Thermodynamics of G-Tetraplex Formation by Telomeric DNAs†

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

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

Received August 7, 1992; Revised Manuscript Received October 20, 1992

ABSTRACT: Telomeres are structures at the ends of eukaryotic chromosomes, the DNA of which contains stretches of tandemly repeated sequences with G clusters along one strand. Model telomeric G-rich DNAs can form different tetraplex structures, stabilized by cyclic hydrogen bonding of four guanines in the presence of metal ions such as Na^+ or K^+. Oligonucleotides with a single copy of the *Oxytricha* sequence dT_4G_4 form a tetramer, with a parallel-stranded, right-handed helical structure. Additional copies favor folded-back structures that associate to form an antiparallel dimer. The parallel-stranded tetramer has all G's in the anti configuration, while the folded-back dimer has alternating syn and anti nucleotide conformations along each strand. Here we have constructed two G-tetraplex structures, containing identical G-tetrad base pairs, from oligonucleotides. One has the truncated telomeric sequence from *Oxytricha*, $dG_4T_4G_4$, which forms an antiparallel G-quartet structure; the second is constrained to form a parallel G-strand arrangement by insertion of a 5′-p-5′ linkage between two dT_2G_4 sequences. Each oligomer forms a defined G-tetraplex dimeric structure in the presence of Na^+. The standard-state enthalpies, entropies, and free energy for formation of these tetraplexes have been determined. The parallel strand structure is thermodynamically more stable than the antiparallel one, primarily because of both greater enthalpy and entropy of formation. In addition, the two molecules differ in their interaction with sodium ions, reflecting a difference in ion binding and therefore in structure between the two forms.

Telomeres are specialized DNA–protein structures located at the ends of chromosomes, essential for the stability and complete replication of chromosomal termini [reviewed in Blackburn (1991)]. Telomeric DNA contains highly repeated simple sequences, with characteristic clusters of G residues along one strand (Blackburn & Szostak, 1984). In addition, two repeats of the G-rich sequence form a 3′ overhang terminating the G strand (Zakian, 1989). This overhang endows the DNA with unusual cohesive properties in solution (Lipps, 1980; Oka & Thomas, 1987). Synthetic DNA oligonucleotides containing one or more copies of the G-rich telomeric sequences have the potential to form four-stranded tetraplex structures in solution that are stabilized by cyclic hydrogen bonding between four guanines (Henderson et al., 1987; Sen & Gilbert, 1988, 1990; Williamson et al., 1989; Sunquist & Klug, 1989; Kang et al., 1992; Smith & Feigon, 1992; Gupta et al., 1992). DNA bearing a single copy of the G-cluster can also self-associate, forming a tetrameric structure in which each strand has an identical conformation—a parallel, right-handed helix, with all nucleotides in the anti configuration (Figure 1A) (Sen & Gilbert, 1988, 1990; Lu et al., 1992; Gupta et al., 1992). Two repeats of the 3′-terminal G strand form a folded-back hairpin that dimerizes to create an antiparallel G-quartet structure in which the nucleotides are alternately syn and anti along each strand (Figure 1B) (Sundquist & Klug, 1989; Kang et al., 1992; Smith & Feigon, 1992). It has been shown that a four-stranded parallel G-tetraplex is more stable than the antiparallel bimolecular one (Sen & Gilbert, 1990; Lu et al., 1992). Questions of obvious interest then are the following: By how much is the antiparallel isomer of a G-tetraplex structure energetically favored over the alternative, and what is the basis for the difference? Comparison of a tetrameric structure with a dimer requires taking into account the difference in molecularity. Our approach here is to use synthetic DNAs constrained to adopt parallel or antiparallel dimer conformations as models for the two isomers, and to investigate the thermodynamics of formation of these two G-tetraplex structures. We have synthesized an oligonucleotide ($dG_4T_4G_4$) corresponding to part of the *Oxytricha* telomere repeat of dT_4G_4, and a parallel-stranded analog (3′-dG_4T_2-5′–5′-T_2G_4-3′), by introducing 5′-p-5′ phosphodiester bonds to enforce a parallel orientation of the G-clusters. A thermodynamic analysis of these two DNA molecules is presented here, which shows that the parallel structure is more stable because of concerted enthalpic and entropic contributions. The two molecules differ in the dependence of their T_m on salt concentration, indicating a structural difference between the parallel and antiparallel molecules.

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. The concentration of DNA strands was determined spectrophotometrically at 260 nm and 80 °C (Cantor et al., 1970).

Circular Dichroism (CD) Spectroscopy. DNA solutions were prepared in a 10 mM sodium phosphate buffer, 0.1 mM EDTA, pH 7, and 200 mM NaCl. The DNA samples were heated to 100 °C for 2 min, cooled slowly, and incubated at 20 °C for 24 h. CD spectra were recorded using an AVIV

† This work was supported by National Institutes of Health Grant CA-24101 (N.R.K.). M.L. is supported by National Research Service Award GM-15405 from the NIH.

* Author to whom correspondence should be addressed.
‡ Harvard University.
§ New York University.
‖ Present address: Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

0006-2960/93/0432-598$04.00/0

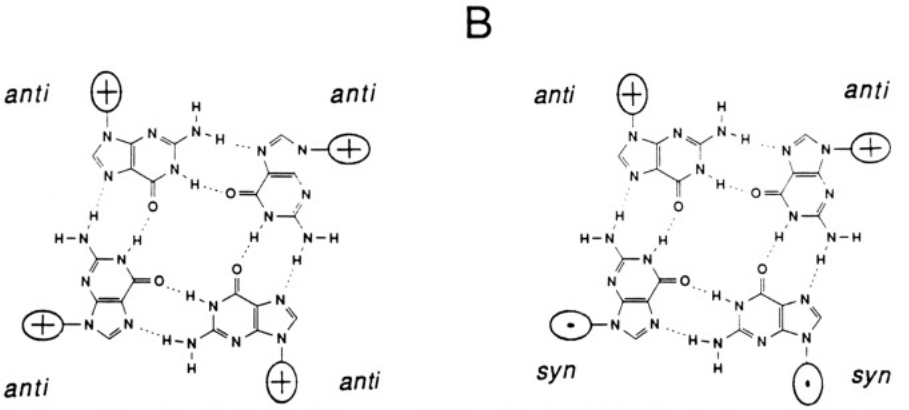

FIGURE 1: Illustration of the base pairing of cyclic G-tetrad in a parallel four-stranded G4 DNA structure of dT_4G_4 in the presence of Na^+ or K^+ (A) (sen & Gilbert, 1988, 1990; Lu et al., 1992; Gupta et al., 1992), and an antiparallel folded-back G-quartet structure of $dG_4T_4G_4$ in the presence of Na^+ (B) (Smith & Feigon, 1992).

Model 60DS CD spectropolarimeter equipped with a programmable, thermoelectrically controlled cell holder. Each spectrum corresponds to an average of five scans from which the buffer background was subtracted. The quantity $\epsilon_L - \epsilon_R$ was calculated by dividing the apparent signal in degrees, θ, by $3.3C$, where C is defined in moles of residues. CD melting curves were recorded at 265 nm for dG_4T_2-T_2G_4 and at 295 nm for $dG_4T_4G_4$. 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.

Differential Scanning Calorimetry (DSC). The total heat of the ordered structure–coil transition of DNA molecules in this study was measured with a Microcal MC-2 differential scanning calometer in 10 mM sodium phosphate buffer, pH 7, 0.1 mM EDTA, and 200 mM NaCl. The total DNA strand concentration was 1×10^{-4} M. Temperatures from 10 to 105 °C were scanned at a heating rate of 45 °C/h. A buffer versus buffer scan was subtracted from the sample scan and normalized for the heating rate. The area of the resulting curve is proportional to the transition heat, and when normalized for the number of moles yields the transition enthalpy ($\Delta H°_{cal}$).

Gel Electrophoresis. DNA solutions were prepared in a 10 mM Tris·HCl buffer/0.1 mM EDTA, pH 7, with or without 50 mM NaCl. DNA samples with a total strand concentration of 1×10^{-4} M were heated to 100 °C for 2 min, cooled slowly, incubated at 20 °C for 24 h, and run on 20% nondenaturing polyacrylamide gels [19:1 acrylamide:bis(acrylamide) ratio] in 0.6× TBE buffer (53.4 mM Tris–borate/1 mM EDTA, pH 8.3) at 4 °C for 17 h (7.5 V cm^{-1}). 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. Gels were stained in a 0.01% Stains-All formamide solution (9:11 formamide:H_2O ratio).

RESULTS AND DISCUSSION

Construction and Characterization of Two Tetraplex Complexes. Recent structural models for the G-tetraplexes formed by dimerization of $dG_4T_4G_4$ define a structure in which the strands run antiparallel, such that the G nucleotides along each strand have alternating syn, anti glycosyl torsion angle pucker (Figure 1B) (Kang et al., 1992; Smith & Feigon, 1992). On the other hand, experiments with single-copy oligonucleotide models of telomeric repeat sequences reveal that these

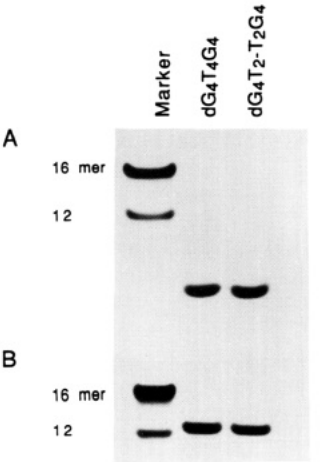

FIGURE 2: Electrophoresis of telomeric oligonucleotides in nondenaturing gels. (A) 0.6× TBE. (B) 0.6× TBE + 50 mM NaCl.

form a four-stranded G-tetrad complex with a helical structure that is different from the X-ray model (Sen & Gilbert, 1988, 1990; Lu et al., 1992; Gupta et al., 1992). This is a right-handed, four-stranded helix, in which all nucleotides are anti, and the sugars have the C2′ endo conformation (Figure 1A) (Gupta et al., 1992). Preliminary analysis of the stability of the latter indicates that it is extremely stable in solution, relative to the folded-back dimer structure (Sen & Gilbert, 1990; Lu et al., 1992). However, it has not been possible to compare the stability of these structures quantitatively, because in the one case the complex formed is a tetramer of identical strands while in the other the structure is a dimer of hairpins. To carry out a direct thermodynamic comparison of these forms, we decided to construct both tetraplex structures as nearly as possible in the same background. Accordingly, we synthesized the strand $dG_4T_4G_4$, which forms an antiparallel-stranded structure in the presence of sodium or potassium. Next we constructed the sequence dT_2G_4 in the usual manner, and added 5′-p-5′ dT_2G_4 to this. We refer to this latter strand as dG_4T_2-T_2G_4 to distinguish it from the normal sequence. Gel electrophoresis of these two DNA strands is shown in Figure 2. In the absence of salt, each oligonucleotide migrates as a monomer species (Figure 2A). In the presence of Na^+, both oligomers have electrophoretic mobility equivalent to a dimer (see Figure 2B). Since CD has proven to be a very sensitive technique for probing the structure of telomeric models (Henderson et al., 1990; Hardin et al., 1991; Guo et al., 1992; Lu et al., 1992), we measured the CD spectra of the two species to determine whether the two sequences differ in conformation. The answer is shown in Figure 3, which clearly

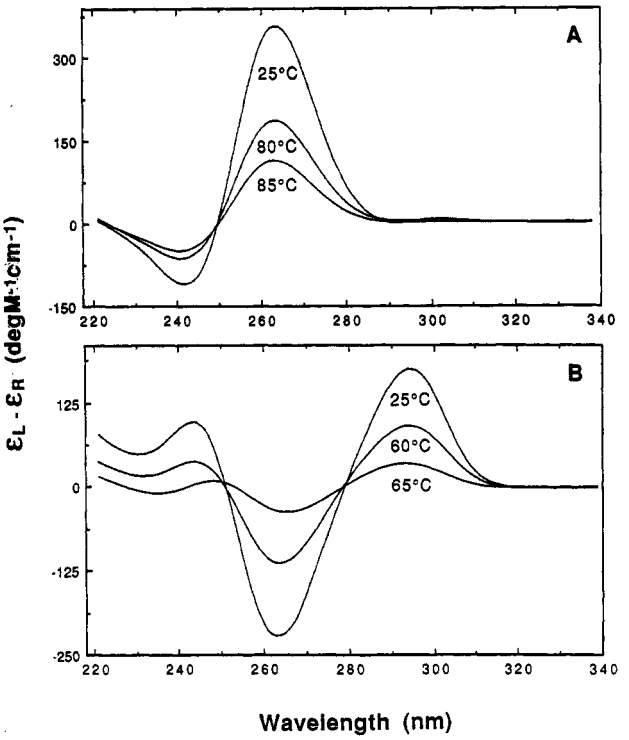

FIGURE 3: Temperature-dependent CD spectra of dG_4T_2-T_2G_4 (A) and $dG_4T_4G_4$ (B) in 10 mM sodium phosphate buffer, pH 7, 0.1 mM EDTA, and 200 mM NaCl. The total DNA strand concentration is 1×10^{-4} M. Samples were preequilibrated at 4 °C for 1 h.

indicates a major difference in the form of the spectra of these molecules in the presence of sodium. The spectrum of $dG_4T_4G_4$ in panel B corresponds with that assigned previously to the tetraplex formed by dimerizing two hairpins (Sundquist & Klug, 1989; Henderson et al., 1990; Guo et al., 1992; Kang et al., 1992; Smith & Feigon, 1992). The spectrum of dG_4T_2-T_2G_4, on the other hand, is identical to that for the parallel-stranded tetramer formed from four strands of dT_4G_4, for example (Lu et al., 1992). The latter is more stable to thermal unfolding, as also indicated in Figure 3. In the next section, the thermodynamic profiles of these two species are studied in more detail.

Thermodynamic Comparison of the Two Tetraplex Structures. The thermodynamics of formation of G-tetraplex structures such as those illustrated in Figure 1 provide information about the interactions responsible for stabilizing the structure, and can also reveal structural differences. If the process of forming the dimers is fully cooperative, the system can be analyzed in terms of the van't Hoff relation:

$$1/T_m = (R \ln C_T + \Delta S)/\Delta H_{VH}$$

where T_m is the transition temperature, R is the gas constant, C_T is the total concentration of single strands, ΔS is the entropy, and ΔH_{VH} is the van't Hoff enthalpy. Figure 4 shows a plot of $1/T_m$ as a function of the single-strand concentration, according to this equation. Both species give linear plots, from which we can extract the enthalpy change $\Delta H°_{VH}$, the entropy change $\Delta S°S_{VH}$, and thus the free energy of forming the dimer, $\Delta G°$. The thermodynamic profiles are shown in Table I. This analysis assumes all or none formation of the tetraplex structures in each case. The thermodynamic parameters indicate that the standard free energy ($\Delta G°$) of forming the parallel-stranded dimer is 5 kcal/mol below that for formation of the antiparallel structure, in agreement with the observation from CD thermal unfolding experiments (Figure 3). This difference reflects roughly equal contributions

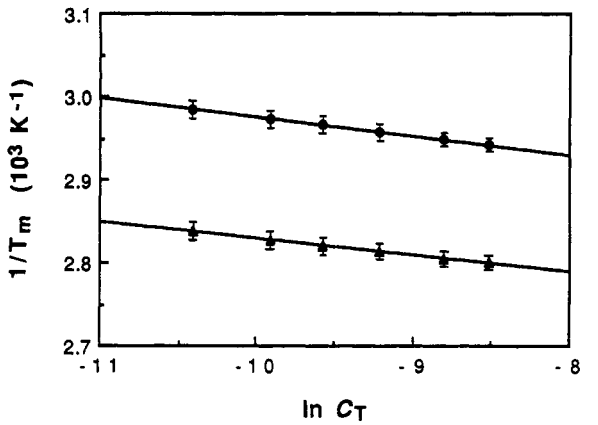

FIGURE 4: Dependence of the transition temperature of dG_4T_2-T_2G_4 (filled triangles) and $dG_4T_4G_4$ (filled circles) in 10 mM sodium phosphate buffer, pH 7, 0.1 mM EDTA, and 200 mM NaCl.

Table I: Thermodynamic Parameters of Tetraplex Formation by dG_4T_2-T_2G_4 and $dG_4T_4G_4$

strand	T_m[a] (°C)	$\Delta H°_{cal}$[b] (kcal/mol)	n	$2\Delta H°_{cal}$ (kcal/mol)	$\Delta H°_{VH}$[c] (kcal/mol)	$\Delta S°_{VH}$[c] [cal/(mol·deg)]	$\Delta G°$[d] (kcal/mol)
dG_4T_2-T_2G_4	82.4	−58.2	2	−116	−99	−260	−21
$dG_4T_4G_4$	65.0	−45.1	2	−90	−86	−237	−16

[a] Calculated for 10^{-4} M single-strand concentration. [b] Average of five DSC scans, and calculated for the total single-strand concentration. [c] Calculated from $1/T_m$ versus ln C_T parameters with the molecularity $n = 2$. [d] Calculated at 25 °C assuming $\Delta C_p = 0$.

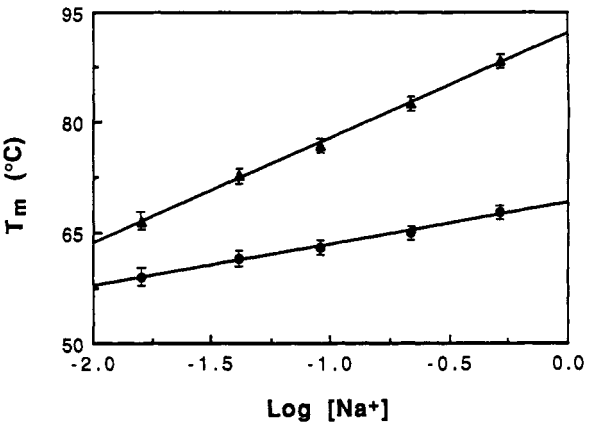

FIGURE 5: Salt dependence of the transition temperature for dissociation of each dimeric structure in 10 mM sodium phosphate buffer, pH 7, and 0.1 mM EDTA: $dG_4T_4G_4$ (filled circles); dGT_2-T_2G_4 (filled triangles).

from the enthalpy ($\Delta H°_{VH}$) and the entropy terms ($\Delta S°$), which are in the same direction (Table I). This means that the bonding and stacking interactions in the parallel structure are more stable while the entropy of this structure also is more favorable than that of the antiparallel one. Binding of ions and water to ordered nucleic acids provides a major contribution to the entropy of forming these structures (Zieba et al., 1991). To assess the ionic contribution, we determined the salt dependence of the midpoint transition temperatures, T_m, of the two species. As can be seen in Figure 5, there is a surprisingly large difference in the slope of T_m vs log [Na^+] plot, indicating that there is a fundamental difference in how sodium ions interact with the two complexes. Since the alternate G-tetrads in the two structures differ in nucleotide conformation, and the relative orientation of each layer, the grooves of these molecules are likely to differ in their geometry and hence their interaction with ligands such as sodium ions.

In addition, the sodium binding site presumed to exist within the tetraplex is expected to be sensitive to the local geometry of the G oxygens (Sundquist & Klug, 1989). The phosphates in the two structures might also differ in charge, with the parallel form having a lower apparent ionic charge than the antiparallel one.

Next we determined the heat of formation of the dimers directly by DSC, as described under Materials and Methods. This gives a model-independent measure of the transition enthalpy, in contrast to the van't Hoff values obtained so far. The results are in close agreement with the van't Hoff values, as shown in Table I. Hence, formation of tetraplex from two dimers is very nearly a fully cooperative reaction, with only minor populations of any intermediates.

What are the biological implications of these results? In the overhang of a single chromosome end, two copies of the cluster are normally available, with the possibility of two more after replication. With no additional "free" copies—uncomplemented by a second strand—this tends to argue that the in vivo association will involve primarily antiparallel strands. It is possible that additional copies become available during replication or some other cellular process (Hardin et al., 1991), so that the potential for forming parallel structures exists. In the presence of damage to the ends of a chromosome, however, in principle either the parallel or the antiparallel structures can form near the ends of chromosomes, depending on the nature of the damage induced (Lu et al., 1992). In this study, we have imposed a constraint on the molecularity of the two structures, allowing us to measure the thermodynamic differences between the two. Since the antiparallel structure requires dimerization rather than slower four-strand association, there is an additional kinetic barrier that prevents formation of the more stable parallel structure. Nevertheless, the latter is the most stable possible state for the telomeric sequence repeats, and since the rates of these processes are strongly dependent on the relative concentrations of Na^+ and K^+, it is a structure that may arise as an intermediate in repair of damage to the ends.

REFERENCES

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. (1992) *Biochemistry 31*, 2451–2455.

Gupta, G., Garcia, A. E., Guo, Q., Lu, M., & Kallenbach, N. R. (1992) *Biochemistry* (submitted for publication).

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.

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.

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

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

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

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

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

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

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

Zieba, K., Chu, T. M., Kupke, D. W., & Marky, L. A. (1991) *Biochemistry 30*, 8018–8026.