

Self-Assembly of DNA Oligomers into High Molecular Weight Species[†]

Tie-Yue Dai, Stephen P. Marotta, and Richard D. Sheardy*

Department of Chemistry, Seton Hall University, 400 South Orange Avenue, South Orange, New Jersey 07079-2694

Received November 4, 1994; Revised Manuscript Received January 6, 1995[®]

ABSTRACT: Thermal denaturation, gel electrophoresis, and circular dichroism methods were used to characterize DNA oligomers possessing one or two segments of four contiguous G bases in order to investigate their environmentally dependent conformational properties. The sequences of the oligomers studied were the following: HP1-T series, C₄T₄G₄T_{5–8}; HP1-TG series, C₄T₄G₄T_{1–4}G₄. In NaCl at concentrations up to 200 mM, the melting profiles of these oligomers are characterized by single inflection points whose T_m values are independent of DNA concentration. In addition, these oligomers run as single bands in polyacrylamide gels under those same conditions as well as in 100 mM K⁺ or 20 mM Mg²⁺. These data suggest that these oligomers exist as intramolecular hairpins comprised of four G:C base pairs in the stems, loops of four T bases, and 3'-overhangs of T_{5–8} or T_{1–4}G₄. In the presence of 100 mM K⁺ plus 20 mM Mg²⁺, however, gel electrophoresis indicates that oligomers of the HP1-T series exist as equilibria between parent hairpins and four-stranded structures (i.e., quadraplexes). Quadraplex formation for any member of the HP1-T series requires unfolding of the hairpin, exposing the G₄ segment prior to quadraplexation. Members of the HP1-TG series self-assemble into multistranded species of high molecular weight in the presence of 100 mM K⁺ plus 20 mM Mg²⁺. For this series of oligomers, the data suggest that these higher order species arise from successive additions of parent oligomer to an initially formed quadraplex. Since the self-assembly is not observed with K⁺ or Mg²⁺ alone, these cations behave in a synergistic manner in the formation and/or stability of the supermolecular self-assemblies.

The formation of four-stranded DNA structures (i.e., quadraplexes) by guanine-rich DNA oligomers has received considerable interest in the recent literature. The use of specifically designed DNA oligomers with sequences similar to those found in natural telomeric DNA affords a means to model the architecture and thermodynamics of these DNA structural motifs. For example, several studies have implicated G:G Hoogsteen pairing involving N-7 of guanine in quadraplex formation (Sen & Gilbert, 1988; Williamson et al., 1989; Kanng et al., 1992; Sundquist & Klug, 1989; Panyutin et al., 1990; Henderson et al., 1990; Wang & Patel, 1992). The orientation of the substituent strands in the quadraplex has also been of considerable interest. Parallel-stranded structures have been noted for sequences such as TG₃T (Jin et al., 1992) and T₄G₄ (Lu et al., 1992; Wang & Patel, 1992). In addition, the CD spectrum of G₄TG₄ under quadraplex formation conditions is also consistent with a parallel structure (Balagurumoorthy et al., 1992). NMR studies on the G₄T₄G₄ quadraplex suggested a structure with alternating parallel and antiparallel strands formed from the complexation of two fold-back hairpins with T loops running diagonally across the G-tetrads (Smith & Feigon, 1992, 1993). Fully antiparallel intermolecular duplexes of fold-back structures have also been observed (Sundquist & Klug, 1989). These investigators suggested that the T residues may serve as spacers to allow hairpin formation and thus reduce steric interactions between the quadraplex and adjacent duplex regions. Venczel and Sen (1993) have noted novel

structural features for both the parallel and antiparallel forms of a complex telomeric sequence.

Thermodynamic characterization of sequences such as G₂T₄G₂ (Jin et al., 1990; Wang et al., 1991), TG₃T and TG₃T₂G₃T (Jin et al., 1992), and G₄T₄G₄ and G₄T₂-5'-5'-T₂G₄ (Lu et al., 1993) has also been carried out. For example, the standard free energies of quadraplex formation at 25 °C have been determined via differential scanning calorimetry for TG₃T and TG₃T₂G₃T to be -6.9 kcal/mol of quadraplex and -20.2 kcal/mol of quadraplex, respectively. In addition, a recent report from Kallenbach's group demonstrated that the stability for quadraplexes arising from T_nG₄ ($n = 1 - 8$) decreases as n increases (Guo et al., 1993).

Quadraplex structures for G-rich DNA oligomers are stabilized by monovalent cations, and the order of ion binding preference is K⁺ > Na⁺ > Cs⁺ > Li⁺ (Williamson et al., 1989; Guschlbauer et al., 1990). Quadraplex structures are also stabilized by divalent ions (Hardin et al., 1991, 1992; Williamson, 1994). With both monovalent and divalent cations, the stabilization decreases with decreasing size, suggesting a model in which the ion which fits best into the cavity created by the G-tetrad stabilizes the most. A recent study, however, suggests that electronic effects may also be involved (Ross & Hardin, 1994).

Studies on the *Tetrahymena* sequence (dT₂G₄)₄ revealed an equilibrium of monomolecular, bimolecular, and tetramolecular species. This equilibrium is influenced by the nature and concentration of the cation, the concentration of the DNA oligomer, and the temperature (Hardin et al., 1991; Williamson, 1994). Hardin et al. (1992) have also reported cation-dependent transitions, and hence equilibria, between the Watson-Crick hairpin form and the four-stranded quadraplex form of CGCG₃CGC. Formation of the quadraplex form for this oligomer is therefore predicated on

[†] A preliminary account of this work was presented in part at the 207th National Meeting of the American Chemical Society (Dai et al., 1994b).

* Address correspondence to this author.

[®] Abstract published in *Advance ACS Abstracts*, March 1, 1995.

denaturation of the hairpin. Thus, the equilibrium is determined by the balance of free energies required for hairpin denaturation and released upon quadruplex formation. We have designed DNA oligomers possessing $C_4T_4G_4$ segments, capable of hairpin formation, abutted to the 5' termini of T_{5-8} and $T_{1-4}G_4$ segments. Thermal denaturation, gel electrophoresis, and circular dichroism studies were carried out in order to investigate the environmentally dependent conformational states of these molecules. The results of these studies demonstrate that in the presence of 100 mM Na^+ , 100 mM K^+ , or 20 mM Mg^{2+} , all oligomers exist predominantly as intramolecular hairpins. In the presence of *both* 100 mM and 20 mM Mg^{2+} , those oligomers with single internal G_4 segments exist as equilibria between parent hairpins and four-stranded quadruplex structures. However, the DNA oligomers which possess two G_4 segments display remarkable tendencies to form regular higher order, multistranded species of very high molecular weights.

MATERIALS AND METHODS

DNA Synthesis and Preparation. The DNA oligomers were synthesized on a 1.0 μ mol scale via the phosphoramidite method (Caruthers, 1982) on an Applied Biosystems 380B DNA synthesizer with purification by trityl select reverse phase HPLC as previously reported (Sheardy, 1988). After the second HPLC purification, the oligomers were exhaustively dialyzed vs water and then lyophilized to dryness. Product purity was verified via analytical HPLC and native and denaturing polyacrylamide gel electrophoresis. Prior to any characterization, lyophilized DNA samples were reconstituted in the appropriate buffer, heated to 80 °C for 2 min, allowed to slowly cool to room temperature, and then stored for 48 h at 4 °C. For the experiments reported here, all DNA samples were freshly prepared from the original lyophilized material. Extinction coefficients at 25 °C for the single strands, ϵ ($L \text{ mol}^{-1} \text{ cm}^{-1}$), were determined using the method described in Fasman (1975): HP1-T5, 8388; HP1-T6, 8372; HP1-T7, 8357; HP1-T8, 8345; HP1-T1G4, 8870; HP1-T2G4, 8827; HP1-T3G4, 8789; and HP1-T4G4, 8755.

UV/VIS Spectroscopy. The UV spectrum of each oligomer was recorded at 20 °C under a variety of buffer conditions. For thermal denaturation studies, samples were prepared in phosphate buffer (10 mM phosphate, 0.1 mM EDTA, pH 7.0) plus NaCl added to the desired concentration. For these studies, the temperature was ramped from 20 to 95 °C at 0.3 °C/min. The spectra at 95 °C were recorded before slow cooling back to 20 °C. The spectra were recorded again at 20 °C for comparison to the original 20 °C spectra. If the absorbance differed by more than 5% between the two 20 °C spectra for any sample, the melting data were discarded, and the experiment was repeated with a new sample. The acceptable melting data were transferred to a PC for analysis according to Marky and Breslauer (1987). DNA oligomer single-strand concentrations at 25 °C were determined by extrapolation of the upper base lines of the melting profiles back to 25 °C to determine the absorbances at that temperature and then dividing by the appropriate calculated extinction coefficients.

Gel Electrophoresis. Nondenaturing polyacrylamide gel electrophoreses [19:1 acrylamide:bis(acrylamide)] at different total percent acrylamide concentrations of the various oligomers were carried out at 4 ± 1 °C at 8 V/cm with a

T Series

HP1-T5:	$C_4T_4G_4T_5$
HP1-T6:	$C_4T_4G_4T_6$
HP1-T7:	$C_4T_4G_4T_7$
HP1-T8:	$C_4T_4G_4T_8$

TG Series

HP1-T1G4:	$C_4T_4G_4T_1G_4$
HP1-T2G4:	$C_4T_4G_4T_2G_4$
HP1-T3G4:	$C_4T_4G_4T_3G_4$
HP1-T4G4:	$C_4T_4G_4T_4G_4$

FIGURE 1: DNA oligomers considered in this report.

running buffer of 100 mM Tris–borate (pH 8.0) with $MgCl_2$, NaCl, or KCl added to the desired concentrations. The acrylamide was polymerized in the same buffer as the running buffer for each different experiment. Visualization of the DNA bands was carried out by treatment with Stainsall (BIO-RAD, Richmond, CA). Permanent records of the gels were made using Kodak Electrophoresis Duplicating Paper (Eastman, Rochester, NY). For a typical oligomer preparation for native PAGE, the particular oligomer was reconstituted in 18 μ L of $1 \times$ TBE buffer with the corresponding cation at the concentration of the running buffer for that experiment. After equilibration for 48 h at 4 °C, 2 μ L of $10 \times$ tracking dye [Xylene Cyanole XX, Bromophenol Blue (BIO-RAD)] in $1 \times$ TBE buffer was added to each sample before being loaded into the well; 10–15 μ g of DNA oligomer was added per lane in order to effect visualization with Stains-all.

UV Cross-Linking. DNA samples for UV cross-linking experiments were prepared in the same fashion as for native gel electrophoresis. After equilibration, the samples were irradiated at 254 nm for 30 min using a Sylvania G8T5 lamp (8 W, Germicidal). The samples were then subjected to electrophoresis on a denaturing 16% polyacrylamide gel (100 mM Tris–borate, pH 8.0, 7 M urea, 65 °C).

Circular Dichroism Studies. The CD spectra of select oligomers were determined in 100 mM Tris–borate buffer (pH 8.0) with 100 mM K^+ and 20 mM Mg^{2+} at 25 °C with an AVIV 60DS CD spectropolarimeter. The concentration of each DNA oligomer, determined as above, was 8.6×10^{-5} M in single strand.

RESULTS

The oligomers designed for this study are shown in Figure 1. Each oligomer has a 5'– $C_4T_4G_4$ –3' segment. Abutted to the 3' terminus of the $C_4T_4G_4$ segment are T_{5-8} tails (for the HP1-T series) or $T_{1-4}G_4$ tails (for the HP1-TG series). It was predicted *a priori* that the $C_4T_4G_4$ segment would form a stable hairpin in each oligomer under non-quadruplex-forming conditions (i.e., low concentrations of Na^+ or K^+). Under quadruplex-forming conditions (i.e., high concentrations of Na^+ or K^+), members of the HP1-T series could only form quadruplexes by unfolding of that hairpin, thereby exposing the internal G_4 segment. However, quadruplex

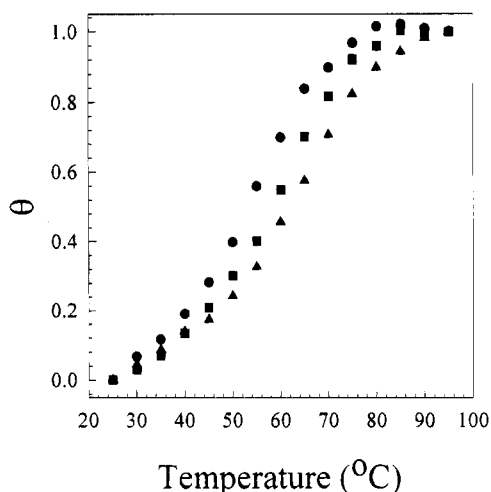


FIGURE 2: Thermal denaturation profiles of HP1-T7 in 10 mM phosphate buffer, pH 7.0, at 15 mM NaCl (●), 50 mM NaCl (■), and 200 mM NaCl (▲). The relative fraction of absorbance change, θ , is calculated as $[(A - A_i)/(A_f - A_i)]H_R$, where A is the absorbance at a particular temperature, A_i and A_f are the absorbances at the initial and final temperatures (i.e., 20 and 95 °C), respectively, and H_R is the relative hyperchromicity. H_R is calculated as the hyperchromicity [$H = (A_f - A_i)/A_i$] of the particular sample divided by the hyperchromicity of the sample in 15 mM NaCl. Similar profiles were obtained with all oligomers of Figure 1 at all NaCl conditions.

formation for the members of the HP1-TG could simply occur at the exposed terminal G₄ segment. The participation of the internal G₄ segment in any quadruplex formation in either series would ultimately depend upon the relative stability of the quadruplex to the hairpin. The conformational state available to any of these oligomers will, thus, have a strong dependence on environmental conditions. Therefore, a series of studies were initiated to characterize their conformational properties under a variety of conditions.

The first set of studies was the determination of the melting profiles of the members of the HP1-T series as a function of NaCl concentration. The nature of these melting profiles can be used to determine if the oligomer is single or double stranded. For duplex DNA, a monophasic melt is typical for a single species while a bi- or multiphasic melt may indicate, for example, the presence of several species or intramolecular triplex. Furthermore, the DNA concentration dependence of the T_m can be used to discriminate between unimolecular transitions and multimolecular transitions. Figure 2 shows the melting profiles of HP1-T7 in buffer with 15, 50, and 200 mM NaCl. As can be seen, the shape of the melting profile and the extent of hyperchromicity (ca. 30%) are quite typical for double-stranded DNA. All oligomers of the HP1-T series in phosphate buffer at DNA concentrations ranging from 6×10^{-5} to 3×10^{-3} M (in single strand) and NaCl concentrations ranging from 15 to 200 mM had similar melting profiles with single inflection points, suggesting that the thermal transitions were monophasic. In addition, the T_m values obtained were independent of DNA concentrations over that range of DNA concentrations.

The melting data indicate that each thermally induced denaturation in the presence of NaCl is unimolecular and arises from a single predominant species present under these conditions, a conclusion borne out by the gel electrophoresis results discussed below. That being the case, analyses of the transition profiles to obtain thermodynamic parameters

Table 1: Thermodynamic Parameters for the HP1 and HP1-T Series^a

oligomer	ΔH°	ΔS°	ΔG°_{298}
In 15 mM NaCl			
HP1-T5	33.3	100	3.4
HP1-T6	32.4	98	3.3
HP1-T7	31.6	95	3.2
HP1-T8	34.3	103	3.5
In 200 mM NaCl			
HP1-T5	31.4	93	3.9
HP1-T6	31.2	92	3.9
HP1-T7	31.4	92	3.9
HP1-T8	34.2	101	4.1

^a ΔH° and ΔG° values are given in kilocalories per mole and ΔS° values in entropy units. The errors in these calculations are typically on the order of $\pm 5\%$.

according to the following equations (applicable to a two-state, unimolecular process) were carried out:

$$\Delta H_{VH} = 4RT_m(\delta\alpha/\delta T)_{T=T_m} \quad (1)$$

$$\Delta G^\circ = \Delta H^\circ(1 - T/T_m) \quad (2)$$

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \quad (3)$$

where ΔH_{VH} is the van't Hoff enthalpy (which is assumed to be equal to ΔH°), T_m is the temperature at $\alpha = 0.5$ (α is the fraction of single strands), ΔG° is the free energy of the transition at a temperature of T , and ΔS° is the entropy change for the transition (Marky & Breslauer, 1987). Using a computer program to determine the upper and lower base lines of the melting transition profiles to obtain α and T_m , ΔH_{VH} was then calculated directly from these profiles (eq 1) and ΔG° and ΔS° were subsequently determined (eqs 2 and 3, respectively). The results for thermal denaturations at 15 and 200 mM NaCl are given in Table 1. As can be seen, the oligomers have similar thermal denaturation free energies at identical NaCl concentrations and the oligomers are slightly more stable ($\Delta\Delta G \approx 0.5$ kcal/mol) at the higher NaCl concentration.

In order to determine the number of species present under our various salt conditions, a series of gel electrophoresis experiments were carried out. As noted above, the DNA samples and the gel were prepared in the running buffer of Tris-borate (pH 8.0) plus added cation (Mg^{2+} , Na^+ , and/or K^+) at select concentrations. Na^+ and K^+ were chosen since these cations are known to stabilize quadruplexes formed from G-rich DNA oligomers [for example, see Hardin et al. (1991, 1992), Sen and Gilbert (1992), Balagurumoorthy et al. (1992), Lu et al. (1992), and Guo et al. (1992)]. The reported concentrations of NaCl or KCl used, methods to prepare DNA quadruplexes, and the electrophoretic conditions vary in the literature. We chose to carry out our initial studies at 100 mM Na^+ , 100 mM K^+ , or 20 mM Mg^{2+} . The concentrations of 100 mM for Na^+ and K^+ should be sufficient for quadruplex formation based upon literature reports [for example, see Hardin et al. (1991, 1992) and Balagurumoorthy et al. (1992)]. Mg^{2+} was chosen since we normally run nondenaturing gels in the presence of Mg^{2+} for enhanced duplex stabilities of the oligomers of interest.

In the presence of 20 mM Mg^{2+} , 100 mM Na^+ , or 100 mM K^+ , all oligomers, prepared as noted under Materials and Methods, run essentially as single species through the

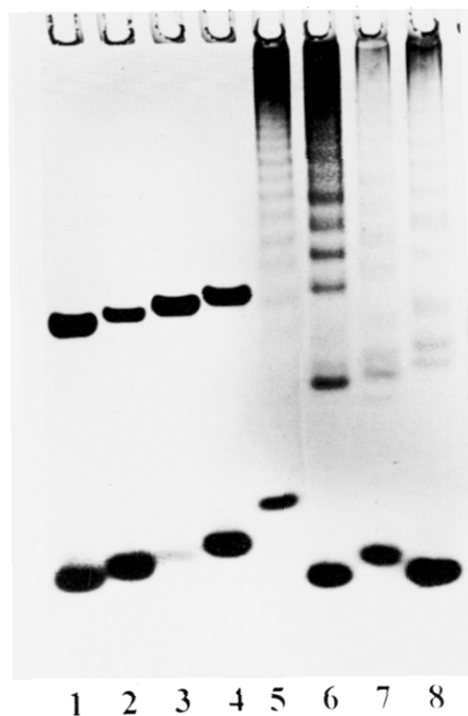


FIGURE 3: Polyacrylamide gel electrophoresis (12%) at 4 °C in TBE buffer plus 20 mM Mg^{2+} with 100 mM K^+ . Lane 1, HP1-T5; lane 2, HP1-T6; lane 3, HP1-T7; lane 4, HP1-T8; lane 5, HP1-T1G4; lane 6, HP1-T2G4; lane 7, HP1-T3G4; lane 8, HP1-T4G4.

gel (gels not shown). Under all three salt conditions, some higher order species are observed for HP1-T1G4 and HP1-T2G4, but the corresponding bands were very faint relative to the major bands. The similarity in electrophoretic mobilities of the HP1-TG series oligomers to those of the HP1-T series oligomers indicates a similarity in structure. Hence, all oligomers of Figure 1 exist as duplexed hairpins under these conditions. The gel electrophoresis was repeated in 100 mM K^+ plus 20 mM Mg^{2+} and gave the electrophoretic pattern shown in Figure 3. As can be seen, the HP1-T series show only two distinct bands, but a ladder of bands is obtained for the HP1-TG series. What is particularly noteworthy about these data is that higher order structure formation is observed in the presence of *both* Mg^{2+} and K^+ . As noted above, 100 mM K^+ alone or 20 mM Mg^{2+} alone was insufficient for higher order structure formation (at the same [DNA]). Hence, a synergistic effect is noted when both cations are present.

The bands of highest mobilities correspond to the intramolecular hairpin observed when the gels are run in Na^+ , K^+ , or Mg^{2+} alone. The mobilities of the Mg^{2+}/K^+ -induced bands for the HP1-T series oligomers are consistent with species based upon four-stranded structures. Perhaps the most striking feature of the gel run in 20 mM Mg^{2+} with 100 mM K^+ is the indication of discreet bands of very high molecular weight DNA for oligomers of the HP1-TG series. Hence, these oligomers self-assemble into supermolecular structures. It is important to note that only those oligomers with two G_4 segments form the supramolecular self-assemblies. The gel patterns shown in Figure 3 suggest that the self-assemblies arise from successive additions of a parent oligomer to an initially formed quadruplex. To address this possibility, plots of log relative mobility vs N (band number) were constructed for the HP1-TG series using only the major Mg^{2+}/K^+ -induced bands and assigning the highest mobility

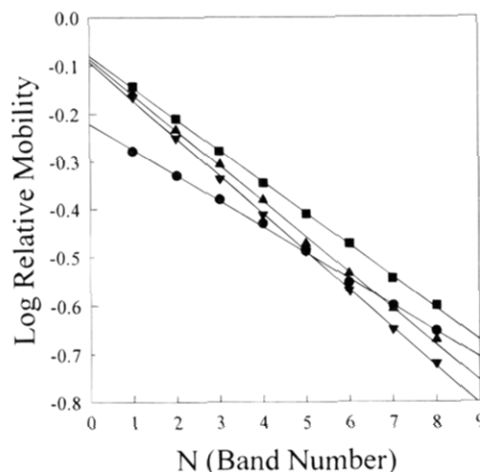


FIGURE 4: Plots of log relative mobility vs N (induced band number) for members of the HP1-TG series run in 12% polyacrylamide gels at 4 °C in TBE buffer with 20 mM Mg^{2+} plus 100 mM K^+ . Relative mobility was calculated as the distance traveled by the particular band divided by the distance traveled by a 12-mer marker band. In addition, we designated the first Mg^{2+}/K^+ -induced band as $N = 1$. HP1-T1G4 (●); HP1-T2G4 (■); HP1-T3G4 (▲); and HP1-T4G4 (▼).

band among these to be $N = 1$. As can be seen in Figure 4, these plots resulted in high correlation ($R^2 > 0.998$) linear fits. These data are consistent with the above notion of successive additions. It should be noted that log relative mobilities for the parent bands (arising from the hairpinned duplex) do not lie on the respective least-squares lines.

In order to learn more about the molecularity of self-assembly for the HP1-TG series, UV cross-linking experiments on HP1-T2G4 were carried out. The products formed from the irradiation at 254 nm of HP1-T2G4 in the presence of 20 mM $Mg^{2+}/100$ mM K^+ were separated on a 12% denaturing polyacrylamide gel. As can be seen in Figure 5A, at least five bands were observed corresponding to monomer, dimer, trimer, tetramer, and pentamer. Such products could only arise from T-T cross-links in multi-stranded structures (Lu et al., 1992). A plot of log mobility vs X (number of cross-linked oligomers) gives rise to a linear fit (Figure 5B) and indicates that the faintest band (assigned $X = 5$) does in fact arise from five cross-linked strands and not six (see Discussion).

The Mg^{2+}/K^+ gels were then repeated at different acrylamide concentrations in order to construct Ferguson plots according to

$$\log M = \log M_0 - K_R C \quad (4)$$

where M is the mobility (calculated as the distance traveled by the band divided by the total electrophoretic time) at gel concentration C (percentage of acrylamide), M_0 is the free mobility of the DNA oligomer (i.e., at $C = 0\%$), and K_R is the retardation coefficient (Rodbard & Chrambach, 1971). K_R , obtained from the slope of the Ferguson plot, is qualitatively related to the shape of the migrating species, and M_0 (from the y-intercept) is thought to vary with the charge density of the migrating species.

The Ferguson plots for the HP1-T and HP1-TG series for gels run in 20 mM Mg^{2+} with 100 mM K^+ were constructed and displayed linear relationships between $\log M$ and C for all oligomers. The line parameters calculated from least-squares analyses are given in Table 2. As can be seen, the

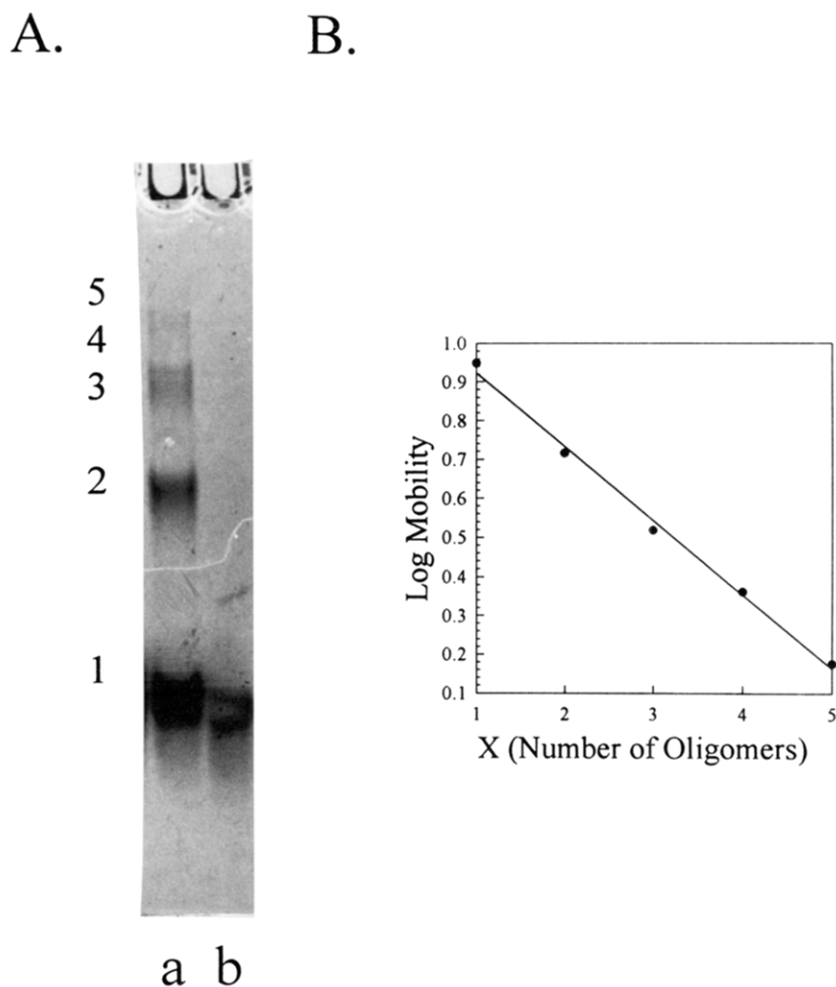


FIGURE 5: (A) Denaturing polyacrylamide gel electrophoresis of the products obtained by the irradiation of HP1-T2G4 at 254 nm in the presence (lane a) or absence (lane b) of 20 mM Mg^{2+} and 100 mM K^+ . (B) Plot of log mobility vs X (number of oligomers) for the bands arising from UV cross-linking.

Table 2: Line Parameters from Ferguson Analysis of PAGE of the Oligomers of This Report^a

oligomer	parent band		first Mg^{2+}/K^+ induced band	
	K_R	$\log M_0$	K_R	$\log M_0$
12-mer	0.0584	1.61		
HP1-T5	0.0425	1.48	0.0767	1.63
HP1-T6	0.0434	1.49	0.0803	1.67
HP1-T7	0.0443	1.49	0.0829	1.69
HP1-T8	0.0449	1.49	0.0858	1.71
HP1-T1G4	0.0513	1.53	0.0649	1.58
HP1-T2G4	0.0432	1.49	0.0681	1.61
HP1-T3G4	0.0440	1.48	0.0692	1.61
HP1-T4G4	0.0415	1.46	0.0801	1.76

^a Values reported are $\pm 2\%$.

bands assigned as intramolecular hairpins (parent band), in general, have similar retardation coefficients and free mobilities, indicating that they have similar shapes and charge densities as they pass through the gel. In addition, the K_R and $\log M_0$ values are quite different from our linear 12-mer marker. The bands corresponding to the first member of the induced higher ordered structures (first Mg^{2+}/K^+ induced bands) have larger frictional coefficients and higher free mobilities than the parent bands. In addition, for this series of bands, the K_R and $\log M_0$ values increase with increasing molecular weight of the parent oligomer. The observed differences in retardation coefficients and free mobilities for the first member of the higher ordered species

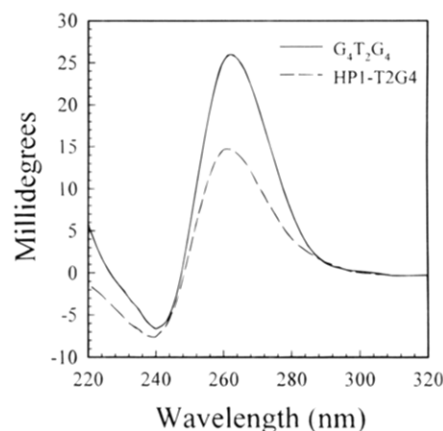


FIGURE 6: Comparison of the CD spectra of HP1-T2G4 and $G_4T_2G_4$ in 100 mM TBE, pH 8.0, with 100 mM K^+ and 20 mM Mg^{2+} .

indicate differences in shapes and/or charge densities from the parent molecules.

Circular dichroism determinations on four-stranded structures can be used to discriminate between parallel and antiparallel strand orientation (Balagurumoorthy et al., 1992; Guo et al., 1993; Lu et al., 1993; Giraldo et al., 1994). Figure 6 shows the CD spectra of HP1-T2G4 (i.e., $C_4T_4G_4T_2G_4$) and $G_4T_2G_4$ in 100 mM TBE buffer with 100 mM K^+ and 20 mM Mg^{2+} . Both samples give CD spectra with peaks at 260 nm and shallow troughs at 240 nm, consistent with

parallel strand orientation (Balagurumoorthy et al., 1992; Guo et al., 1993; Lu et al., 1993).

DISCUSSION

Due to the indiscrimination in the types of associations available to G bases, members of the HP1-TG series could form a variety of structures such as hairpinned duplex, intramolecular triplex, or inter- and/or intramolecular quadruplexes. It is also likely that the members of the HP1-T series could also form quadruplexes, by virtue of their internal G₄ segments, at high Na⁺ or K⁺ concentrations. The types of species present would depend upon the DNA concentration, the sequence of the DNA, and the nature and concentration of added cations (Hardin et al., 1991, 1992; Williamson, 1994). It should be pointed out, though, that one of the goals with this set of oligomers was the determination of the nature of the species present under particular conditions.

The thermal denaturation data for the HP1-T series oligomers indicate behavior as unimolecular species in NaCl at concentrations up to 200 mM. A hairpinned duplex is the most likely structure for each oligomer under these conditions. The similarity of the electrophoretic patterns for the HP1-TG oligomers to the corresponding HP1-T oligomers in 100 mM Na⁺ indicates that the HP1-TG oligomers also exist as hairpinned duplexes. Thus, in the presence of NaCl, each oligomer of Figure 1 possesses a duplex domain in the form of a hairpin formed from the C₄ segment with the *internal* G₄ segment creating a T₄ loop and T tails or TG tails of varying lengths at the 3' end. The incorporation of only four T bases in the stem of the potential hairpin was chosen by design to favor the intramolecular hairpin over the potential intermolecular association to bulged duplex, a process favored by longer stem lengths. Each oligomer of the HP1-TG series could also form a hairpinned duplex between the C₄ segment and the terminal G₄ segment, creating T₄G₄T₁₋₄ loops. However, previous studies investigating the effect of loop size on hairpin stability noted that increasing loop size resulted in decreasing hairpin stability (Panar et al., 1992; Rentzeperis et al., 1993). Hence, the most stable hairpin would be that with the four base stem and four base loop as suggested above.

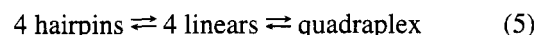
The observation of single bands in the electrophoresis of these oligomers at 100 mM Na⁺, 100 mM K⁺, or 20 mM Mg²⁺ is consistent with the presence of a single species for each oligomer. These NaCl and KCl concentrations are higher than those used to induce quadruplex formation in such oligomers as d(CGCG₃GCG), which formed a quadruplex at 40 mM K⁺ (Hardin et al., 1992), or d(G₄T₄G₄), which formed a quadruplex at 70 mM Na⁺ (Balagurumoorthy et al., 1992). The duplexed domains of our oligomers are thus more stable than the corresponding quadruplexes at those electrophoretic conditions. Although gel electrophoresis indicates the homogeneity of the sample, it cannot indicate the nature of the species responsible for any particular band. However, the presence of a single species for each oligomer in the electrophoretic gels coupled to the previously described characteristics of the corresponding thermal denaturation profiles indicates that each oligomer under these conditions exists solely as an intramolecular hairpin.

For gels run in the presence of both Mg²⁺ and K⁺, all oligomers show the formation of higher ordered structures.

The highest mobility band of each lane is assigned as the unquadruplexed intramolecular hairpin since these bands have relative mobilities similar to those under conditions giving rise to only single bands. The low mobility bands of the higher order structures observed in the presence of Mg²⁺ and K⁺ must therefore arise from some form of intermolecular multistrand formation, as verified by the UV cross-linking experiments. The CD data indicate a parallel strand orientation for these higher order structures.

The Ferguson data clearly indicate that each type of possible species has characteristic gel behavior. For example, the intramolecular hairpins have lower frictional coefficients and higher free mobilities, and thus higher charge densities, than the linear 12-mer. Consistent with multistrand formation, the higher order structures have higher frictional coefficients and higher charge densities than their respective parent hairpin. These data indicate that the higher ordered structures have shapes similar to each other but quite different from their parent structures.

Quadruplex formation for any member of the HP1-T series can only be obtained via unfolding of the hairpin, thereby exposing the internal G₄ segment, followed by quadruplexation via association of four of these unfolded linear oligomers. We, thus, propose a three step mechanism:



The minimum free energy required for the first step can be estimated from the data in Table 1 (3–4 kcal/mol of duplex at 25 °C times 4 duplexes = 12–16 kcal/mol). The free energy released upon quadruplex formation must be sufficient to overcome that required for unfolding of the hairpins. A recent thermodynamic study by Jin et al. (1992) determined ΔG° for *tetrad* formation to be –2.4 kcal/mol for TG₃T and –3.4 kcal/mol for TG₃TG₃T at 25 °C. Thus, for the oligomers of the HP1-T series, ΔG° for quadruplex formation (i.e., creation of four tetrads) would be on the order of –9.6 to –13.6 kcal/mol. The free energy released upon quadruplex formation may thus be insufficient to drive reaction 5 to completion. The observation of bands arising from both the parent and higher order structures is consistent with this conclusion. Examination of the gel pattern for the HP1-T series oligomers (Figure 3, lanes 1–4) reveals nearly equal band intensities for hairpin and quadruplex forms of HP1-T5, HP1-T6, and HP1-T8 while the quadruplex form predominates for HP1-T7. All initial strand concentrations were identical for this experiment. Since the rate of quadruplex formation may be slow (Sen & Gilbert, 1990), the variations in band intensities for our oligomers may be due to insufficient incubation resulting in a nonequilibrium distribution of species.

The thermodynamics and architectures of the higher ordered structures formed from the HP1-TG series are more complex. The most thermodynamically favorable process would be the formation of a four-stranded structure with a single quadruplex formed at the free 3'-terminal G₄ segment in the TG tail. If unfolding of the hairpin also occurs, allowing the participation of both G₄ segments, two possible structures can be considered: a quadruplex formed from the association of two "fold-back" species; or a four-stranded structure possessing two contiguous quadruplexes separated by T linkers of 1–4 bases. Both of these scenarios require that the free energy of quadruplex formation is sufficient to overcome that required for hairpin unfolding.

The formation of a fold-back structure is ruled out for several reasons. First, more than two products are observed arising from the UV cross-linking of HP1-T2G4. Fold-back structures have been observed for sequences such as $G_4T_2G_4$ (Balagurumorthy, 1992; Smith & Feigon, 1992, 1993) but not for G_4TG_4 . Thus, HP1-T2G4, HP1-T3G4, and HP1-T4G4 could form fold-back structures, but such formation is unlikely for HP1-T1G4. A fold-back structure and the four-stranded structure should also display quite different gel behavior. The Ferguson data, however, indicate that all members of the HP1-TG series form similar higher order structures. Furthermore, K^+ favors the four-stranded structure of $d(T_3G_4)_4$ over the fold-back structure (Hardin et al., 1991). The CD spectrum of HP1-T2G4 is also consistent with a parallel strand orientation. Fold-back structures are predetermined to have antiparallel strand orientation. Finally, the DNA concentrations used for these studies should also favor a tetramolecular association (Williamson, 1994).

The indication of very high molecular weight DNA in those lanes containing members of the HP1-TG series suggests higher order structures beyond simple quadraplexes. The gel resolution of 10–12 Mg^{2+}/K^+ -induced bands corresponds to discreet species possessing 14–16 oligomers. Hence, for HP1-T2G4, the lowest mobility bands that are still visually resolvable arise from species with molecular weights in the range of ca. 78 000–89 000. Previously, Sen and Gilbert (1990) observed some higher order structures with DNA oligomers possessing two G_4 segments separated by T linkers, but not as many as noted here nor of such high molecular weight. What differentiates our studies from previous studies is the presence of the C_4 run in the oligomers and the presence of both Mg^{2+} and K^+ in the gel electrophoresis.

The data plotted in Figure 4 suggest that the slower migrating bands of the HP1-TG series arise from successive additions of parent oligomer to an initially formed quadraplex. Successive additions could arise through two possible mechanisms. Watson–Crick associations of the free 5'- C_4 segments of the four-stranded quadraplex with the free terminal 3'- G_4 segments of parent hairpins would give rise to a concatomerization-like process. When HP1-T2G4 is run on polyacrylamide gels under the Mg^{2+}/K^+ conditions at temperatures up to 75 °C, the ladder of bands is still observed. Furthermore, a ladder of bands quite similar to that of HP1-T2G4 is also observed for $T_8G_4T_2G_4$ at 4 °C in the presence of both Mg^{2+} and K^+ (Sheardy and Marotta, unpublished results). These two observations indicate that Watson–Crick associations cannot be responsible for the self-assembly.

The most likely scenario is the linking of successive quadraplexes via one or two strands such as depicted in Figure 7. This conclusion is completely consistent with that of Marsh and Henderson (1994) in their assessment of electrophoretic data observed for $G_4T_2G_4$. Although there are fundamental differences between the sequences reported here and $G_4T_2G_4$, as well as experimental protocols, the data strongly suggest that multistranded species are responsible for the bands of high molecular weight observed in the electrophoretic gels. Marsh and Henderson observed increasing self-assembly as the temperature was increased from 4 to 37 °C for $G_4T_2G_4$ in either 50 mM Na^+ or 50 mM K^+ . We did not observe any self-assembly in either 100 mM Na^+ or 100 mM K^+ at 4 °C. This is most likely due to the free

A.

```

RGGGGTT-GGGG  R-GGGG--TT-GGGG  R-GGGG--TT-GGGG
RGGGGTT-GGGG  R-GGGG--TT-GGGG  R-GGGG--TT-GGGG
RGGGGTT-GGGG  R-GGGG--TT-GGGG  R-GGGG--TT-GGGG
R-GGGG--TT-GGGG  R-GGGG--TT-GGGG

```

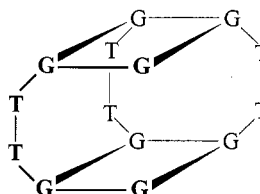
B.

```

RGGGGTT-GGGG  R-GGGG--TT-GGGG  R-GGGG--TT-GGGG
RGGGGTT-GGGG  R-GGGG--TT-GGGG  R-GGGG--TT-GGGG
R-GGGG--TT-GGGG  R-GGGG--TT-GGGG
R-GGGG--TT-GGGG  R-GGGG--TT-GGGG

```

C.



D.

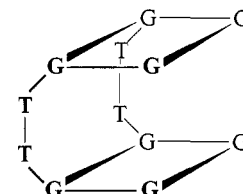


FIGURE 7: Schematic representation of the bridging of successive quadraplexes by one (A) or two (B) strands. Such self-assembly would result in a linear superstructure with periodic "pockets" such as shown in (C) for a one-strand bridge or (D) for a two-strand bridge.

energy of quadraplex formation being insufficient to drive the denaturation of the hairpin form of our oligomers under these conditions. We have also reported that the gel patterns observed for the HP1-TG series oligomers, in terms of band mobilities and intensities, are strongly dependent on sample preparation (Dai et al., 1994a). Thus, how the sample is prepared plays a role in the types of structures generated and their relative abundances. What is particularly noteworthy is the observation of self-assembled superstructures in the presence of both 100 mM K^+ and 20 mM Mg^{2+} at 4 °C. Apparently, Mg^{2+} plays a role in catalyzing the self-assembly and/or stabilizing the supermolecular structures.

The structure of the "G-wire" proposed by Marsh and Henderson (1994) has two strands bridging successive quadraplexes. This would give rise to supermolecular assemblies possessing 4, 6, 8, ..., $2N$ (where N is any integer) oligomers. However, self-assembly involving only one bridging strand (such as depicted in Figure 7A) would give rise to supermolecular structures possessing 4, 5, 7, 8, 9, 11, 12, 13, 15, ..., etc. oligomers. That is, assemblies possessing 6, 10, 14, ..., $4N+2$ ($N > 1$) oligomers would not be generated. The data presented in Figures 4 and 6B indicate supermolecular structures possessing $4 + N$ (N is any integer) oligomers. It is therefore likely that one- or two-strand bridges are possible within a single assembly. It should be noted that a single assembly could not, however, possess both one- and two-strand bridges.

The self-assembly of oligomers possessing $G_4T_xG_4$ segments (where $x = 1 - 4$ or higher) thus generates a linear scaffolding with periodically spaced "pockets" such as depicted in Figure 7C (for the one-strand bridge) and Figure 7D (for the two-strand bridge). In addition to electrostatic binding sites on the DNA backbones, these pockets possess hydrophobic and hydrogen bonding sites. Thus, these self-assemblies should be ideal candidates for binding small molecules within those pockets.

The observation of the higher order structures with the HP1-TG series oligomers, with sequences analogous to those found in nature, suggests that the structures of the ends of

chromosomes may be more complicated than originally thought. It is obvious that the exact nature of these particular higher order structures, the kinetics, thermodynamics, and mechanisms of their formation, and the roles of Mg^{2+} and K^+ in their stabilizations are of interest. We are currently undertaking further investigations with respect to these questions.

ACKNOWLEDGMENT

We thank Neville Kallenbach and Luis Marky for the use of their CD instrument and UV lamp and for their helpful discussions concerning this work.

REFERENCES

- Balagurumoorthy, P., Brahmachari, S. K., Mohanty, D., Bansal, M., & Sassisekharan, V. (1992) *Nucleic Acids Res.* 20, 4061–4067.
- Caruthurs, M. H. (1982) in *Chemical and Enzymatic Synthesis of Gene Fragments* (Gassen, H. G., & Lang, A., Eds.) Verlag Chemie, Weinheim.
- Dai, T.-Y., Marotta, S., & Sheardy, R. D. (1994a) *Biophys. J.* 66, A158.
- Dai, T.-Y., et al. (1994b) *Abstracts of Papers*, 207th National Meeting of the American Chemical Society, San Diego, CA, March 13–17, 1994; American Chemical Society, Washington, DC: ORG 66.
- Fasman, G., Ed. (1975) in *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed., p 589, CRC Press, Cleveland, OH.
- Giraldo, R., Suzuki, M., Chapman, L., & Rhodes, D. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 7658–7662.
- Guo, Q., Lu, M., Marky, L. A., & Kallenbach, N. R. (1992) *Biochemistry* 31, 2451–2455.
- Guo, Q., Lu, M., & Kallenbach, N. R. (1993) *Biochemistry* 32, 3596–3603.
- 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.
- Hardin, C. C., Watson, T., Corregan, M., & Bailey, C. (1992) *Biochemistry* 31, 833–841.
- Henderson, E. R., Moore, M., & Malcolm, B. A. (1990) *Biochemistry* 29, 732–737.
- Jin, R., Breslauer, K. J., Jones, R. A., & Gaffney, B. L. (1990) *Science* 250, 543–546.
- Jin, R., Gaffney, B. L., Wang, C., Jones, R. A., & Breslauer, K. J. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 8832–8836.
- Kanng, C., Zhang, X., Ratliff, R., Moyzis, R., & Rich, A. (1992) *Nature* 356, 126–131.
- Lu, M., Guo, Q., & Kallenbach, N. R. (1992) *Biochemistry* 31, 2455–2459.
- Lu, M., Guo, Q., & Kallenbach, N. R. (1993) *Biochemistry* 32, 598–601.
- Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers* 26, 1601–1620.
- Marsh, T. C., & Henderson, E. (1994) *Biochemistry* 33, 10718–10724.
- Paner, T. M., Amaratunga, M., & Benight, A. S. (1992) *Biopolymers* 32, 881–892.
- Panyutin, I. G., Kovalsky, O. I., Budowsky, E. I., Dickerson, R. E., Rokhrev, M. E., & Lipanov, A. A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 867–870.
- Rentzeperis, D., Alessi, K., & Marky, L. A. (1993) *Nucleic Acids Res.* 21, 2683–2689.
- Rodbard, D., & Chrambach, A. (1971) *Anal. Biochem.* 40, 95–134.
- Ross, W. S., & Hardin, C. C. (1994) *J. Am. Chem. Soc.* 116, 6070–6080.
- 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.
- Sheardy, R. D. (1988) *Nucleic Acids Res.* 16, 1153–1167.
- Smith, F. W., & Feigon, J. (1992) *Nature* 356, 164–168.
- Smith, F. W., & Feigon, J. (1993) *Biochemistry* 32, 8682–8692.
- Sundquist, W. I., & Klug, A. (1989) *Nature* 342, 825–829.
- Venczel, E. A., & Sen, D. (1993) *Biochemistry* 32, 6220–6228.
- Wang, Y., & Patel, D. J. (1992) *Biochemistry* 31, 8112–8119.
- Wang, Y., Jin, R., Gaffney, B., Jones, R. A., & Breslauer, K. J. (1991) *Nucleic Acids Res.* 19, 4619–4622.
- Williamson, J. R. (1994) in *Annual Reviews of Biophysics and Biomolecular Structure* 23. (Stroud, R. M., Cantor, C. R., & Pollard, T. D., Eds.) pp 703–730, Annual Reviews, Inc., Palo Alto, CA.
- Williamson, J. R., Raghuraman, M. K., & Cech, T. R. (1989) *Cell* 59, 871–880.

BI942584L