Copy Oligomers[a]

| | oligomer | | |
|---|---|---|---|
| | $d(T_2G_4)_2$ | $(T_3G_4)_2$ | $d(T_4G_4)_2$ |
| $T_m$ (°C) | 67.1 | 64.0 | 60.9 |

[a] The thermal melting experiment was done in 10 mM sodium phosphate buffer, pH 7, 0.1 mM EDTA, and 50 mM NaCl, with a total strand concentration of $1 \times 10^{-4}$ M.

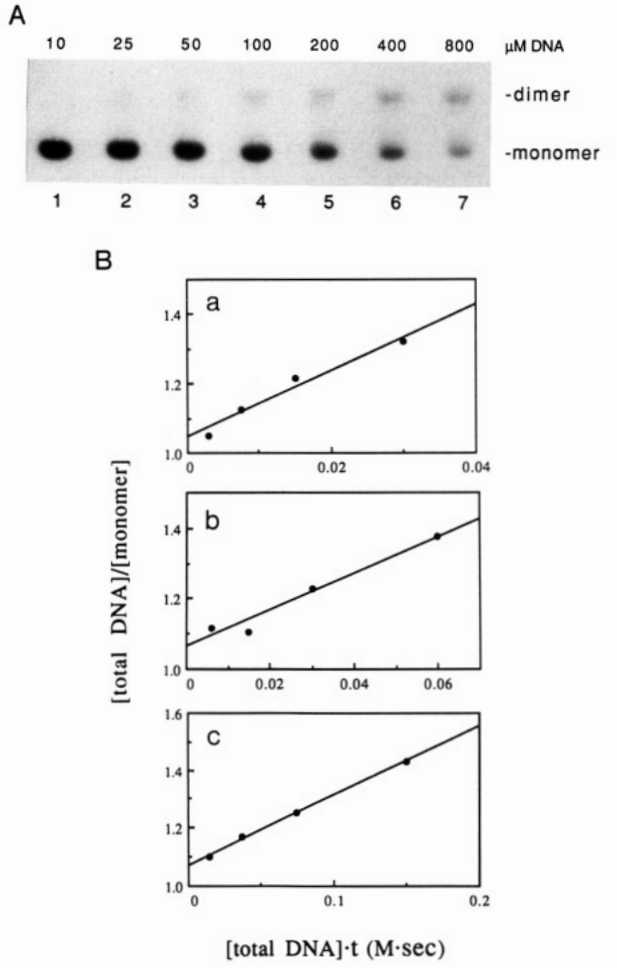


FIGURE 8: Kinetic analysis of G-quartet DNA formation of $d(T_4G_4)_2$ in 50 mM NaCl, performed as described in the text. (A) Dependence of rate of antiparallel folded-back dimer formation on DNA concentration. (B) Plots of the total DNA strand concentration divided by the monomer concentration as the total DNA strand concentration multiplied by time are shown at 12.0 °C (25 min) (a), 26.0 °C (10 min) (b), and 38.5 °C (5 min) (c).

Table V: Second-Order Rate Constants of Formation of the Dimer $d(T_4G_4)_2$

| $T$ (°C) | $k$ ($M^{-1}$ $sec^{-1}$) | $T$ (°C) | $k$ ($M^{-1}$ $sec^{-1}$) |
|---|---|---|---|
| 12.0 | 2.42 | 38.5 | 9.65 |
| 26.0 | 5.18 | | |

incubated at 38.5 °C for 5 min. The products were run at 4 °C on a native gel, to separate the monomers from the dimers. These bands were cut out and counted to monitor the extent of reaction. This procedure was repeated at 12.0 and 26.0 °C. Panel B shows a log–log plot of the monomer concentration as a function of the total strand concentration. The linearity of this graph indicates second-order reaction kinetics, with the slope measuring the rate constant, $k$ (Table V). A plot of the rates in Table V as a function of $1/T$ yields the apparent activation energy for this process, which is 9 kcal/mol. Thermodynamic studies indicate that a single pair in each G layer requires 8 kcal/mol for disrupting each G base pair in the hairpin dimer (Riesner & Romer, 1973). Our results place a strong restriction on the extent of structural opening that accompanies tetraplex formation, on the face of it arguing that no extensive disruption of base pairs occurs on hairpin dimerization. Consequently, we favor the edge-structural arrangement in the looped dimer for $d(T_2G_4)_2$, $d(T_3G_4)_2$, and $d(T_4G_4)_2$ in our system.

*Role of T Residues in Telomeric DNA Sequences.* The telomeric DNA repeats from a variety of organisms have the consensus sequence $d(T/A)_{1-4}(G)_{1-8}$ (Blackburn & Szostak, 1984). We have previously shown that the presence of A has a strong effect on the stability and structure of such sequences (Guo et al., 1992b). Here we find that the T tract influences the structure and stability of the tetraplex structures that are formed in the presence of $Na^+$ or $K^+$. The role of the T tract has been postulated to provide conformational flexibility in the complexes formed from these sequences (Sundquist & Klug, 1989). This is clearly so in terms of the difference in structures formed from $dTG_4TG_4$ and $dT_2G_4T_2G_4$ or the longer variants. A single T imposes a constraint that inhibits folding over, and the only structure formed is the parallel one. A T tract of length 8 clearly destabilizes all the structures in these models. In the case of the parallel-strand tetramer structure formed by single-copy oligomers (Lu et al., 1992), at high temperature there is a clear destabilizing effect as the T tract length increases. As shown above, however, this is not the case at lower temperatures. A distinct interaction between the T loop sequences can be detected using the electrophoretic assay in Figure 4. These experiments further demonstrate a real difference in the structures stabilized by $Na^+$ and by $K^+$: on the basis of the CD analysis, $K^+$ tends to favor formation of a parallel-strand structure (Lu et al., 1992) rather than the folded-back models (Kang et al., 1992). This has also been observed in connection with truncated two-copy models (Balagurumoorthy et al., 1992).

Our results on the two-copy oligomer series $d(T_nG_4)_2$ differ in fundamental respects from the results of a similar study of sequences of truncated two-copy models, $dG_4T_nG_4$ (Balagurumoorthy et al., 1992). In particular, the transition profiles for the latter system show an opposite behavior from comparable members of the series we have studied; the order of stability observed is $dG_4T_4G_4 > dG_4T_3G_4 > dG_4T_2G_4$. The last species, in fact, forms only a partial folded-back structure relative to the others, implying that loop closure with two T's is difficult, while it is relatively facile for three and four T's. Assuming no kinetic or equilibrium artifacts in either analysis, we are forced to conclude that the initial T tract in the two-copy sequence has a major role in establishing the conformation and stability of these molecules. This is unexpected. The truncated models provide us with the most detailed structural information so far concerning folded-back dimers (Kang et al., 1992). One possible explanation is that the T flanking the 5′ G-quartet clearly interacts strongly with the G layer, as shown above, and this affects the stability of the structure that can form 3′ to this sequence, in particular the loop structure. The order of stability in two-copy models $d(T_nG_4)_2$ that we observe might then not apply to the truncated species, including the model based on NMR (Smith & Feigon, 1992). Our attempts to investigate members of the two-copy series using NMR reveal multiple species at these concentrations, perhaps consistent with the idea that the two systems are fundamentally different. Nevertheless, the 3′-terminal overhang contains two copies of the sequence, not a truncated version (Zakian, 1989). This makes the conclusions of the present study relevant in attempting to understand the conformational repertoire in telomeric DNAs. It must further be noted that none of these models addresses the role of the G-rich sequence in the context of the complementary sequence present, except for the 3′-terminal overhang. This is a complex issue. It can be argued that the opposite strand is locally unpaired on replication of the DNA, allowing formation of a variety of folded-back structures such as are seen here and

have been reported by others. This is particularly likely to occur on ablation of the ends in cells lacking telomerase activity, which lose sequences at the ends of chromosomes as they replicate (Levy et al., 1992). Telomeres include proteins that interact with these sequences, also (Blackburn, 1991), and the role of these in the strand equilibrium has not been elucidated. The perspective of this study is that the G-rich strand alone can form a variety of different stable structures, rather than one single structure, depending on the context and the ambient conditions. Single copies can associate to form a stable parallel-strand tetramer stabilized by $K^+$. Two copies tend to favor a folded-back structure, except if only one T is present in the sequence. We show here that the pathway for forming the ordered $Na^+$ structure in $d(T_2G_4)_2$ proceeds via dimerization, emphasizing the role of the local strand concentration. Truncation of the 5′ T tract appears to alter the nature and determination of the structure in these molecules, on the basis of the apparent differences between our results and those of Balagamoorthy et al. (1992). The high-resolution structure of a complete two-copy model has not yet been reported.

## REFERENCES


Balagurumoorthy, P., Brahmachari, S. K., Mohanty, D., Bansal, M., & Sasisekharan, V. (1992) *Nucleic Acids Res. 20*, 4061–4067.

Blackburn, E. (1991) *Nature 350*, 569–571.

Blackburn, E. H., & Szostak, J. W. (1984) *Annu. Rev. Biochem. 53*, 163–194.

Cantor, C., Warshaw, M. W., & Shapiro, H. (1970) *Biopolymers 9*, 1059–1077.

Caruthers, M. H. (1991) in *Chemical and Enzymatic Synthesis of Gene Fragments* (Gassen, H. G., & Lang, A., Eds.) pp 71–79, Verlag Chemie, Weinheim.

Guo, Q., Lu, M., Marky, L. A., & Kallenbach, N. R. (1992a) *Biochemistry 31*, 2451–2455.

Guo, Q., Lu, M., & Kallenbach, N. R. (1992b) *J. Biol. Chem. 267*, 15293–15300.

Gupta, G., Garcia, A. E., Guo, Q., Lu, M., & Kallenbach, N. R. (1993) (submitted for publication).

Guschlbauer, W., Chantot, J. F., & Thiele, D. (1990) *J. Biomol. Struct. Dyn. 8*, 491–511.

Hardin, C. C., Henderson, E. R., Watson, T., & Prosser, J. K. (1991) *Biochemistry 30*, 4460–4472.

Henderson, E. R., Hardin, C. C., Walk, S. K., Tinoco, I., & Blackburn, E. H. (1987) *Cell 51*, 899–908.

Henderson, E. R., Moore, M., & Malcolm, B. A. (1990) *Biochemistry 29*, 732–737.

Kang, C., Zhang, X., Ratliff, R., Moyzis, R., & Rich, A. (1992) *Nature 356*, 126–131.

Levy, M. Z., Allsopp, R. C., Futcher, A. B., Greider, C. W., & Harley, C. B. (1992) *J. Mol. Biol. 225*, 951–960.

Lipps, H. J. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4104–4108.

Lu, M., Guo, Q., & Kallenbach, N. R. (1992) *Biochemistry 31*, 2455–2459.

Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601–1620.

Maxam, A., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Oka, Y., & Thomas, C. A. (1987) *Nucleic Acids Res. 15*, 8878–8899.

Riesner, D., & Romer, R. (1973) in *Physico-Chemical Properties of Nucleic Acids* (Duchesne, J., Ed.) Vol. II, Academic Press, London.

Sen, D., & Gilbert, W. (1988) *Nature 334*, 364–366.

Sen, D., & Gilbert, W. (1990) *Nature 344*, 410–414.

Sen, D., & Gilbert, W. (1992) *Biochemistry 31*, 65–70.

Smith, F. W., & Feigon, J. (1992) *Nature 356*, 164–168.

Sundquist, W. I., & Klug, A. (1989) *Nature 342*, 825–829.

Wang, Y., & Patel, D. J. (1992) *Biochemistry 31*, 8112–8119.

Williamson, J. R., Raghuraman, M. K., & Cech, T. R. (1989) *Cell 59*, 871–880.

Zakian, V. A. (1989) *Annu. Rev. Genet. 23*, 579–604.