5004–5012 *Nucleic Acids Research, 1996, Vol. 24, No. 24*

Duplex–tetraplex equilibrium between a hairpin and two interacting hairpins of d(A-G)$_{10}$ at neutral pH

Mary Claire Shiber, Emory H. Braswell[1], Horst Klump[+] and Jacques R. Fresco*

Department of Molecular Biology, Princeton University, Princeton, NJ 08544-1014, USA and [1]The National Analytical Ultracentrifugation Facility, University of Connecticut, Storrs, CT 06269-3140, USA

Received July 16, 1996; Revised and Accepted November 6, 1996

ABSTRACT

d(A-G)$_{10}$ forms two helical structures at neutrality, at low ionic strength a single-hairpin duplex, and at higher ionic strength a double-hairpin tetraplex. An ionic strength-dependent equilibrium between these forms is indicated by native PAGE, which also reveals additional single-stranded species below 0.3 M Na$^+$, probably corresponding to partially denatured states. The equilibrium also depends upon oligomer concentration: at very low concentrations, d(A-G)$_{10}$ migrates faster than the random coil d(C-T)$_{10}$, probably because it is a more compact single hairpin; at high concentrations, it co-migrates with the linear duplex d(A-G)$_{10}$·d(C-T)$_{10}$, probably because it is a two-hairpin tetraplex. Molecular weights measured by equilibrium sedimentation in 0.1 M Na$^+$, pH 7, reveal a mixture of monomer and dimer species at 1°C, but only a monomer at 40°C; in 0.6 M Na$^+$, pH 7, only a dimer species is observed at 4°C. That the single- and double-stranded species are hairpin helices, is indicated by preferential S1 nuclease cleavage at the center of the oligomer(s), i.e., the loop of the hairpin(s). The UV melting transition below 0.3 M Na$^+$ or K$^+$, exhibits a $dT_m/d\log[Na^+/K^+]$ of 33 or 36°C, respectively, consistent with conversion of a two-hairpin tetraplex to a single-hairpin duplex with extrahelical residues. When [Na$^+$/K$^+$] ≥ 0.3 M, $dT_m/d\log[Na^+/K^+]$ is 19 or 17°C, respectively, consistent with conversion of a two-hairpin tetraplex directly to single strands. A two-hairpin structure stabilized by G-tetrads is indicated by differential scanning calorimetry in 0.15 M Na$^+$/5 mM Mg^{2+}, with ΔH of formation per mole of the two-hairpin tetraplex of –116.9 kcal or –29.2 kcal/mol of G-tetrad.

INTRODUCTION

The frequency of occurrence of alternating repetitive homopurine·homopyrimidine stretches in distinct functional regions, i.e., promoters, within eukaryotic genomes is much higher than predicted based upon genome complexity (1). Since these simple repetitive sequences have been shown to be polymorphic, it is intriguing to consider possible roles of alternative structures in gene regulation.

Several observations have shown that the repetitive sequence d(A-G)$_n$ has a potential to form self-structures. Based upon thermal melting and CD spectra, it was assumed that long strands of this sequence form a pH-independent four-stranded structure involving G- and A-tetrads (2–4). However, based upon CD spectral changes, it was suggested that the same d(A-G) self-structure is promoted by either acidic pH or high monovalent salt concentrations at neutrality (5). Complexes at pH 4–9 in 10 mM Mg^{2+} of the related sequences d(G-A)$_7$G and d(G-A)$_{12}$G were interpreted as linear parallel duplexes with G·G and A·A base pairs (6). A related homodimer structure was also reported (7). But, flanked by self-complementary sequences, d[-(A-G)$_{20}$-)] forms a hairpin stabilized by the terminal Watson–Crick base pairs (8). Some type of folded structure was also suggested for the repetitive sequences d(T$_4$GAGA)$_4$ and d(T$_4$AGAG)$_4$ at pH 8 (9); however, their electrophoretic behavior is different than that of the known fold-back tetraplex adopted by d(T$_4$G$_4$)$_4$ (9). In attempting to utilize d(A-G)$_{10}$ as a homopurine third strand directed to the target duplex d(A-G)$_{10}$·d(C-T)$_{10}$, we were led to ascribe uncommonly weak third strand binding to its involvement in competing equilibria. Subsequently, anomolous behavior for d(A-G)$_{10}$ was revealed, resulting in the identification of two pH-dependent classes of structures. Low pH and ionic strength stabilize a concentration-independent, single-stranded, α-helix-like structure for d(A-G)$_{n \leq 10}$ (10–12). Based upon CD spectroscopy, this helix was differentiated from the other that forms at higher ionic strength at neutrality. Under such neutral pH conditions, this sequence adopts alternate conformations in a concentration- and ionic strength-dependent manner. In this report, we characterize the single- and double-hairpin conformations of d(A-G)$_{10}$ at neutrality, and in the accompanying report (13), provide direct spectroscopic evidence for a G-tetrad H-bonding scheme in the double-hairpin conformation.

MATERIALS AND METHODS

Oligonucleotides

d(A-G)$_6$, d(A-G)$_{10}$ and d(C-T)$_{10}$ were synthesized and purified to homogeneity (10); excess salts were removed and the oligomers purified as the Na$^+$ salt using SepPak C18 (Waters) columns. Oligomer concentrations are reported on a residue basis.

*To whom correspondence should be addressed. Tel: +1 609 258 3927; Fax: +1 609 258 6730; Email: jrfresco@princeton.edu

[+]Present address: Department of Biochemistry, University of Capetown, South Africa

Native gel electrophoresis

The six sample conditions (all containing 0.01 M Tris–HCl, pH 7.4) included: (i) 0.01 M Na^+, (ii) 0.01 M Na^+/5 mM Mg^{2+}, (iii) 0.15 M Na^+, (iv) 0.15 M Na^+/5 mM Mg^{2+}, (v) 0.6 M Na^+ or 0.6 M Na^+/5 mM Mg^{2+}. Except where otherwise indicated, all cations were included as the chloride salt. Concentrations of $d(A\text{-}G)_{10}$ were varied as indicated; $d(C\text{-}T)_{10}$ and $d(A\text{-}G)_{10}\cdot d(C\text{-}T)_{10}$ were 6×10^{-1} mM. Each oligomer was labeled, purified (12) and an equal concentration of the appropriate 5′-^{32}P-end-labeled oligomer (6×10^{-4} mM) was included in all mixtures. Samples were heated to 75°C, cooled slowly to 25°C, and equilibrated at 4°C overnight. At 4°C, aliquots (4.5 µl) of each sample were added to 20% Ficoll-400 (1.5 µl), mixed briefly, and loaded onto a pre-electrophoresed 16% native polyacrylamide gel. Gels and electrophoresis buffers contained 0.089 M Tris–borate, pH 7.4 in addition to each of the aforementioned ionic conditions. Gels (0.4 mm × 31.0 cm × 38.5 cm) were run at 5–7 V/cm at 4°C with recirculation until bromophenol blue in a marker lane had migrated 16 cm; the samples contained no dye. Autoradiography of the wet gels was carried out for 18 h at 4°C.

Sedimentation equilibrium studies

Solutions of $d(A\text{-}G)_{10}$ in 0.01 M cacodylate, pH 7 with 0.6 M Na^+ or 0.15 M Na^+/5 mM Mg^{2+} at loading concentrations of 0.0042, 0.0083, 0.016, 0.05 and 0.13 mg/ml were centrifuged at 30 000 and 40 000 r.p.m. at 4°C. The ultracentrifuge and solution loading were as described (11). The quantities of solution and solvent used produced solution columns ~5 mm in height. Appropriate combinations of solution concentration, cell pathlengths of 4 and 12 mm, and wavelengths of light of 290, 280 and 260 nm were used to measure the concentration gradients by absorption. The oligomer in 0.01 M cacodylate, pH 7, with 0.1 M Na^+ was studied similarly at both 1 (30 000 and 40 000 r.p.m.) and 40°C (40 000 r.p.m.), with loading concentrations of 0.005, 0.015, 0.05 and 0.13 mg/ml.

Data analysis was performed (11) using a non-linear least squares program (NONLIN, 14) which fits all the data (speeds, wavelengths and optical paths) globally to various assumed association models. A good fit produces the reduced molecular weight $M' = [M(1 - \phi'\rho)]$, where ϕ' is the specific volume at constant free energy (15) and ρ is the solvent density, the values of the other constants related to the model (i.e., second virial coefficient, B; equilibrium constants for association, K; etc.), and the value of the r.m.s. error.

Solvent densities were estimated from density tables to be 1.004 and 0.999 g/ml at 1 and 40°C respectively for the 0.1 M Na^+ solvent, and 1.008 and 1.026 g/ml respectively for the 0.15 M Na^+/5 mM Mg^{2+} and 0.6 M Na^+ solvents at 4°C.

S1 nuclease digestion

Reaction mixtures contained 0.01 M Tris–HCl, pH 7.4, 0.15 M Na^+, 1 mM Zn^{2+} and 5 U S1 nuclease (US Biochemicals). Those with 0.64 mM $d(A\text{-}G)_{10}$ included $7–8 \times 10^{-3}$ mM 5′-^{32}P-end-labeled oligomer, with and without 5 mM Mg^{2+}; 6×10^{-4} mM samples contained only labeled oligomer. Mixtures were heated to 90°C and slowly cooled to 25°C; after 5 min at 5, 15 or 40°C (as indicated), digestion was initiated by adding S1 nuclease and Zn^{2+}. While pH 4.0–4.3 is optimal for enzyme activity, digestions were performed at pH 7.4, where the activity is dramatically diminished (16), allowing detection of the earliest and most sensitive cleavage sites. Digestions were halted by removing 5µl aliquots at 0.5, 1, 2, 5, 10, 15, 30 and 60 min, adding them to 4 µl loading dye (70% formamide, 0.1% xylene cyanole, 0.1% bromophenol blue, 57 mM EDTA, pH 7.5), and freezing. Frozen samples were heated at 90°C for 5 min prior to loading on a 22% denaturing gel, and electrophoresed at 50 V/cm until the bromophenol blue migrated 20 cm. Products were visualized by autoradiography.

UV thermal melting

Samples contained 0.16 or 0.64 mM $d(A\text{-}G)_{10}$ in 0.01 M cacodylate, pH 7.0, a total Na^+ or K^+ concentration of 0.05, 0.10, 0.15, 0.30, 0.60 or 1.2 M, and where indicated 5 mM Mg^{2+}. Solutions were equilibrated at 25°C for 1 h followed by 10 min at 3°C prior to thermal melting in 0.1 or 1.0 cm pathlength cuvettes. After melting, samples in the Na^+ series were brought to 25°C and equilibrated for 4 weeks, followed by 10 min at 3°C and remelted. Profile reproducibility was excellent.

Absorbances were recorded using a computer-driven AVIV 14DS spectrophotometer (AVIV Associates, Lakewood, NJ) equipped with a thermoelectrically controlled cuvette holder. Spectra recorded between 320 and 230 nm at 1 nm intervals every 2°C from 3 to 89°C were used to generate melting profiles and their derivatives at desired wavelengths. Oligonucleotide spectra were corrected for solvent and cuvette contributions at each temperature. Melting profiles were corrected for thermal expansion of water, smoothed by a least-squares polynomial fit (third order) and normalized to an absorbance of 1.0 at 3°C.

Differential scanning calorimetry (DSC)

Measurements were made as described (11). A baseline was obtained between 10 and 60°C with the solvent (0.01 M cacodylate, pH 7, 0.15 M Na^+, 5 mM Mg^{2+}) in both reference and sample cells, and used to correct the DSC plot for the sample. The sample cell with 1.59 mM $d(A\text{-}G)_{10}$ in solvent and the reference cell with solvent alone were scanned from 10 to 60°C. After slow cooling to 10°C, and equilibration overnight, they were re-scanned.

RESULTS

Concentration dependence of $d(A\text{-}G)_{10}$ structure

In contrast to the single-stranded helix without base stacking formed by this sequence in 0.01 M Na^+ at acidic pH (10), preliminary experiments on this oligomer in 0.15 M Na^+, 5 mM Mg^{2+} at pH 7 revealed cooperative UV melting attended by a large hyperchromic change and oligomer concentration-dependence of T_m. Also, at neutrality a unique intense CD spectrum characteristic of a helical structure was observed (Figure 10 in ref. 14).

Native PAGE was utilized to determine how oligomer concentration influences formation of the $d(A\text{-}G)_{10}$ helical structure(s) at pH 7, using the duplex $d(A\text{-}G)_{10}\cdot d(C\text{-}T)_{10}$ and single-stranded random coil $d(C\text{-}T)_{10}$ as mobility standards of known size and conformation. [While $d(C\text{-}T)_n$ has been reported to exist as a duplex at very high concentrations (17), we have found that in the concentration and ionic strength range studied here, it does not form a multi-stranded stacked structure. Thus, at pH 7.4, $d(C\text{-}T)_{10}$ migrates equally in 0.15 M Na^+ with and without Mg^{2+} at both 6×10^{-3} and 1 mM oligomer on native PAGE, i.e., in a concentration-independent manner. Also, $d(C\text{-}T)_{10}$ exhibits no

UV hyperchromism with increasing temperature.] Figure 1 shows the electrophoretic mobility of $d(A\text{-}G)_{10}$ over a 1000-fold oligomer concentration range, from 6×10^{-4} to 6×10^{-1} mM (residues) and under a range of ionic conditions, in comparison with the two standards. The relative mobilities reveal two salient features: (i) there is oligomer concentration and ionic strength dependence of structure formation, and (ii) there is an equilibrium between multiple monomer (i.e., single-stranded) species and one dimer species of $d(A\text{-}G)_{10}$. These features are manifest in several ways. In general, $d(A\text{-}G)_{10}$ migrates as very broad band(s), indicative of continual re-equilibration among structures as oligomer is diluted during the course of electrophoresis. There are a minimum of three types of species with different mobilities, the equilibrium between them depending on oligomer concentration and ionic conditions: a duplex species migrating like $d(A\text{-}G)_{10}\cdot d(C\text{-}T)_{10}$, a random coil single-stranded species migrating like $d(C\text{-}T)_{10}$, and still faster migrating species.

The species co-migrating with the duplex standard is favored by high oligomer concentration and ionic strength, being most clearly observed in 0.6 M Na^+ with and without Mg^{2+} (Fig. 1A and B, lanes 10–16). As addition of Mg^{2+} to 0.6 M Na^+ does not appear to enhance the stability of this species, the equilibrium must already be maximally shifted by 0.6 M Na^+ towards the duplex when oligomer concentration is $\geq 4 \times 10^{-2}$ mM. The duplex is also evident in 0.15 M Na^+/5 mM Mg^{2+} or 0.01 M Na^+/5 mM Mg^{2+} at oligomer concentrations $\geq 4 \times 10^{-2}$ mM (Fig. 1C and D, respectively). While some duplex is formed at $\geq 4 \times 10^{-2}$ mM oligomer in 0.15 M Na^+ without Mg^{2+}, the resulting bands are more diffuse, indicating a decrease in the stability of the structure at lower ionic strength (Fig. 1E).

Bands with mobilities like those for the single-stranded random coil standard are observed at lower oligomer concentrations at both low and high ionic strength (Fig. 1). However, diffuse banding patterns with intermediate mobilities, i.e., between single strand and duplex are also apparent. A broad banding pattern probably results from duplex dissociation due to oligomer dilution as electrophoresis proceeds (Fig. 1A–F).

A third species, which migrates fastest, is most favored and clearly observed at lower oligomer concentration in 0.01 M Na^+ (Fig. 1F, lanes 1–8). Diffuse mobilities between those of the single-strand standard and the fastest species probably result from an equilibrium distribution of monomer species intermediate between the true random coil [like $d(C\text{-}T)_{10}$] and a more compact fully formed hairpin helix (see below). Intramolecular hairpin helices have been observed to migrate faster than random coil single strands (18), while the duplex formed by $d(A\text{-}G)_{10}$ corresponds not to a linear duplex, but to an unusual two-hairpin duplex (see below).

In sum, native PAGE has revealed a monomer–dimer oligomer equilibrium that depends both on oligomer and ion concentration.

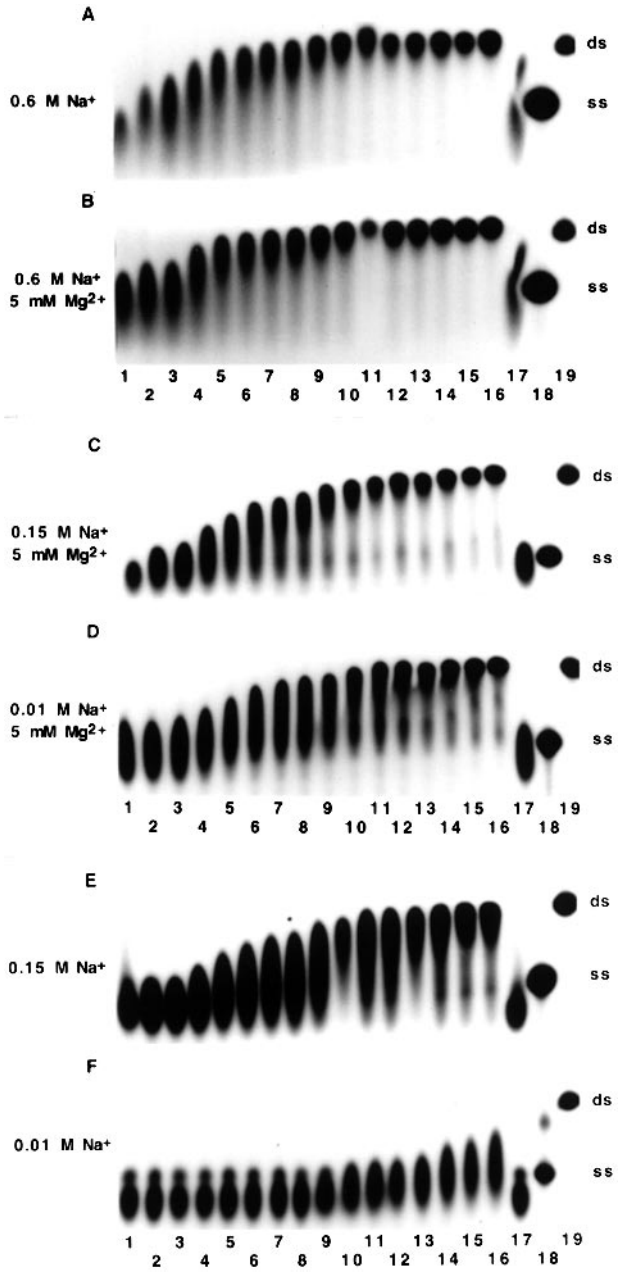

Figure 1. Concentration dependence of $d(A\text{-}G)_{10}$ structure at pH 7.4 under various ionic conditions: (**A**) 0.6 M Na^+, (**B**) 0.6 M Na^+/5 mM Mg^{2+}, (**C**) 0.15 M Na^+/5 mM Mg^{2+}, (**D**) 0.01 M Na^+/5 mM Mg^{2+}, (**E**) 0.15 M Na^+, (**F**) 0.01 M Na^+. Lanes 1–17 in (A)–(F) contain $d(A\text{-}G)_{10}$ at the following concentrations: lanes 1 and 17, 6×10^{-4} mM; lane 2, 8×10^{-4} mM; lanes 3–7, 1, 2, 4, 6 and 8×10^{-3} mM; lanes 8–12, 1, 2, 4, 6 and 8×10^{-2} mM; lanes 13–16, 0.1, 0.2, 0.4 and 0.6 mM. In (A)–(F), lane 18 contains single-stranded (ss) $d(C\text{-}T)_{10}$ at 0.6 mM and lane 19 contains the (1:1) duplex (ds) $d(A\text{-}G)_{10}\cdot d(C\text{-}T)_{10}$ at 0.6 mM. All gels are 16% native polyacrylamide and electrophoresis was carried out as described in the Materials and Methods section.

Molecularity of $d(A\text{-}G)_{10}$ structures

Equilibrium sedimentation analysis was used to obtain molecular weights directly and thereby the number of oligomer units associated with the three types of structures revealed by native PAGE. For this purpose, sedimentation experiments were performed under three ionic conditions, in one case at two temperatures.

(i) In 0.6 M Na^+ at 4°C, where the gels reveal a high propensity for $d(A\text{-}G)_{10}$ to form a stable duplex structure (Fig. 1A), 10 sets of sedimentation data covering a range from ~10 to ~0.6 mg/ml oligomer were fitted. An ideal single species (ISS) model fit the data, producing a molecular weight of 6.583×10^3 g/mol, corresponding to a dimer, assuming ϕ' is ~0.56 ml/g.

(ii**)** In 0.15 M Na^+/5 mM Mg^{2+} at 4°C, where the gels reveal the same predominant molecular species (Fig. 1C), 10 sets of data were fitted covering a range from <10 μg/ml to >0.6 mg/ml. The

best fit was obtained by the ISS model, producing a molecular weight corresponding to a dimer, assuming a ϕ' of ~0.59 ml/g.

(iii) At 1°C in 0.1 M Na^+, 10 sets of data were fitted spanning a range of a few μg/ml to ~0.18 mg/ml, close to that in the gel studies. The fit of the data to the ISS model was poor, giving a Z average molecular weight between that of the monomer and dimer. The best fit was to a monomer–dimer model wherein different equilibrium constants are allowed for each loading concentration and speed. This fit produced a molecular weight corresponding to the actual monomer, assuming a ϕ' of 0.59 ml/g. The range of values for $\ln K_2$ (l/g) was from 3.4 to 12.9 (avg. ~6), indicating heterogeneity in association tendencies within the predominant monomer–dimer equilibrium. The range of values for the equilibrium constant indicates that ≥80% of the strands are present as dimers at 1 mg/ml (3.15 mM) at 1°C.

(iv) Seven data sets spanning a range up to ~0.8 mg/ml in 0.1 M Na^+ at 40°C, best fitted a non-ideal single species model, with a molecular weight corresponding to the single chain oligomer if ϕ' is 0.55 ml/g, and a small positive virial coefficient, probably resulting from a small effective charge on the molecule under these conditions. The ISS model fit the data almost as well, however, resulting in a molecular weight corresponding to the monomer if $\phi' = 0.56$ ml/g. These results indicate that under these conditions the molecule exists as a single chain with a small effective charge.

The values of ϕ' reported for a DNA double helix (19) are ~0.53 and 0.55 ml/g respectively in 0.15 and 0.6 M Na^+. Those for $d(A\text{-}G)_{10}$ at neutrality are somewhat different; at low temperatures and higher salt concentrations, where it is present as a hairpin dimer in a tetraplex structure (see below), $\phi' = 0.57$–0.59 ml/g, and for the monomer strand at 40°C, the value is 0.55–0.56 ml/g. Moreover, the single-stranded $d(A^+\text{-}G)_{10}$ helix at low temperature and its random coil form at elevated temperature exhibit a ϕ' of 0.54 ml/g (11). Apparently, structure and base composition affect ϕ', though exactly how is not obvious.

These direct observations of the molecular weights of the $d(A\text{-}G)_{10}$ structures at neutrality completely corroborate the interpretations made on the basis of native PAGE.

S1 nuclease-sensitive sites

To distinguish between hairpin and linear structures for the putative helical conformations indicated by the native PAGE and sedimentation equilibrium studies, the earliest sites of S1 nuclease cleavage were determined under various conditions of ionic strength, oligomer concentration and temperature under which those conformations had been identified. Because of the dynamic equilibrium between the different structures, it was necessary to find digestion conditions that would produce cleavage patterns characteristic of the predominant species. Realizing this goal was complicated because conditions for optimal structure stability do not correspond to those for optimal activity of the nuclease (see Materials and Methods). Moreover, near 0°C, where the helical conformations are most stable, the longer digestion time required to generate specific cleavage products also contributes to the background of observable products.

Figure 2A shows the results in 0.64 mM $d(A\text{-}G)_{10}$ in 0.15 M Na^+, where the dimer helix is the predominant species; and Figure 2B shows the results in 6×10^{-4} mM $d(A\text{-}G)_{10}$ in 0.15 M Na^+, which favors the monomer helix. A cleavage pattern essentially the same as for Figure 2A was obtained when 5 mM Mg^{2+} was added to 0.64 mM $d(A\text{-}G)_{10}$ in 0.15 M Na^+ (data not shown). For both of these structures, the major nuclease sensitive sites are at the center of the molecule. S1 nuclease appears to cleave more readily after dG than after dA residues, with clearly preferred cleavage sites at G10 and G12, as established by comparison with the undigested $d(A\text{-}G)_6$ and $d(A\text{-}G)_{10}$ chainlength markers. Though the kinetics of digestion are obviously faster at higher temperatures, parallel experiments at 15 and 40°C or 5 and 15°C for each oligomer concentration reveal the same cleavage sites. Cleavage at the 5′ and 3′ ends of the strand(s) is also observed, in keeping with the propensity of helices to 'fray' at their termini. However, at lower temperatures, cleavage at the termini is preferentially diminished relative to cleavage at the center of the molecule.

These results are only consistent with hairpin structures for $d(A\text{-}G)_{10}$ at both low and high oligomer concentration, in 0.15 M Na^+ without or with Mg^{2+}. Together with native PAGE analysis and equilibrium molecular weight observations, these findings lead to their characterization as single- and double-hairpin helices rather than linear ones.

UV thermal melting

Analysis of thermal melting profiles for $d(A\text{-}G)_{10}$ over a range of ionic conditions is not straightforward. Although determined over the entire near-UV range, profiles are shown only at 260 and 280 nm. Since 280 nm is isosbestic for the unstacking of dA residues (19), changes in the stacking of only the dG residues are observable there; and at 260 nm, the hyperchromic contribution from unstacking of dA residues is approximately twice that from dG residues (19). These hyperchromic changes can be due to unstacking of bases that may or may not be H-bonded. Only if the bases are in helices involving two or more strands, will the changes be responsive to ionic strength. Against this background, and with the knowledge of the single- and double-hairpin helices identified above at low temperatures, the melting profiles are qualitatively interpretable.

In 0.6 M Na^+ or K^+, or 0.15 M Na^+/5 mM Mg^{2+}, where the prevailing species at low temperature is the two-hairpin tetraplex, the melting profiles are quite similar (Fig. 3B, D and E). The 260 and 280 nm profiles essentially coincide up to 25°C, suggesting that the small absorbance increase constitutes some 'pre-melting' adjustment of tetraplex conformation. Between ~25 and 55°C, the dA residues apparently become extrahelical and the dG-tetrads begin to melt. These processes are not coincidental, based upon the profiles in Figure 3B, D and E; and there is no reason why they should be. For, with tetrad melting, the liberated dG residues can stack with their dA nearest neighbors, and these dA–dG stacks will then melt with increasing temperature. Above ~60°C, the further melting of dA and dG residues must overlap sufficiently, since all that is observed is a gradual hyperchromic change at 280 nm.

In 0.15 M Na^+ or K^+, where there is an equilibrium mixture at low temperature favoring the two-hairpin tetraplex over the single-hairpin duplex, the profiles in the two cations (Fig. 3A and C) are again similar, but the nature of the melting process is quite different from that at high ionic strength. Thus, the profiles at 280 nm differ qualitatively from those at 260 nm, and their T_m values are substantially higher, e.g., ΔT_m ($T_{m\ 280nm} - T_{m\ 260nm}$) = 28.4°C in 0.15 M Na^+. Moreover, melting of the dA residues to an extrahelical state is complete by ~40°C, largely preceding the melting of the dG residues present in a mix of tetraplex and

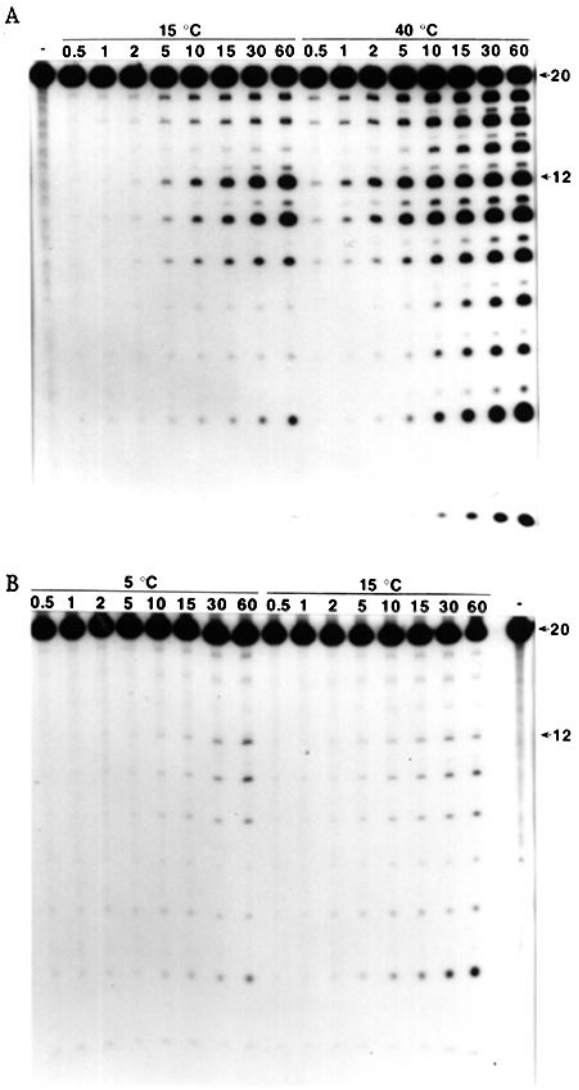

Figure 2. S1 cleavage products of d(A-G)$_{10}$ at pH 7.4. Patterns obtained for cleavage at 15 and 40°C at the indicated times (min) of 0.64 mM d(A-G)$_{10}$ in (**A**) 0.15 M Na^+. Patterns obtained for 5 and 15°C at the indicated times (min) of 6×10^{-4} mM d(A-G)$_{10}$ in (**B**) 0.15 M Na^+. Lanes labeled (–) contain uncleaved d(A-G)$_{10}$, i.e., $n = 20$. The oligomer band corresponding to cleavage at residue G12 of d(A-G)$_{10}$ was identified by co-migration with a band corresponding to uncleaved d(A-G)$_6$, i.e., $n = 12$ (not shown). All gels are 22% denaturing polyacrylamide.

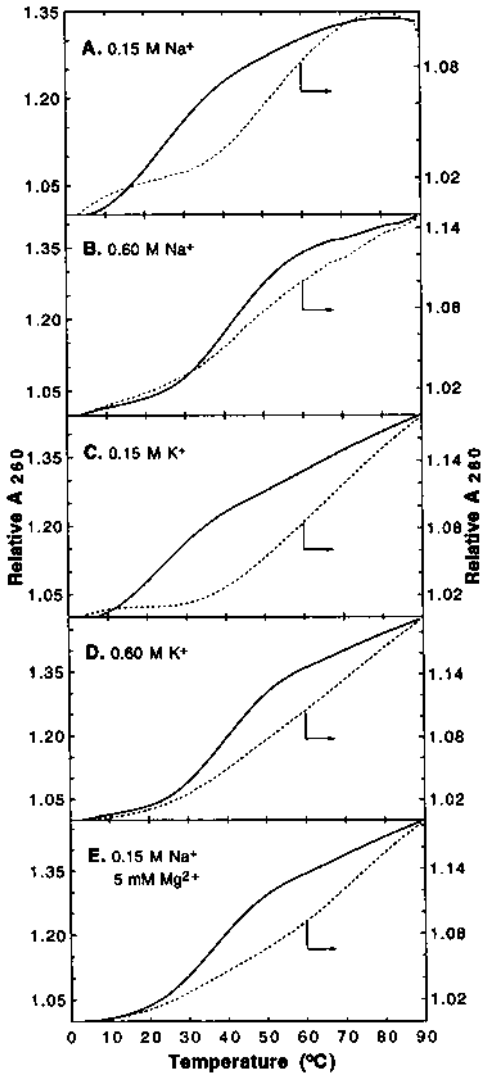

Figure 3. UV melting profiles at 260 (——) and 280 (- - -) nm for 0.16 mM d(A-G)$_{10}$ in (**A**) 0.15 M Na^+, (**B**) 0.6 M Na^+, (**C**) 0.15 M K^+, (**D**) 0.6 M K^+, (**E**) 0.15 M Na^+/5 mM Mg^{2+}, at pH 7.0. The arrows indicate the profiles related to the ordinate on the right.

duplex, that itself shifts with increasing temperature to single-hairpin duplex intermediates. Again, as the dG pairs or tetrads melt, these residues can stack with their nearest neighbor dA residues, and then the resultant dA–dG neighbors gradually unstack. Therefore, melting at >40°C must be a complex non-cooperative process for which the profiles at both 260 and 280 nm above 40°C reflect the cumulative absorbance changes. This differential melting at 260 and 280 nm argues against H-bonding between the dA and dG residues. This dissimilarity in the thermally induced unstacking of the dA and dG residues is not unlike that observed by UVRR in 0.6 M Na^+ (0.3 M Na_2SO_4) at 2 mM d(A-G)$_{10}$ using selective excitation of dA or dG residues (13).

It is interesting that the UVRR profile in 0.6 M Na^+ (13) is much more similar to the disparate melting of those residues in 0.15 M rather than 0.6 M monovalent cation. Conceivably, this relates to different ratios of cation to oligomer concentration in the UV melting [0.16 mM d(A-G)$_{10}$] and the UVRR melting [2 mM d(A-G)$_{10}$] experiments, which in turn affects the equilibrium between single-hairpin duplex and double-hairpin tetraplex. Additionally, since in the simplest model resonance Raman intensity is proportional to the square of the absorptivity (20,21), subtle changes in absorption can lead to more dramatic changes in the resonance Raman spectrum. At 0.6 M monovalent cation, it is therefore possible that the UVRR melting profiles are more sensitive to the differential melting behavior of the dA residues than the UV absorption melting profiles. In any case, both the UV and the UVRR melting observations indicate that at low temperature both the dA and the dG residues are stacked in helices that the UVRR data show to contain G-tetrads (13).

The changing nature of the melting process as a function of cation concentration is further indicated by an analysis of the UV

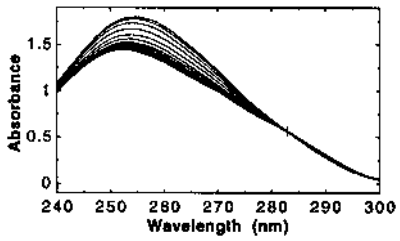

Figure 4. UV spectra of d(A-G)$_{10}$ in 0.60 M Na^+ as a function of temperature. Scans are shown every 4°C from 3 to 55°C. An isosbestic point (wavelength) is indicated by a vertical line which intersects the spectra.

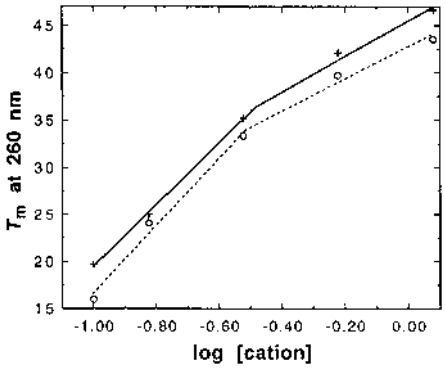

Figure 5. Plot of T_m versus log [Na^+] (+ — +) and log [K^+] (∘- - -∘) at 260 nm.

spectra at various temperatures. It was thereby possible to define temperature ranges over which the melting process is seemingly two-state, which means operationally that as a function of temperature the spectra pass through an isosbestic point (example in Fig. 4). The upper temperature limit of the range increases from ~35°C in 0.1 M Na^+ to ~60°C in 1.2 M Na^+ (data not shown), demonstrating a substantial increase in stability for the low temperature structure(s) with higher salt concentration. These upper limits coincide roughly with the temperatures where the seemingly two-state melting observed at 260 nm ends. Also, with increasing Na^+ concentration, the isosbestic point gradually red shifts. Similar behavior is seen when the cation is K^+ (data not shown).

The ionic strength-dependence of thermal melting provides additional insights. Plots of T_m versus log [Na^+ or K^+] for the putative d(A-G)$_{10}$ double hairpin present at low temperature reveal that the structure is extraordinarily sensitive to salt concentration. And in contrast to poly d(G-A), which displays monophasic semilog plots of comparable slope for Na^+ or K^+ (4), d(A-G)$_{10}$ shows two distinct linear dependencies. At ≤0.3 M, $dT_m/d\log [Na^+]$ is 33°C, while for ≥0.3 M it is 19°C (Fig. 5); for K^+ the comparable values are 36 and 17°C, respectively (Fig. 5). For natural DNA duplexes, the values are 18–22°C (22). Hence, the high sensitivity of T_m upon salt concentration at ≤0.3 M monovalent cation for d(A-G)$_{10}$ is reasonable for a helical structure with a negative charge density greater than that of a duplex, i.e., consistent with the presence of at least some tetraplex, in this case formed from two hairpins. However, it is puzzling that this sensitivity seems to decrease at concentrations >0.3 M.

This difference in $dT_m/d\log[Na^+$ or $K^+]$ below and above 0.3 M suggests that dissociation occurs by different pathways, at <0.3 M, double-hairpin (tetraplex)→single-hairpin (duplex)→single-strand, and at >0.3 M, double-hairpin (tetraplex)→single-strand. A similar difference in the pathway of dissociation has been observed for the triplex poly(U:A·U), whose melting occurs via a duplex intermediate to the single-strands at lower ionic strengths, but directly to single-strands at higher ionic strengths (23). Moreover, the linear dependence of T_m upon log ionic strength for the 3→1 and 3→2 transitions is different, the slope being lower for the former than for the latter, just as observed for the tetraplex→single-strand transition for d(A-G)$_{10}$ relative to that of the tetraplex→duplex transition. What these results particularly emphasize is that at higher ionic strengths, the two-hairpin tetraplex is more stable than the hairpin duplex.

Both values for $dT_m/d\log[K^+]$ for d(A-G)$_{10}$ contrast the extremely low values for $dT_m/d\log[K^+]$ reported for the G-tetraplexes formed by the consecutive dG residues of d(TG_3T) and by d(TG_3T)$_2$ (24). In that study, it was suggested that the presence of site bound K^+ might be the reason for the uncommonly low ionic strength dependence of the stability of the tetraplex. This is consistent with findings that in sequences containing runs of G residues, K^+ preferentially stabilizes G-tetrad formation (25,26). The K^+ is thought to coordinate the carbonyl groups and has been localized between the planes of G-tetrads (27). Clearly, such selective K^+ site-binding does not occur with the tetraplex formed by d(A-G)$_{10}$, as indicated by the marked similarity of the ionic strength dependence of T_m in Na^+ and K^+. This must be because the G-tetrads are not contiguous, in which case the alternating arrangement of dA amino and dG carbonyl moieties in the axial channel of the tetraplex probably precludes selective interplanar localization of site-bound K^+ in the core of this tetraplex.

Thus, d(A-G)$_{10}$ forms a tetraplex that melts similarly in the presence of either Na^+ or K^+. In addition to native PAGE, the overwhelming similarity between all the analyses of the thermal melting data in Na^+ and K^+ implies that both cations similarly stabilize helices that undergo similar thermal dissociation pathways (those pathways being different below and above 0.3 M Na^+/K^+).

Thermodynamic parameters of d(A-G)$_{10}$

The foregoing observations, together with those of Mukerji *et al.* (13), indicate that d(A-G)$_{10}$ forms a bimolecular hairpin at low temperature, stabilized by alternating G-tetrads and intercalated dA residues. Since the complex exhibits a reversible UV melting profile, values for the thermodynamic parameters of the complex in 0.15 M Na^+/5 mM Mg^{2+} (Fig. 3E) were first obtained by assuming a two-state cooperative transition (28,29). Shape analysis of this thermal melting profile gave an equilibrium constant for complex formation at T_m (36.7°C) of 1.25×10^5, and thermodynamic values for formation of the complex of ΔH_{vH}(shape) = –73.0 kcal/mol, ΔG_{298} = –3.4 kcal/mol and ΔS = –233.4 cal/mol·K. However, the apparent breadth of the melting profile and wavelength dependence of T_m (see below) suggest that complications are introduced in assuming a two-state transition, even at higher ionic strength, as presupposed by the van't Hoff analysis.

DSC was therefore used to directly obtain thermodynamic parameters. The calorimetric transition curve (Fig. 6), corresponding to excess heat capacity of the sample versus temperature, is almost symmetrical and centered about a T_m of 43.97°C (higher than the value from the UV profile due to higher oligomer concentration), with flat and reliable (data not shown) pre- and post-transitional baselines. The area under the curve indicates a favorable enthalpy of formation (ΔH_{cal}) of –116.9 kcal/mol

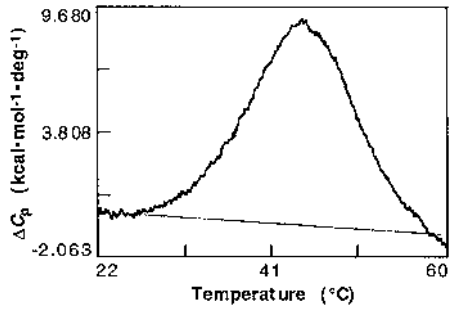

Figure 6. Molar excess heat capacity (Δ*C*p) versus temperature for d(A-G)$_{10}$ in 0.15 M Na^+/5 mM Mg^{2+}, pH 7 as obtained by differential scanning calorimetry.

complex. The stability of the complex at 25°C is reflected in the value of $\Delta G = -7.0$ kcal/mol, while the large entropy of formation, $\Delta S_{cal} = -368.8$ cal/mol·K, is in keeping with the requirement for the association of four dG residues to form tetrads in the complex.

A comparison of the transition enthalpies obtained by the model-dependent and independent methods shows that ΔH_{vH}(shape) < ΔH_{cal}, confirming that the transition involves intermediate populations and is not a two-state process. This is corroborated by a van't Hoff analysis of the calorimetric heat capacity curve itself, which yields a value of ΔH_{vH}(cal) = –62.3 kcal/mol, also much smaller than that of ΔH_{cal}. However, the ratios ΔH_{vH}(shape) / ΔH_{cal} and ΔH_{vH}(cal) / ΔH_{cal} suggest that the cooperative melting unit is two, supporting the proposed model. This suggests that two hairpin helices interact to form a tetraplex with the four G-tetrads in the structure remaining after the intercalated dA residues have become extrahelical. That unpaired residues can be accomodated extrahelically is well established (30–32). This interpretation is also supported by the dissimilarity in thermal melting of the dA and dG residues observed by both UV and UVRR melting experiments.

Schematic diagrams for possible structures are depicted in Figure 7A and B (see below); the maximal incorporation of four G-tetrads in a two-hairpin helix is in keeping with our observations and likely represents the most stable possibilities. In this connection, with a ΔH_{cal} for d(A-G)$_{10}$ of –116.9 kcal/mol and four such G-tetrads in a two-hairpin tetraplex, a corresponding value of –29.2 kcal/mol of tetrad is obtained. This value is in keeping with the reported values for enthalpy of formation of a G-tetrad, which range from –21 to –26 kcal/mol (24).

DISCUSSION

A variety of techniques have been used to provide evidence that at neutrality and low temperature d(A-G)$_{10}$ is involved in multiple *equilibria* between monomer and dimer *molecularities*, whose principal *conformations* are a single hairpin duplex and a two-hairpin tetraplex. The *equilibria* depend upon oligomer and salt concentrations in addition to temperature. Under a variety of conditions, native PAGE reveals fast-, slow- and co-migrating species relative to d(C-T)$_{10}$, a random coil of identical chainlength. The fast species is most likely accounted for by the compact nature of a fully formed single hairpin. The slow-migrating species, which approximately co-migrates with d(A-G)$_{10}$·d(C-T)$_{10}$, is best explained by some type of duplex. These *molecularities* are confirmed by sedimentation equilibrium measurements. In 0.1 M Na^+ at 1°C, the oligonucleotide monomer is in equilibrium with the dimer; at higher concentrations, larger aggregates may be present to a very small extent. However, at 40°C in this buffer, the oligonucleotide is present as monomer only. In 0.15 M Na^+/5 mM Mg^{2+} at 4°C, the oligonucleotide is present as a dimer. In 0.6 M Na^+ at 4°C, the oligonucleotide is present as a dimer, with a very small amount of higher aggregates (possibly linear tetraplex).

One explanation for the approximate co-migration with the Watson–Crick duplex is a linear d(A-G)$_{10}$·d(A-G)$_{10}$ model, as suggested by Rippe *et al.* (6); another possibility is a hairpin duplex in which the two strands interact to form a tetraplex. Although a hairpin duplex is more compact than a linear duplex, the homopurine nature and unusual stacking in this complex might make its mobility relative to the linear Watson–Crick duplex hard to predict. Hence, on the basis of gel migration patterns, neither the linear nor the hairpin models could be ruled out. Discrimination between these *conformations* was sought using dimethylsulfate (DMS) modification to determine the presence of any preferred cleavage site (data not shown). However, this approach only afforded inconclusive results, probably because methylation at the hairpin loops and termini causes a shift in the conformational equilibrium. A more selective conformational probe was found in S1 nuclease, which under conditions previously determined to stabilize the monomer or dimer complex, preferentially cleaves at positions G10 and G12 (the center of the molecule). This result establishes that the strand(s) of both the monomer and dimer species are hairpins, in one case forming a duplex, in the other a tetraplex.

A two-hairpin tetraplex could be interlocked, with the loops at opposite ends, as shown in Figure 7A. This type of complex has been observed using ^{1}H NMR spectroscopy in a structure containing G-tetrads and T residues at the turns (33). However, such an interlocked tetraplex should melt in an all-or-none fashion, which is not what is observed (Fig. 3). This absence of a two-state transition in the melting profiles at either low or even at higher ionic strength is most clearly demonstrated by the inequality of the transition enthalpies, ΔH_{vH}(shape) and ΔH_{cal}. Therefore, side-by-side arrangements appear most likely.

Dimeric side-by-side and monomeric hairpins are schematically depicted in Figure 7B. Homogeneous associations of single hairpins *a*, *f* and *k* are shown in horizontal rows *b–e*, *g–j* and *l–o*, respectively. Heterogenous associations of hairpins *f* and *k* are shown in horizontal rows *p–s*. Since dissimilar melting profiles at wavelengths more selective for A or G residues were indicated by both UV absorption and UVRR spectroscopies (13), the lack of evidence for A·G base pairs would seem to eliminate structures *a–e*.

The preferred cleavage of both monomers and dimers at G10 and G12 can most easily be explained by single hairpins *f* and *k* and their homogeneous associations in *g–j* and *l–o*, respectively; however, heterogeneous associations of hairpins *f* and *k*, as in *p–s*, might yield similar results. While the earliest timepoints at which cleavage could be observed indicate equal cleavage at G10 and G12, the resolution of the assay is not good enough to allow discrimination between these models.

Based upon fluorescence energy transfer studies of d(G-A)$_7$G and d(G-A)$_{12}$G under similar ionic conditions, Rippe *et al.* (6) concluded that these oligomers form linear parallel duplexes stabilized by A·A and G·G base pairs. However, their results are not incompatible with our hairpin model, so long as the two hairpins are parallel. Moreover, the possibility exists that the presence of fluorescent tags at the 5′ termini may shift the

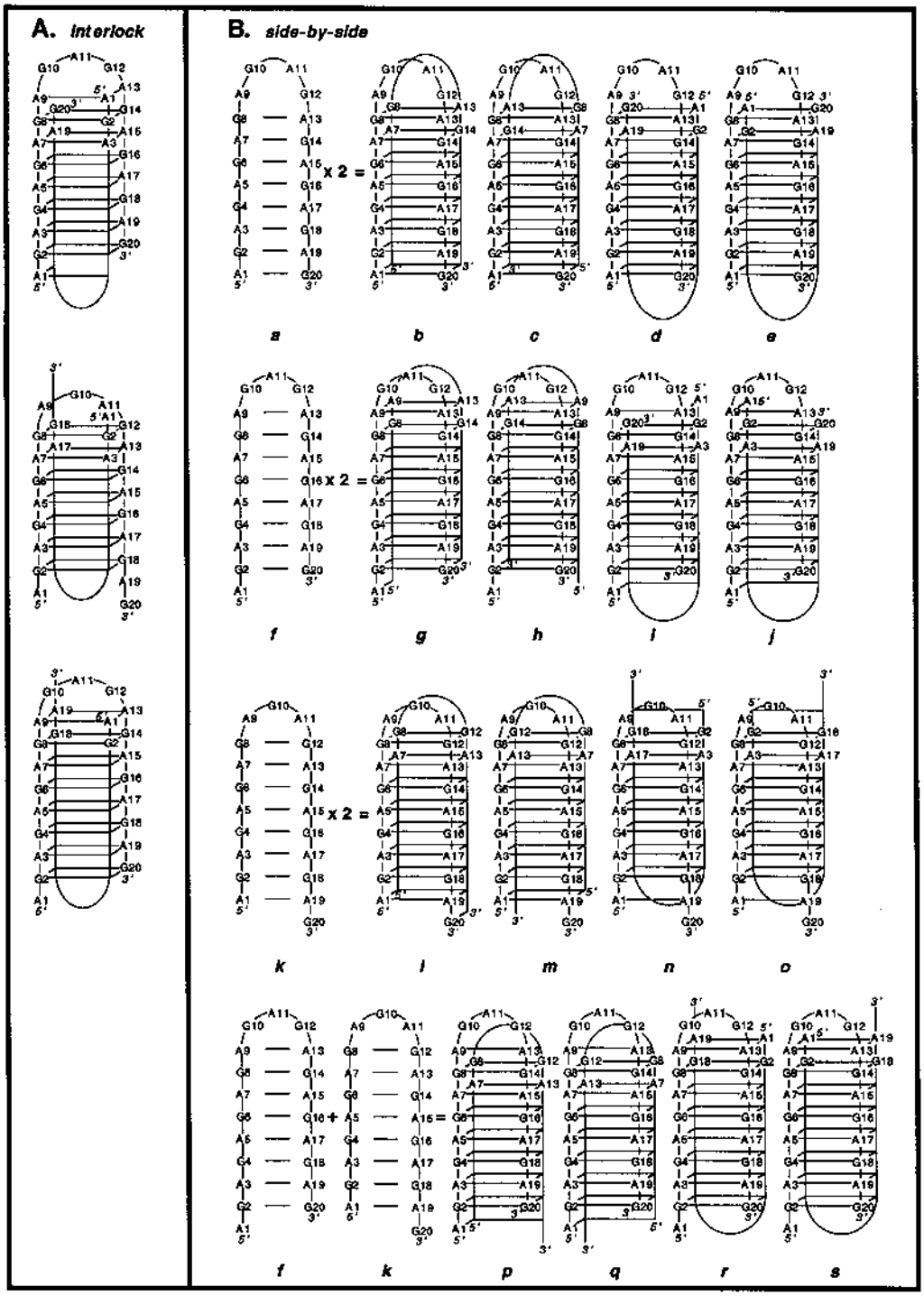

Figure 7. Conformational possibilities for (**A**) interlocking and (**B**) side-by-side hairpin structures of d(A-G)$_{10}$ at pH 7. Planes containing base residues are indicated to facilitate orientation; however, H-bonded base tetrads are known to be formed only by the dG-residues; the dA-residues are intercalated but not H-bonded [see accompanying paper, Mukerji *et al.* (13)].

equilibrium from the two-hairpin tetraplex we have observed towards their suggested linear duplex. This interpretation is supported by the absence of a specific cleavage pattern when methylation by DMS was used to probe the conformation of d(A-G)$_{10}$ in this study, or of d(G-A)$_7$G and d(G-A)$_{12}$G in the work of Rippe *et al.* (6). Using unmodified d(A-G)$_{10}$, UV absorption and UVRR data (13) support only the interaction of G-residues.

G-tetrad helices have been recognized for a long time (38–40), and recently hairpin helices containing G-tetrads with T-loops have been described and possibly implicated in telomere physiology (25,27,33,37). This work, together with that of Mukerji *et al.*

(13), demonstrates that G-tetrads separated by non-H-bonded intercalated A residues provides another variant of this type of four-stranded helical structure.

Promoter regions comprised of GAGA·TCTC duplex repeats have been shown to be specifically recognized by transcriptionally related proteins (38,39), and may play a role in nucleosomal organization (40). The types of hairpin structures described here seem worthy of consideration in developing mechanisms to explain recognition by transcriptional proteins of DNA sequences containing GAGA·TCTC repeats, and how such repeats may play a role in nucleosomal organization.

ACKNOWLEDGEMENTS

This work was supported by grants to J.R.F. from the National Institutes of Health (GM42936) and to E.H.B. from the National Science Foundation (BIR9318373). M.C.S. was supported in part by a Sterling postdoctoral fellowship. We thank D. Zhu for help with the ultracentrifugation measurements, and R. Friedman and Y. Kim for preliminary thermal melting studies and helpful discussions.

REFERENCES

1 Manor, H., Rao, B. S. and Martin, R. G. (1988) *J. Mol. Evol.*, **27**, 96–101.
2 Lee, J. S., Johnson, D. A. and Morgan, A. R. (1979) *Nucleic Acids Res.*, **6**, 3073–3091.
3 Lee, J. S., Evans, D. H. and Morgan, A. R. (1980) *Nucleic Acids Res.*, **8**, 4305–4320.
4 Lee, J. S. (1990) *Nucleic Acids Res.*, **18**, 6057–6060.
5 Antao, V. P., Gray, D. M. and Ratliff, R. L. (1988) *Nucleic Acids Res.*, **16**, 719–738.
6 Rippe, K., Fritsch, V., Westhof, E. and Jovin, T. M. (1992) *EMBO J.*, **11**, 3777–3786.
7 Noonberg, S. B., François, J. -C., Garestier, T. and Hélène, C. (1995) *Nucleic Acids Res.*, **23**, 1956–1963.
8 Casasnovas, J. M., Huertas, D., Ortiz-Lombardía, M., Kypr, J. and Azorín, F. (1993) *J. Mol. Biol.*, **233**, 671–681.
9 Murchie, A. I. H. and Lilley, D. M. J. (1994) *EMBO J.*, **13**, 993–1001.
10 Dolinnaya, N. G. and Fresco, J. R. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 9242–9246.
11 Dolinnaya, N. G., Braswell, E. H., Fossella, J. A., Klump, H. and Fresco, J. R. (1993) *Biochemistry*, **32**, 10263–10270.
12 Shiber, M. C., Lavelle, L., Fossella, J. A. and Fresco, J. R. (1995) *Biochemistry*, **34**, 14293–14299.
13 Mukerji, I., Shiber, M. C., Fresco, J. R. and Spiro, T. G. (1996) *Nucleic Acids Res.*, **24**, 5013–5020.
14 Johnson, M. L., Correia, J. J., Halvorson, H. R. and Yphantis, D. A. (1981) *Biophys. J.*, **36**, 575–588.
15 Eisenberg, H. (1990) In Landolt-Bornstein (ed.) *Numerical Data and Functional Relationships in Science and Technology, New Series, Group VII Biophysics*. Springer-Verlag, New York and Berlin, Vol. 1, Nucleic Acids, Subvolume c, pp. 257–272.
16 Vogt, V. M. (1973) *Eur. J. Biochem.*, **33**, 192–200.
17 Jaishree, T. N. and Wang, A. H. -J. (1993) *Nucleic Acids Res.*, **21**, 3839–3844.
18 Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., Jr and Blackburn, E. H. (1987) *Cell*, **51**, 899–908.
19 Fresco, J. R., Klotz, L. C. and Richards, E. G. (1963) *Cold Spring Harbor Symp. Quant. Biol.*, **28**, 83–90.
20 Fodor, S. P. A. and Spiro, T. G. (1986) *J. Am. Chem. Soc.*, **108**, 3198–3205
21 Perno, J. R., Grygon, C. A. and Spiro, T. G. (1989) *J. Phys. Chem.*, **93**, 5672–5678.
22 Record, M. T., Anderson, C. F. and Lohman, T. M. (1978) *Q. Rev. Biophys.*, **11**, 103–178.
23 Blake, R. D., Massoulie, J. and Fresco, J. R. (1967) *J. Mol. Biol.*, **30**, 291–308.
24 Jin, R., Gaffney, B. L., Wang, C., Jones, R. A. and Breslauer, K. J. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 8832–8836.
25 Sundquist, W. I. and Klug, A. (1989) *Nature*, **342**, 825–829.
26 Sen, D. and Gilbert, W. (1990) *Nature*, **344**, 410–414.
27 Kang, C. H., Zhang, X., Ratliff, R., Moyzis, R. and Rich, A. (1992) *Nature*, **356**, 126–131.
28 Breslauer, K. J. (1986) In Hinz, H. J. (ed.), *Thermodynamic Data for Biochemistry and Biotechnology*. Springer-Verlag, New York, pp. 402–427.
29 Marky, L. A. and Breslauer, K. J. (1987) *Biopolymers*, **26**, 1601–1620.
30 Fresco, J. R. and Alberts, B. M. (1960) *Proc. Natl. Acad. Sci. USA*, **46**, 311–321.
31 Lomant, A. J. and Fresco, J. R. (1975) *Progress in Nucleic Acids Research*, **15**, 185–218.
32 Nikonowicz, E., Roongta, V., Jones, C. R. and Gorenstein, D. G. (1989) *Biochemistry*, **28**, 8714–8725.
33 Smith, F. W. and Feigon, J. (1992) *Nature*, **356**, 164–168.
34 Fresco, J. R. and Massoulie, J. (1963) *J. Am. Chem. Soc.*, **85**, 1352–1353.
35 Arnott, S., Chandrasekaran, R. and Marttila, C. M. (1974) *Biochem. J.*, **141**, 537–543.
36 Zimmerman, S. B., Cohen, G. H. and Davies, D. R. (1975) *J. Mol. Biol.*, **92**, 181–192.
37 Schultze, P., Smith, F. W. and Feigon, J. (1994) *Structure*, **2**, 221–233.
38 Biggin, M. D. and Tjian, R. (1988) *Cell*, **53**, 699–711.
39 Soeller, W. C., Oh, C. E. and Kornberg, T. B. (1993) *Mol. Cell. Biol.*, **13**, 7961–7970.
40 Tsukiyama, T., Becker, P. B. and Wu, C. (1994) *Nature*, **367**, 525–532.
41 Mukerji, I., Shiber, M. C., Spiro, T. G. and Fresco, J. R. (1995) *Biochemistry*, **34**, 14300–14303.

This is paper no. 24 in the series Polynucleotides, of which the last is ref. 41.