Structure of $d(GT)_n \cdot d(GA)_n$ Sequences: Formation of Parallel Stranded Duplex DNA†

Markus W. Germann,*,‡ Bernd W. Kalisch,§ and Johan H. van de Sande§

Kimmel Cancer Institute and Department of Microbiology and Immunology, Thomas Jefferson University, 233 South 10th Street, Philadelphia, Pennsylvania 19107, and Department of Medical Biochemistry, The University of Calgary, 2500 University Drive N.W., Calgary, Alberta, Canada T2N 1N4

Received October 29, 1997; Revised Manuscript Received July 16, 1998

ABSTRACT: Alternating polypurine sequences exhibit remarkable polymorphism. In this study, we report that dGA·dGT sequences form parallel stranded duplex DNA at neutral pH. Using two model hairpins, 3′-$d(GT)_3$-5′5′-$T_4(AG)_3$-3′ (I) and 3′-$d(GT)_4$-5′5′-$T_4(AG)_4$-3′ (II), containing 5′5′ linkages which direct parallel strand formation, we systematically explored the spectroscopic and thermodynamic properties of parallel stranded $d(GA)_n \cdot d(GT)_n$. The parallel stranded hairpins are remarkably stable structures with T_M's of 41.5 and 47.5 °C (in 0.4 M NaCl) for the shorter and longer hairpins, respectively. The van't Hoff enthalpies of 80.7 and 114 kJ mol^{-1} are relatively low but are comparable to a parallel stranded $d(GA)_n$ duplex. On the basis of the spectroscopic and electrophoretic characteristics, we conclude that parallel strand formation is not restricted to hairpin systems, but also readily occurs in unconstrained dimeric duplexes with the appropriate sequence homologies. Both melting curves and electrophoretic analyses of parallel stranded heteroduplexes in which the sequence enforces specific base pairing demonstrate that G-G and A-T base pairs are formed in $d(GA)_n \cdot d(GT)_n$ segments.

In parallel stranded (PS)[1] DNA, a recent addition to the family of duplex DNA structures, the two complementary chains have the same polarity, and the bases form reverse Watson−Crick or Donohue base pairs (*1*−*3*). The sequence requirements and stability of PS DNA have been investigated in several laboratories (*1*, *4*−*8*). AT base pairs form PS DNA most readily while single GC base pairs embedded in AT tracts generally destabilize the resulting duplex (*9*; Germann et al., unpublished results). Parallel stranded duplex structures are not restricted, however, to reverse Watson−Crick pairing. Homooligomeric polyribopurine or polyribopyrimidine PS duplex structures were first observed under acidic conditions (*10*, *11*). In recent years, many examples of PS duplexes have been reported (*7* and references cited therein). Other reports demonstrated that oligo- and polydeoxynucleotides are also capable of forming PS duplexes at low pH (*12*, *13*). The sequence requirement for the formation of PS structures was further relaxed by reports that some mixed sequences, again under acidic conditions, could also from parallel stranded structures (*14*, *15*).

Polypurine regions, notably $d(GA)_n$ repeats, are abundant in eukaryotic DNA (*16*). These repeats are located predominantly in noncoding regions, introns, and 5′/3′ flanking regions. Polypurine·polypyrimidine stretches in chromatin are hypersensitive to single strand specific nucleases, indicative of a non-B-DNA duplex conformation. Some of these sequences are highly conserved and are required for biological processes (*17*−*21*). It has been known for some time that oligonucleotides containing $d(GA)_n$ repeats are inefficient agents for either triplex formation or antisense approaches, partly due to the inherent tendency of these sequences to self-associate (*22*). These observations have provided an impetus for the study of $d(GA)_n$ sequences which exhibit a surprising polymorphism (*5*, *23*−*26*). The unusual behavior of $d(GA)_n$ sequences has led to several proposed structures ranging from single-stranded helices (*24*), antiparallel duplexes (*27*), and multistranded structures (*28*) to parallel stranded duplexes (*5*, *25*). Parallel strand disposition of alternating d(GA) sequences was established from pyrene fluorescence quenching, while additional structural information such as base pairing and dinucleotide repeat was deduced from chemical modification studies, model building, and CD studies (*5*). This observation prompted us to examine additional sequences for their potential to form PS structures and to further explore the conformational variability of homopurine sequences. Previous work on parallel stranded duplex structures demonstrated that AT base pairs are compatible with PS duplex formation in either a homoligomeric or an alternating sequence environment (*29*, *30*). Based on our own work and that of others (*5*, *31*) for $d(GA)_n$, we hypothesize that $d(GA)_n$ and $d(GT)_n$ are logical candidates for PS heteroduplex formation.

† This work was in part supported by grants from the NIH GMOD55401 to M.W.G. and NSERC to J.H.v.d.S.

* Corresponding author: Kimmel Cancer Institute, Thomas Jefferson University, BLSB 809, 233 S. 10th St., Philadelphia, PA 19107. FAX: 215-923-2117. Phone: 215-503-4581. EMAIL: mwg@lac.jci.tju.edu.

‡ Thomas Jefferson University.

§ The University of Calgary.

[1] Abbreviations: PS, parallel stranded; APS, antiparallel stranded; bpb, bromophenol blue; CD, circular dichroism; UV, ultraviolet; Tris, tris(hydroxymethyl)aminomethane; TBE, Tris−borate−ethylenediaminetetraacetate; TBM, Tris−borate−magnesium chloride; DMS, dimethyl sulfate; DEPC, diethyl pyrocarbonate; CCD, cyclic capture and dissociation.

S0006-2960(97)02684-6 CCC: $15.00
Published on Web 08/25/1998

In this investigation, our goal was to study the nature of the interaction of $d(GA)_n$ sequences with $d(GT)_n$ tracts. Given the conformational variability of $d(GA)_n$ sequences, it is preferable to first investigate a well-defined model system which will provide unambiguous results, and subsequently study less constrained systems. This approach has proven useful in the characterization of several parallel stranded duplex DNA structures (*1*, *6*, *29*, *32*, *33*). We will demonstrate that parallel $d(GT)_n$·$d(GA)_n$ structures can form intramolecularly. However, we have also found that PS heteroduplexes are not confined to hairpin systems containing unnatural linkages, but also form from unconstrained duplexes.

MATERIALS AND METHODS

Synthesis and Purification. Oligodeoxynucleotides were synthesized using phosphoramidite chemistry. 3′-DMT phosphoramidites were prepared as described previously (*1*).

Gel Electrophoresis. Synthesis products were purified under denaturing conditions on a 15% polyacrylamide gel (5% cross-linking) in 8 M urea, TBE (90 mM Tris−borate and 5 mM EDTA, pH 8.3). Bands containing the product were visualized by UV-shadowing, excised, eluted by diffusion, and desalted on a Sephadex G-25 column. Purified oligodeoxynucleotides were analyzed on native polyacrylamide gels (pH 8.3, 90 mM Tris−borate, 15 mM $MgCl_2$). The samples were applied in 10 μL of running buffer containing 10% sucrose. Native polyacrylamide gels (15 or 20% with 5% cross-linking) were run at 4 °C and 10 V/cm, and bands were visualized by UV-shadowing or autoradiography. Gels were also stained with 1 μg/mL ethidum bromide in 50 mM NaCl at 4 °C.

Concentration Determination. The concentrations of the oligonucleotides were determined from the absorbance of the mononucleotides at 90 °C in 5.0 M $NaClO_4$, pH 7.2 (*34*). [*E*(r) dpA, 15 500 M^{-1} cm^{-1}; dpC, 6980 M^{-1} cm^{-1}; dpG, 11 200 M^{-1} cm^{-1}; dpT, 9360 M^{-1} cm^{-1}]. Calculated oligonucleotide extinction coefficients at 260 nm under these conditions were as follows: I, 1.79×10^5 M^{-1} cm^{-1}; II, 2.26×10^5 M^{-1} cm^{-1}; IV, 1.98×10^5 M^{-1} cm^{-1}; V, 1.55×10^5 M^{-1} cm^{-1}; VI, 2.17×10^5 M^{-1} cm^{-1}; VII, VIII, 2.36×10^5 M^{-1} cm^{-1}; IX, 1.65×10^5 M^{-1} cm^{-1}; X, XI, 1.28×10^5 M^{-1} cm^{-1}.

Chemical Modifications. ^{32}P-Labeled VI and VII were incubated in 20 mM Tris-HCl, pH 7.0, with either 10 mM $MgCl_2$ (duplex form) or 1 mM EDTA (single-strand form). For heteroduplex reactions, the second strand was added in a 4-fold excess to yield 0.2 OD/ml duplex in the reaction mixture. All modification reactions were done at 4 °C under the following conditions: DEPC, 10% (v/v) for 30 min; DMS, 0.05% for 20 min. Potassium permanganate reaction: 1 μL of a 1:10 dilution of a fresh 100 mM $KMnO_4$ in 3 M tetraethylammonium chloride was added to 10 μL of oligonucleotide solution for 20 min (*35*). Following precipitation, the oligonucleotides were cleaved in 1 M piperidine (30 min at 95 °C), dried, and applied to 20% sequencing gels. Autoradiograms were scanned with a Mustak 12000 II EP scanner and analyzed using ImagePC.

Spectroscopy. Ultraviolet (UV) absorption spectra and thermal denaturation profiles were recorded with a Varian Cary 3E spectrophotometer equipped with thermostated cuvette holders. Absorbance was measured at 260 nm once a minute, and the temperature was increased at 0.5 °C/min.

3'-d(GT)$_3$-**5'5'**-T$_4$(AG)$_3$-3'	I
3'-d(GT)$_4$-**5'5'**-T$_4$(AG)$_4$-3'	II
3'-d(GA)$_4$-**5'5'**-T$_4$(AG)$_4$-3'	III
5'-d(GA)$_7$G	IV
5'-d(GT)$_7$G	V
5'-d(GT)$_7$GA$_4$	VI
5'-d(GA)$_7$GT$_4$	VII
5'-dT$_4$(GA)$_7$G	VIII
5'-dAAAGAGAAATAT	IX
5'-dTTTTGTGTATAT	X
5'-dTTTGTGTTATAT	XI

FIGURE 1: Sequences of $d(GT)_n$ and $d(GA)_n$ oligonucleotides investigated in this study.

The concentration-independent melting curves of the DNA duplexes were analyzed using a two-state model in which the absorbances of the helix and coil forms vary linearly with temperature. The thermal behavior of the two states is described by α, the absorbance at 0 K, and the temperature coefficient β (i.e., $A_H = \alpha_H + \beta_H T$ and $A_C = \alpha_C + \beta_C T$), where the indices H and C refer to helix and coil forms respectively, and T is absolute temperature.

$$A = A_H(1 - \Theta) + A_C\Theta \quad (1)$$

The degree of transition, Θ, for a hairpin to coil transition is given by $K/(1 + K)$, where K is the equilibrium constant for the helix−coil transition for intramolecular hairpin structures. In addition to a 6-parameter fit of the absorbance versus temperature, we have also analyzed the differentiated melting profile, which results in the 5-parameter fit described in our earlier work (*6*). Entropies for the hairpin denaturations are calculated from $\Delta S^o = \Delta H^o_{vH}/T_M$.

Circular dichroism spectra were recorded on a Jasco J-500C spectropolarimeter, equipped with thermostated cuvette holders and interfaced to a 386PC. The circular dichroism $\Delta\epsilon$ (M^{-1} cm^{-1}) is expressed per M nucleotide.

RESULTS AND DISCUSSION

Description of the Systems. To clearly establish what type of structures are formed by $d(GT)_n$·$d(GA)_n$ sequences, we have utilized hairpins as model systems, exploiting the ease of intramolecular duplex formation. The parallel strand orientation is enforced by the use of a 5′5′ linkage at the stem−loop interface. In such a system, concentration-independent behavior is due to the presence of either single strands or intramolecularly formed structures. The latter constitutes strong evidence for PS hairpin formation, if the sequences also migrate similarly to previously characterized hairpin structures of comparable size on nondenaturing gels. The oligonucleotides studied are shown in Figure 1. Sequences I and II are designed to promote intramolecular PS hairpin formation. The T_M and gel electrophoretic mobilities of I and II are predicted to change gradually and predictably if hairpins are formed that only differ in the stem length. Sequences IV+V and VI+VII are candidates or components,

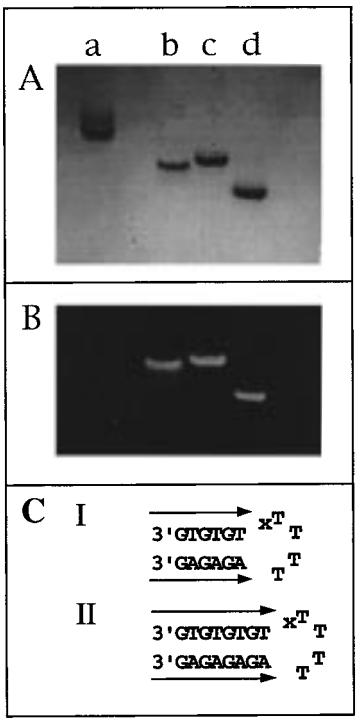

FIGURE 2: Native gel electrophoretic analysis of $d(GT)_n$,$d(GA)_n$ hairpin sequences. 0.2 $A_{260\ nm}$ unit of oligonucleotides was applied in 10 μL of running buffer to a 15% native gel and run for 14 h at 4 °C. Lane a, oligo $dT_{19}-dT_{24}$ single-stranded markers; lane b, 5′-$dT_8C_4A_8$-3′; lane c, oligonucleotide II; lane d, I. Panel A: Visualization of the gel by UV shadowing. Panel B: Staining of the same gel with ethidium bromide at 4 °C. Panel C: Potential intramolecular parallel stranded hairpin structures formed by I and II.

IX, X, XI, for intermolecular PS duplex formation by $d(GT)_n$·$d(GA)_n$ sequences.

Hairpin Substrates. (*A*) *Gel Electrophoresis*. Oligonucleotides I and II were analyzed by gel electrophoresis under native conditions (Figure 2, panel A). All oligonucleotides showed only the presence of a single species. No bands due to multimers or bands with unexpected mobilities were found for any of the samples. It should be noted that the oligonucleotides were applied at high strand concentration (90−110 μM) which would promote intermolecular structure formation. This is particularly relevant since it is known that $d(GA)_n$ tracts, which are an integral component of our sequences, have been show to self-associate (*5*, *22*, *25*, *36*). All sequences migrated faster than the single-stranded dT_n control. Sequence I exhibits the fastest mobility on the gel while the slightly larger sequence II migrates quite similarly to 5′d-$T_8C_4A_8$, which runs as a hairpin structure (*32*, *33*). This indicates that oligonucleotides I and II adopt a hairpin conformation, with the former migrating faster because of the shorter hairpin stem (Figure 2, panel C). This behavior is reminiscent of previously characterized parallel stranded hairpins with different stem lengths (*6*, *32*). The observed staining of hairpins I and II with ethidium bromide is also consistent with a double-helical structure (Figure 2, panel B). Previous results have established that parallel stranded DNA intercalates ethidium bromide readily (*1*, *32*). In contrast, the parallel stranded duplex structure reported for alternating d(GA) sequences which are thermally less stable was reported to show very little fluorescence upon staining with ethidium bromide (*5*). Taken together, this provides strong evidence for hairpin formation and therefore for parallel stranded $d(GT)_n$·$d(GA)_n$.

(*B*) *UV and CD Spectroscopy*. Previous reports have demonstrated that the spectroscopic properties of parallel stranded DNA differ significantly from those of APS DNA of the same sequence composition (*29*, *30*). This was in fact useful for the characterization and detection of parallel stranded DNA formation under different conditions (*6*, *33*). In the present case, comparisons of the spectroscopic properties with the earlier parallel stranded DNA sequences must be made with caution because of the different base composition. In addition, with the present sequence composition, a corresponding antiparallel control cannot be constructed, since antiparallel strand formation between $d(GA)_n$ and $d(GT)_n$ sequences does not occur.

To aid in the characterization of the structures that may be formed under different conditions, we have recorded hyperchromicity profiles (*A*85 °C/*A*5 °C) under different conditions (Figure 3A). The parallel stranded hairpin II shows 20% hyperchromicity at 270 nm, consistent with the formation of a stacked structure. The striking similarity of the profiles recorded in either 400 mM NaCl or 15 mM $MgCl_2$ supports the hypothesis that the same structures were formed under both of these conditions. As expected, the hyperchromicity profiles differ from those of regular antiparallel DNA and also from previously characterized parallel stranded duplexes containing AT base pairs (*29*, *30*). The maxima of the hyperchromicity plots for hairpin II were at 273 nm. Similar data were also obtained for hairpin I which underlines the structural similarity between hairpins I and II (data not shown). In contrast, the profile recorded for the control, sequence III, exhibited a λ_{max} at 262 nm, and differs markedly from that obtained for the parallel stranded hairpins I and II.

The recorded CD spectra of hairpin II, with a positive maximum at 263 nm, zero-crossover at 251 nm, and a negative maximum at 241 nm, provide future support that sequence II forms a stable parallel stranded hairpin structure that can be denatured at high temperature (Figure 3B). Both positive and negative lobes of the CD spectrum of hairpin II have nearly the same intensity. The CD spectrum of hairpin II again differs markedly from that obtained for sequence III (data not shown).

(*C*) *UV Melting Experiments*. Important information about the thermodynamic parameters that govern the formation of this novel structure can be obtained from UV melting profiles (Table 1). The initial experiments, carried out in 0.4 M NaCl solutions, established that the melting temperatures for hairpins I and II are concentration independent and reversible (Figure 3C). This further supports the conclusion that these sequences form hairpin structures. The salt dependence of the melting (dT_M/d log [NaCl] = 13.5 °C) is within the range characteristic for parallel or antiparallel duplex DNA (*30*). Melting of single-stranded structures, with their inherently lower charge density, is expected to exhibit a much lower salt dependence of the T_M. Within the framework of counterion condensation, the thermal denaturation of triplexes and tetraplexes are expected to show a much higher d T_M/d log [NaCl] unless specifically bound cations lower the intrinsic high charge density (*37*).

The UV melting data were analyzed using a two-state melting model (Table 1). As is evident from Figure 3D,

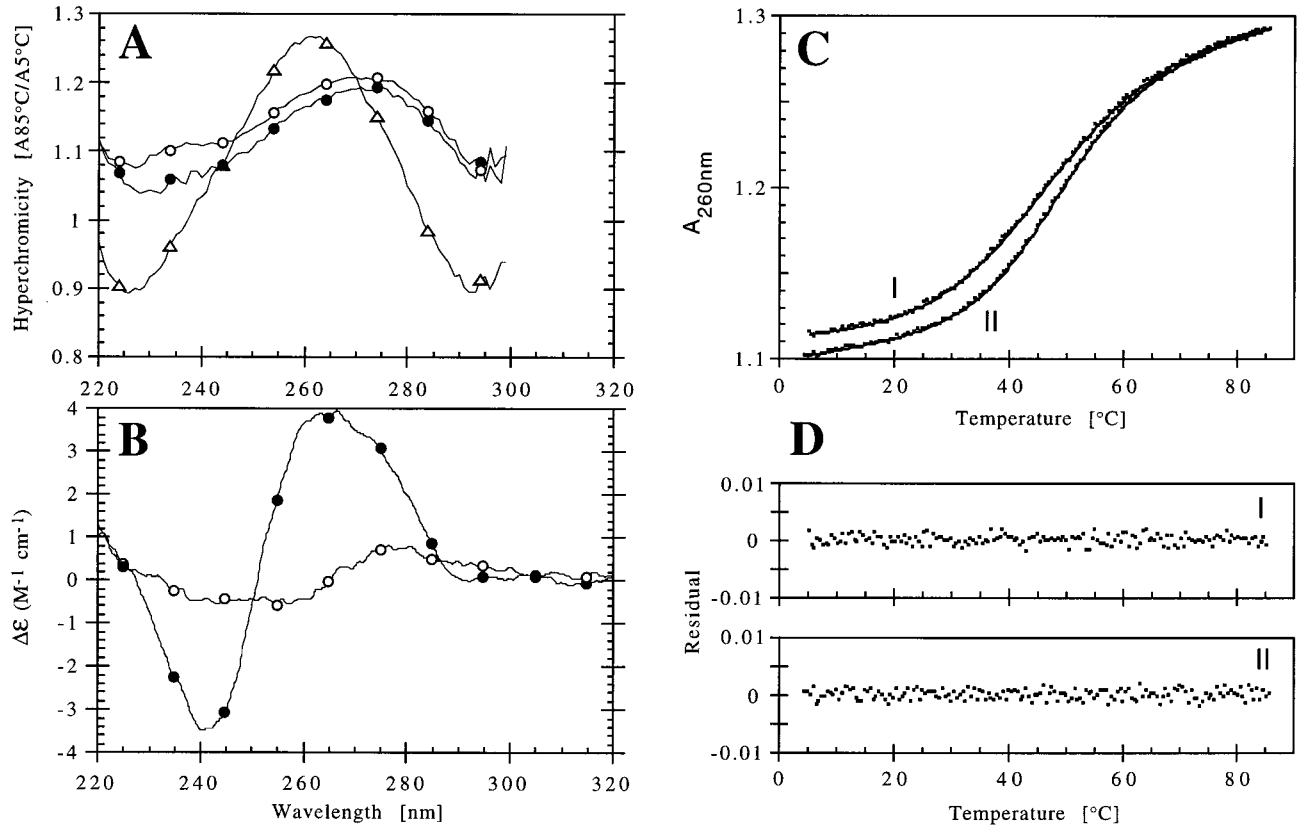

FIGURE 3: Hyperchromicity and CD spectra of II. Panel A: Hyperchromicity profile of 4 μM oligonucleotide II recorded in 400 mM NaCl, 0.1 mM EDTA, 10 mM sodium phosphate buffer, pH 7.0 (●), and 15 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.0 (○). Sequence III was also analyzed in 400 mM NaCl, 0.1 mM EDTA, 10 mM sodium phosphate buffer, pH 7.0 (△). Panel B: CD spectra of 4 μM II, recorded under native 5 °C (●) and denatured conditions 85 °C (○), in 400 mM NaCl, 0.1 mM EDTA, 10 mM sodium phosphate buffer, pH 7.0. The circular dichroism $\Delta\epsilon$ is expressed per M nucleotide. Panel C: Melting curves of I and II in 400 mM NaCl, 0.1 mM EDTA, 10 mM sodium phosphate buffer, pH 7.0. Panel D: Deviation of the melting profile and fitted T_M curve for I and II.

Table 1[a]

	oligonucleotide	
	3′−d(GT)$_4$−5′5′−T$_4$(AG)$_4$−3′ (**II**)	3′−d(GT)$_3$−5′5′−T$_4$(AG)$_3$−3′ (**I**)
	0.4 M NaCl[b]	
T_M (°C)	47.5 (0.5)	41.5 (0.7)
enthalpy (kJ mol^{-1})	114 (9)	80.7 (14)
h[c] (kJ mol^{-1})	16.2	16.1
entropy (kJ mol^{-1} K^{-1})	0.354 (0.03)	0.257 (0.04)
dT_M/d log [NaCl][d]	13.5	nd
	15 mM $MgCl_2$[e]	
T_M (°C)	47.7 (0.4)	nd
enthalpy (kJ mol^{-1})	112 (7)	nd
h (kJ mol^{-1})	16.0	nd
entropy (kJ mol^{-1} K^{-1})	0.349 (0.02)	nd

[a] T_M data are the average of 2−4 independent measurements. Values in parentheses indicate the standard deviation of the measurements. [b] 10 mM sodium phosphate, pH 7.0. The melting curves were found to be concentration independent for the strand concentrations investigated (3.6−20 μM). [c] Enthalpy per nearest-neighbor interaction. [d] Salt dependence of the melting point determined for 0.1−0.4 M NaCl. [e] 10 mM Tris, pH 7.0.

this model fits the data well throughout the entire melting curve. The observed T_M for these hairpins is remarkably high, at 41.5 °C for hairpin I and 47.5 °C for hairpin II. In contrast, the van't Hoff enthalpy for both hairpins is rather low at 80.7 and 114 kJ mol^{-1}, respectively. A previously characterized PS hairpin with a homooligomeric AT stem of the same length as hairpin II exhibits a significantly higher enthalpy, 190 kJ mol^{-1}, and slightly higher T_M, 54 °C. For a PS hairpin with an alternating AT stem the enthalpy is slightly lower at 106 kJ mol^{-1} with a much lower T_M of 32.3 °C (*6*, *32*).[2] The observed low enthalpy and high T_M either is an intrinsic feature of these novel structures or alternatively may be due to multistate melting due to slipped hairpin structures. It should be noted that we have been unable to detect any intermolecular structures even under very high strand conditions (data not shown).

A comparison of the enthalpy calculated per nearest-neighbor interaction of 16.2 and 16.0 kJ mol^{-1} for hairpins I and II, respectively, shows that very similar results are obtained for both parallel stranded hairpins. This is the expected result for hairpin structures differing only in stem length and argues against alternate structures. Thermodynamic data collected for II in 15 mM $MgCl_2$ and 10 mM Tris, the running buffer for native gel electrophoresis, are almost identical to those obtained in 0.4 M NaCl, corroborating the conclusion drawn from the analysis of the hyperchromicity profiles that similar structures are formed under these conditions. It is instructive to compare the thermodynamic data obtained for hairpin II to the parallel stranded d(GA)$_n$ duplex reported previously (*5*). The d(GA)$_n$ duplex is characterized by a relatively low enthalpy of 15−15.7 kJ mol^{-1} per nearest-neighbor interaction. Nearly identical data were obtained for the parallel stranded heteroduplexes investigated in this study (16 kJ mol^{-1}) despite their high

[2] T_M's were recalculated for 0.4 M NaCl solutions.

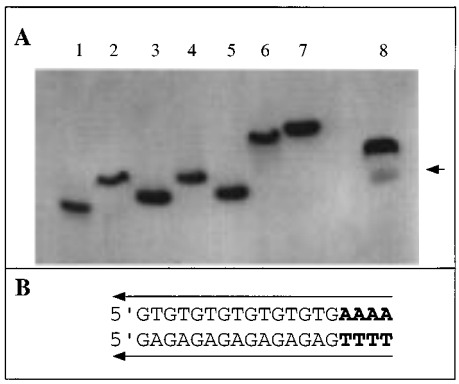

FIGURE 4: Native gel electrophoretic analysis of oligonucleotides IV−VII on a 20% native gel, run for 19 h at 4 °C. Labeled samples were applied in 10 μL of running buffer containing 10% sucrose at a concentration of 0.1 $A_{260\ nm}$/mL. Unlabeled complementary oligonucleotides were present at a concentration of 0.12 $A_{260\ nm}$/mL except for lane 8. (Panel A) Lane 1, [^{32}P]-d(T)$_{14}$; lane 2, [^{32}P]-IV; lane 3, [^{32}P]-V; lane 4, ^{3}[^{32}P]-VI; lane 5, [^{32}P]-V+IV; lane 6, [^{32}P]-VI+VII; lane 7, antiparallel duplex control [^{32}P]-V+d(CA)$_7$C; lane 8, [^{32}P]-IV + a 10-fold excess of unlabeled V. The arrow indicates that a small amount of IV remains which comigrates with the product in lane 2. The majority of [^{32}P]-IV is engaged in a complex with V. (Panel B) Parallel stranded duplex model structure formed from oligonucleotides VI and VII.

T_M, which is due to hairpin formation. A similar trend was also observed for the entropy of nearest-neighbor interaction.

Duplex Substrates. (*A*) *Gel Electrophoresis*. The hairpin model system clearly illustrated that d(GA)$_n$•d(GT)$_n$ can form parallel stranded duplex structures and provided a standard for investigating unconstrained systems. Here an additional complication arises, due to the inherent ability of d(GA)$_n$ sequences to form homoduplexes. In our initial attempt, we used IV and V to form a parallel stranded heteroduplex. Under native gel electrophoresis conditions, we observed the presence of a single species for IV and V alone (Figure 4, lanes 2, 3). However, sequence IV exhibits a mobility that is very different from that of V. The latter migrates with a mobility similar to a single-stranded control (Figure 4, lane 1). The retarded gel electrophoretic mobility of sequence IV is ascribed to the formation of a parallel stranded dimer structure (*5*, *25*, *31*; Germann et al., unpublished results).

Electrophoretic analysis of a 1:1 mixture of sequences IV and V did not produce evidence for a new duplex, as might be expected from our hairpin data. This implies that the homoduplex structure formed by IV is more stable than the parallel stranded heteroduplex with putative GG and AT base pairing (Figure 4, lane 5). This was investigated in more detail by challenging a preformed, 5′-^{32}P-labeled IV homoduplex with a 10-fold excess of unlabeled V (Figure 4, lane 8). Under these conditions, we indeed observed the formation of a new product with significantly slower mobility that contains both V and IV as components. In addition, we observe a small amount of unreacted end-labeled IV (marked with an arrow). Because no heteroduplex could be detected in the 1:1 mixture while in the 10:1 mixture more than 90% heteroduplex was observed, we conclude that the stability difference between the homoduplex formed by IV and the heteroduplex (IV+V) is relatively small. Indeed, if both constituent stands are modified to contain either A or T tails (i.e., VI and VII), which are known to allow parallel stranded DNA formation, only the heteroduplex is observed (Figure 4, lane 6). The mobility of this structure is similar to that of an antiparallel control duplex (Figure 4, lane 7). The sequence alignment for the heteroduplex is shown in Figure 4, panel B. This figure indicates that if the T_4 tail is introduced at the 5′ end of sequence VII, the stabilizing parallel stranded AT segment could not be formed. It is expected that with a 5′ T_4 tail the heteroduplex would not be stable, which supports the parallel strand alignment. As shown in Figure 5 heteroduplex formation is only observed for (VI+VII) but not for (VI+VIII). This demonstrates that the proper orientation of the T_4 tail is important and supports the parallel strand alignment shown in Figure 4.

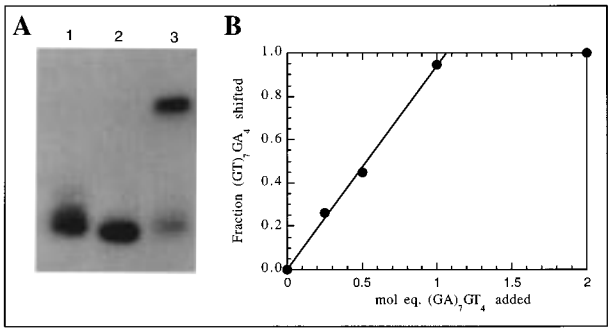

FIGURE 5: Stabilization and stoichiometry of heteroduplex formation. (Panel A) Native gel electrophoretic analysis. Lane 1, [^{32}P]-VI+VIII; lane 2, [^{32}P]-VI; lane 3, [^{32}P]-VI+VII. Samples were applied in 10 μL of running buffer containing 10% sucrose. The concentration of labeled VI was 1.0 $A_{260\ nm}$/mL; the second strands in lanes 1 and 3 were present at an equimolar ratio. (Panel B) Stoichiometry of heteroduplex (V+VI) formation. Unlabeled VII was added to a solution of 6.3 μM [^{32}P]-VI. Products corresponding to heteroduplex and unreacted [^{32}P]-VI were separated by native gel electrophoresis. The extent of heteroduplex formation was obtained from scintilation counting of shifted and unshifted bands.

Sequences VI and VII also allow us to address the stoichiometry of the structure formed. Our hairpin data demonstrated intramolecular duplex formation which strongly suggests a 1:1 strand stoichiometry. This was examined by gel shift experiments titrating a know amount of ^{31}P-labeled VI with increasing concentrations of VII (Figure 5). The results show that 1 equiv of VI is needed for the formation of a single product, the heteroduplex. Under none of the gel electrophoretic conditions could we observe bands that might be due to higher order structures.

(*B*) *UV and CD Spectroscopy*. As anticipated, the hyperchromicity profile of VII (Figure 5A), which forms a homoduplex structure with $\lambda_{max} = 262$ nm, is remarkably similar to that observed for III (Figure 3A). In contrast, the profile of VI ($\lambda_{max} = 270$ nm) shows only very little intensity (Figure 5A). The hyperchromicity profile of the 1:1 mixture of these sequences, with a λ_{max} of 272 nm, differs markedly from that of its constituent strands, but resembles the profiles recorded for hairpins I and II. These observations indicate that in the presence of VI the homoduplex formed by VII dissociates to form a parallel stranded heteroduplex. This is consistent with our gel electrophoretic data.

This is further substantiated by the CD spectra shown in Figure 5. The CD spectra of the constituent strands VI and VII are shown in Figure 5B. Particularly striking in the spectrum, reminiscent of an A-type structure, is the high intensity observed for VII at 265 nm and the low intensity at 242 nm. In addition, we observed a small negative dip at

290 nm. Similar CD spectra were also recorded for sequence III and have also been reported for other d(GA)$_n$ sequences (*5*, *38*). Not surprisingly, sequence VI showed very little intensity. At high temperature, under denaturing conditions, these characteristics all disappeared, as expected.

The CD spectrum of the heteroduplex (VI+VII) differs noticeably from that of the constituent strands. Particularly striking is the relatively low intensity of the CD signal at 5 °C. This is not due to melting since the T_M of the heteroduplex under these conditions is about 24 °C (data not shown). Under native conditions, the positive peak near 260 nm and negative peak at 242 nm have nearly the same intensity and are very similar to the CD spectrum of the parallel stranded hairpin II (Figure 5C). The similar hyperchromicity profiles as well as CD spectra of the heteroduplex (VI+VII) compared to the hairpin II provide strong evidence that similar structures are formed.

(*C*) *Base Pairing*. Our data demonstrate that d(GA)·d(GT) sequences form parallel stranded duplex DNA. The structure and base pairing in these duplexes cannot be deduced from the data presented thus far. In particular, we are interested in determining whether GG and AT base pairs occur, or, alternatively, if GA and GT base pairs are formed. Both parallel GG and AT base pairs have been reported previously, although not within the context of an alternating sequence motif (*31*, *39*). GA and GT base combinations, in contrast, have been reported to be destabilizing if embedded in an oligo-d(A)·oligo-d(T) stretch of a parallel stranded duplex (*39*).

To resolve the question of what base pairs form in a parallel d(GA)$_n$·d(GT)$_n$ duplex, three additional sequences (IX, X, XI, Figure 1) were designed in which a (GT)$_2$ and (GA)$_2$ core is embedded within a sequence which both stabilizes and enforces the alignment of the central core. The function of the alternating AT stretch is to ensure proper sequence alignment and destabilize slipped structures, thereby simplifying the analysis. The two potential duplexes are shown in (Figure 6A,B). For duplex A, the central core was designed to enable GT and GA base pairs while for duplex B, GG and AT base pairs are possible. The UV melting profile of these sequences (Figure 6) demonstrates that only duplex B, in which the core alignment allows GG and AT base pairing, was stable. Nondenaturing gel electrophoretic analysis of these sequences (Figure 7) also provides clear evidence for the formation of duplex B (lane 2). The mobility of duplex B was comparable to a decamer control duplex (lane 6). By themselves, all oligonucleotides migrated as single strands.

Chemical modification studies provide additional information by probing which functional groups are accessible and therefore not involved in base pairing. For these experiments, the thermally more stable duplex (VI+VII) was used. $KMnO_4$ oxidizes unpaired or unstacked T (T > C ≫ G > A). This reagent reacts more readily with single strands than double-stranded DNA and has been used to detect mismatches and insertions and to characterize PS DNA (*42*, *43*). The reactivity of single-stranded VI is considerably higher than for the heteroduplex (VI+VII) as suggested from the amount of unreacted material as well as the protection of T in the heteroduplex (Figure 8A).[3] This implies that T is involved in secondary structure formation.

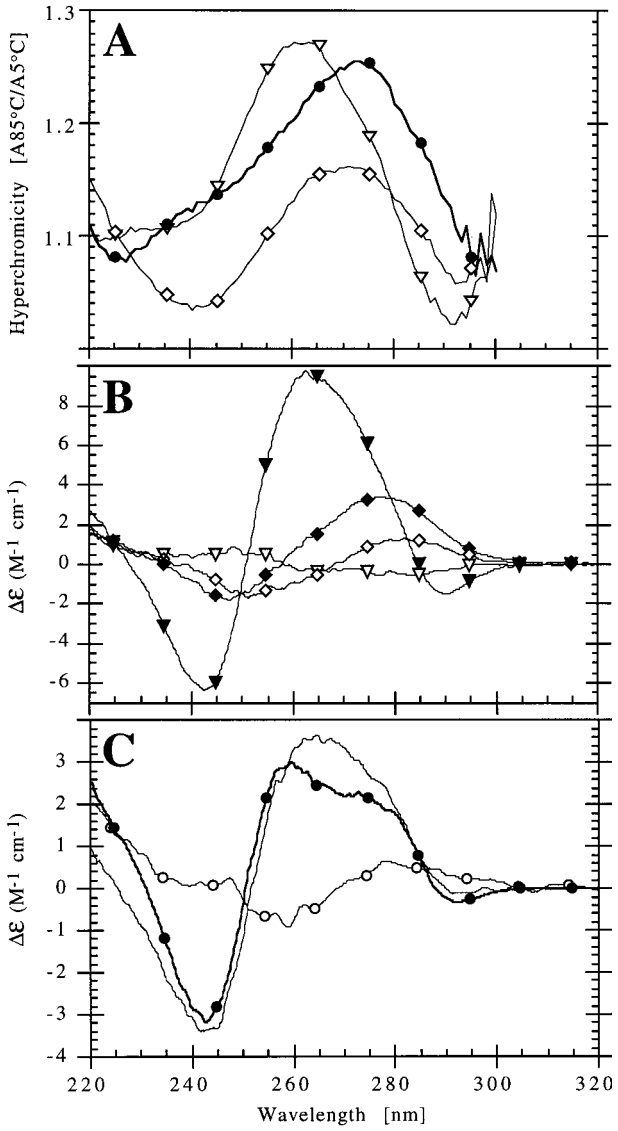

FIGURE 6: Hyperchromicity profiles and CD spectra of VI and VII and the parallel heteroduplex formed from these two sequences. All samples were analyzed in 400 mM NaCl, 0.1 mM EDTA, 10 mM sodium phosphate buffer, pH 7.0. Panel A: Hyperchromicity profiles of VII (▽), VI (◇), and the 1:1 strand mixture of VI and VII (●). Panel B: CD spectra of VII at 5 (▼) and 85 °C (▽) and of VI at 5 (◆) and 85 °C (◇). Panel C: CD spectra of heteroduplex (VI+VII) at 5 (●) and 85 °C (○). The CD spectrum of the parallel stranded hairpin II (—) is shown for comparison. The circular dichroism $\Delta\epsilon$ is expressed per M nucleotide.

DMS methylates the N7 position of guanine unless the N7 group is shielded from the reagent (*44*). Under single-stranded conditions, the reactivities of the G's were somewhat lower for VII than VI (Figure 8B). This indicates an intrinsic difference in reactivity for the two strands. Using labeled VII to monitor the heteroduplex reactivity, we observe only a modest reduction in reactivity (30%). A similar result was also obtained for the accessibility of N7 of G's in d(GA)$_n$ sequences (*5*). If VI is monitored, approximately 40% reduction in reactivity is observed.

[3] The formation of the homoduplex of VII at low ionic strength is dependent on the Mg^{2+} concentration as was observed previously for alternating d(GA)$_n$ sequences (*5*). To obtain single-stranded control sequences for VII, all experiments on single stranded sequences were carried out in the absence of Mg^{2+}.

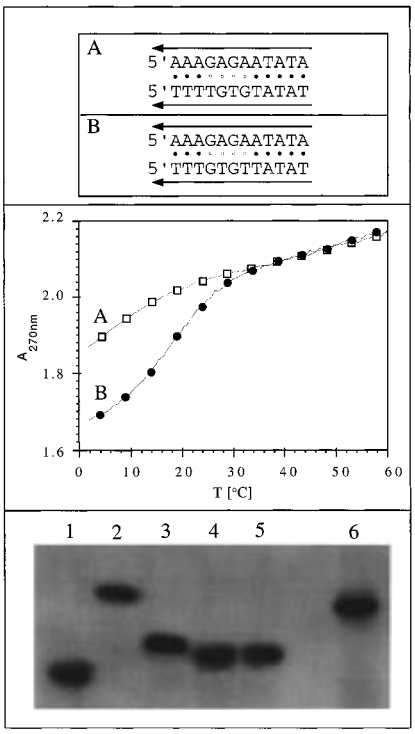

FIGURE 7: Potential parallel stranded bimolecular duplex structures for determination of GG and AT vs GT and AG base pairing. Top panel: (A) Sequence alignment for IX and X. (B) Sequence alignment for IX and XI. Middle panel: UV melting profile of (IX+XI) (●) and (IX+X) (□). The T_M of B is estimated at 17 °C with a van't Hoff enthalpy of about 185 kJ mol^{-1}. All samples were in 400 mM NaCl, 0.1 mM EDTA, 10 mM Na_2HPO_4, pH 7.0, at 92 μM duplex concentration. Bottom panel: Native gel electrophoretic analysis of sequences described in the top panel. All oligonucleotides were applied in 5 μL of running buffer (TBM) containing 10% sucrose to a 18% polyacrylamide gel. The concentration of the applied labeled samples was 10 $A_{260\ nm}$/mL for all lanes. Unlabeled strands were added at a 20% excess where indicated. The gel was run for 17 h at 4 °C. Bands were located by autoradiography. Lane 1, ^{32}P-IX; lane 2, ^{32}P-IX+XI; lane 3, ^{32}P-IX+X; lane 4, ^{32}P-XI; lane 5, ^{32}P−X; lane 6, [^{32}P-dGCGAATTCGC]$_2$ duplex. The mobility shift of sequence ^{32}P-IX in the presence of X is probably the result of cyclic capture and dissociation (CCD) of the labeled strand with X, resulting in the retardation of ^{32}P-IX (40, 41). In order for CCD to be observed, several conditions must be met, inluding that the faster migrating strand is labeled, the second strand in excess, and that the only transiently stable complex (a duplex containing several unpaired bases) migrates slower than either single strand.

These data suggest that N7 of G is not involved in base pairing in either strand.

The reactivities of A and G residues toward DEPC were lowered in the heteroduplex, indicative of secondary structure formation (*5*, *43*) (data not shown). The extent of modification in strands VI and VII was 48% and 57%, respectively, of the single-stranded controls for G_{13}-G_5, while for A_{14}-A_6 41% was detected. In addition, we also observed 51% reactivity for A_{16} which is part of the A tail of the duplex (Figure 8). These data, in conjunction with the DMS results, suggest that N7 of G is accessible and likely not involved in hydrogen bonding. The situation is, however, less clear for A.

The relatively high reactivity of G_{15} in strand VI with DMS may reflect a slight perturbation at the junction of the $(GT)_n$ tract and the A_4 tail in strand VI but not VII of the heteroduplex.[4] It is noteworthy that, in both DEPC and

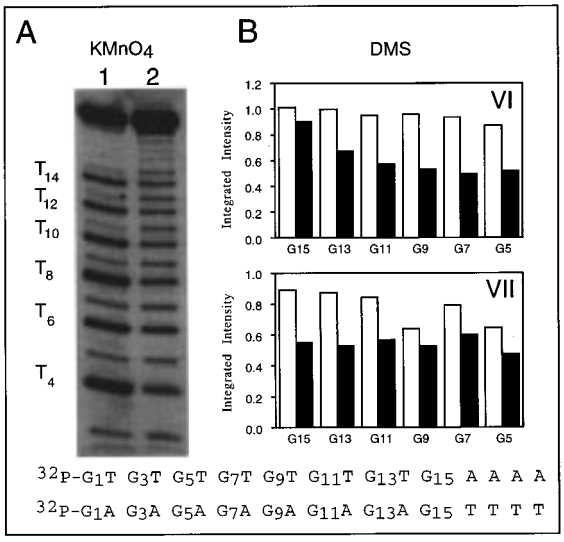

FIGURE 8: Chemical modification of single strands and heteroduplex (VI+VII). Panel A: $KMnO_4$ modification of (lane 1) VI and (lane 2) VI+VII. Panel B: Results of the DMS modification for ^{32}P-labeled VI and VII strands, respectively. Single strands are indicated by open columns while heteroduplex data are represented with filled columns. Strands are numbered starting at the 5′ end.

$KMnO_4$ reactions carried out on the heteroduplex, G_{15} (VI) shows similar reactivity as G_{13} through G_7, supporting base pairing at this junction.

Both UV melting studies and gel electrophoretic analysis indicate that GG and AT base pairs were formed. The similar thermodynamic properties of $d(GA)_n$·$d(GT)_n$ and $d(GA)_n$ duplex formation in conjunction with the inherent sequence resemblance between these two systems prompted us to compare the base pairing for $d(GA)_n$·$d(GT)_n$ to that of the $d(GA)_n$ duplex for which GG and AA base pairing was proposed (*5*). This duplex was reported to have a dinucleotide repeat with the G in a syn and the A in an anti conformation (Figure 9, top). Both the syn conformation and also the dinucleotide repeat are features that are expected to show characteristic NMR spectra which would provide additional evidence for this model. NMR structural data are available for a related sequence, d(CGAGAGA), recorded under low-pH conditions (*45*). In this parallel stranded homoduplex, GG pairs form with hydrogen bonds between N6−N3 and AA pairs with N6−N7 hydrogen bonds. The reported structure differs from the previously proposed model (*5*) in that the exocyclic amino group of G is involved in base pairing. In addition, all bases are in the anti conformation (Figure 9).

Our data suggest that $(GA)_n$·$(GT)_n$ may form G_{anti}-G_{anti} base pairs as reported by Robinson et al. (*45*, *46*). The AT base pair can be accommodated with a similar geometry as the A_{anti}-A_{anti} base pair; however, this places T in a syn orientation. Such a conformation for a pyrimidine is unusual and requires experimental verification. These base pairing schemes are useful for rationalizing the effect of modified bases, which in turn may refine or confirm the proposed base pairing. For instance, the GG base pair proposed for $d(GA)_n$, in the Rippe et al. (*5*) model, predicts that the exocyclic

[4] One of the reviewers suggested that the higher reactivity of G_{15} (strand VI), compared to G_{13} in the heteroduplex, may be indicative of an unpaired base.

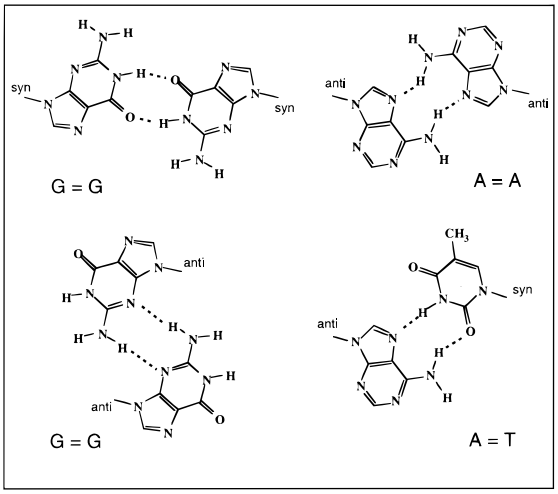

FIGURE 9: Base pairing schemes. G-G and A-A base pairing schemes reported for parallel stranded duplexes (5, 45). Possible base pairing for the A_{anti}-T_{syn} base pair in $(GA)_n$•$(GT)_n$ duplexes with parallel strand disposition.

amino group is not involved in direct hydrogen bonding between the bases. Similarly, in the case of the AA as well as the AT base pair, N7 of A is implicated in hydrogen bonding. Experiments to explore these predictions are currently underway in this laboratory. The most definitive results, however, must await the determination of either an NMR or a crystal structure of these sequences. It is noted, however, that $d(GA)_n$ (*26*; Jovin, personal communication) as well as $d(GA)_n$•$d(GT)_n$ sequences (Germann, unpublished results) are notoriously difficult to study by NMR.

CONCLUSION

In light of the interest in developing antisense oligonucleotides with predictable and exploitable base pairing, it is necessary to understand the polymorphic behavior of nucleic acids. Previous studies have established the extreme polymorphism of $d(GA)_n$ sequences, particularly self-association, which curtails the use of longer stretches of such sequences for potential antisense applications. In this study, we demonstrate that $d(GA)_n$•$d(GT)_n$ sequences form parallel stranded duplex DNA, at neutral pH, in either hairpin or dimeric systems.

The base pairing was investigated by UV melting as well as by gel electrophoresis studies which support G-G and A-T base pair formation. The enthalpic stability of the parallel stranded heteroduplex is very similar to that of the parallel stranded $d(GA)_n$ duplex.

These new findings are important in the design of novel antisense therapeutics as these agents should be available for base pairing with the mRNA target and not engage in self-complex formation that will reduce their efficacy.

These results also indicate that $d(GT)_n$, which in contrast to $d(GA)_n$ sequences does not self-associate, could potentially serve as probes to detect single-stranded $d(GA)_n$ repeats as are present in H-DNA structures. Repetitive $d(GA)_n$ and $d(GT)_n$ sequences are found throughout the eukaryotic genome; the observation that single-stranded $d(GA)_n$ and $d(GT)_n$ sequences are able to recognize and bind to each other may have biological ramifications.

Our findings exemplify the remarkable polymorphism of DNA and add a new member to the family of parallel stranded-duplex DNA.

ACKNOWLEDGMENT

We thank Dr. E. Wickstrom for helpful discussion and critical reading of the manuscript. We are grateful to Mr. M. Brown for technical assistance.

REFERENCES

1. van de Sande, J. H., Ramsing, N. B., Germann, M. W., Elhorst, W., Kalisch, B. W., Kitzing, E. V., Pon, R. T., Clegg, R. C., and Jovin, T. M. (1988) *Science 241*, 551−557.
2. Otto, C., Thomas, G. A., Rippe, K., Jovin, T. M., and Peticolas, W. L. (1991) *Biochemistry 30*, 3062−3069.
3. Zhou, N. Germann, M. W., van de Sande, J. H., Pattabiraman, N., and Vogel, H. J. (1993) *Biochemistry 32*, 646−656.
4. Shchyolkina, A. K., Lysov, Y. P., Il′ichova, I. A., Chernyi, A. A., Golova, Y. B., Ckernov, B. K., Gottikh, B. P., and Florentiev, V. L. (1989) *FEBS Lett. 244*, 39−42.
5. Rippe, K., Fritsch, V., Westhof, E., and Jovin, T. M. (1992) *EMBO J. 11*, 3777−3786.
6. Germann, M. W., Kalisch, B. W., Pon, R. T., and van de Sande, J. H. (1990) *Biochemistry 29*, 9426−9432.
7. Germann, M. W., Zhou, N., van de Sande, J. H., and Vogel, H. J. (1995) *Methods Enzymol. 261*, 207−225.
8. Borisova, O. F., Shchyolkina, A. K., Chernov, B. K., and Tchurikov, N. A. (1993) *FEBS Lett. 322*, 304−306.
9. Rippe, K., Ramsing, N. B., Klement, R., and Jovin, T. M. (1990) *J. Biol. Struct. Dyn. 7*, 1199−1209.
10. Rich, A., Davies, D. R., Crick, F. H., and Watson, J. D. (1961) *J. Mol. Biol. 3*, 71−86.
11. Langridge, R., and Rich, A. (1963) *Nature 198*, 725−728.
12. Brown, T., Hunter, W. N., Kneale, G., and Kennard, O. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2402−2406.
13. Sarma, M. H., Gupta, G., and Sarma, R. H. (1986) *FEBS Lett. 205*, 223−229.
14. Robinson, H., van der Marel, G. A., van Boom, J. H., and Wang, A. H. (1992) *Biochemistry 31*, 10510−10517.
15. Liu, K., Miles, H. T., Frazier, J., and Sasisekharan, V. (1993) *Biochemistry 32*, 11802−11809.
16. Manor, H., Sridhara, R. B., and Martin, R. G. (1988) *J. Mol. Evol. 27*, 96−101.
17. Lyamichev, V. J., Mirkin, S. M., and Frank-Kamenetskii, M. D. (1986) *J. Biol. Mol. Struct. Dyn. 3*, 667−669.
18. Htun, H., and Dahlberg, J. E. (1988) *Science 243*, 1571−1576.
19. De Jonckheere, J., Nelson, J. E., Ginsburg, K. A., Martin, L., and Krawetz, S. A. (1994) *Hum. Mol. Genet. 3*, 10.
20. Yagil, C. (1991) *CRC Crit. Rev. Biochem. Mol. Biol. 26*, 151−226.
21. Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., and Wohlrab, F. (1988) *FASEB. J. 2*, 2939−2949.
22. Noonberg, S. B., François, J.-C., Garestier, T., and Hélene, C. (1995) *Nucleic Acids Res. 23*, 1956−1963.
23. Casasnovas, J. M., Huertas, D., Ortiz-Lombardia, M., Kypr, J., and Azorin, F. (1993) *J. Mol. Biol. 233*, 671−681.
24. Dolinnaya, N. G., and Fresco, J. R. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9242−9246.
25. Dolinnaya, N. G., Ulku, A., and Fresco, J. R. (1997) *Nucleic Acids Res. 25*, 1100−1107.
26. Rajagopalan, P., and Feigon, J. (1989) *Biochemistry 28*, 7859−7870.
27. Huertas, D., Bellsolell, L., Casanovas, J. M., Coll, M., and Azorin, F. (1993) *EMBO J. 10*, 4029−4038.
28. Mukerji, I., Shiber, M. C., Fresco, J. R., and Spiro, T. G. (1996) *Nucleic Acids Res. 24*, 5013−5020.
29. Ramsing, N. B., and Jovin, T. M. (1988) *Nucleic Acids Res. 14*, 6659−6676.
30. Germann, M. W., Kalisch, B. W., and van de Sande, J. H. (1988) *Biochemistry 27*, 8302−8306.
31. Evertsz, E. M., Rippe, K., and Jovin, T. M. (1994) *Nucleic Acids Res. 22*, 3293−3303.

32. Germann, M. W., Vogel, H. J., Pon, R. T., and van de Sande, J. H. (1989) *Biochemistry 28*, 6220−6228.
33. Germann, M. W., Kalisch, B. W., and van de Sande, J. H. (1996) *J. Biol. Mol. Struct. Dyn. 13*, 953−962.
34. Germann, M. W. (1989) Ph.D. Thesis, The University of Calgary.
35. Roberts, E., Deeble, V. J., Woods, C. G. and Taylor, G. R. (1997) *Nucleic Acids Res. 15*, 3377−3378.
36. Feigon, J., Gilbert, D., Rajagopalan, P., Wang, E., van der Marel, G. A., and van Boom, J. H. (1990) In *Structure and Methods 3, DNA and RNA* (Sarma, R. H., and Sarma, M. H., Eds.) pp 207−224, Adenine Press, Guilderland, NY.
37. Jin, R., Gaffney, B. L., Wang, C., Jones, R. A., and Breslauer, K. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8832−8836.
38. Anato, V. P., Gray, D. M., and Ratliff, R. (1988) *Nucleic Acids Res. 16*, 719−738.
39. Germann, M. W., Kalisch, B. W., Quong, B. Q., and van de Sande, and J. H. (1995) *J. Biol. Struct. Dyn. 12*, a066.
40. Belotserkovskii, B. P., and Johnston, B. H. (1996) *Electrophoresis 17*, 1528−1534.
41. Belotserkovskii, B. P., and Johnston, B. H. (1997) *Biophys. J. 73*, 1288−1298.
42. Gogos, J. A., Karayigou, M., Abutatani, H., and Kafatos, F. C. (1990) *Nucleic Acids Res. 18*, 6807−6814.
43. Klysik, J., Rippe, K., and Jovin, T. M. (1990) *Biochemistry 29*, 9831−9839.
44. Sen, D., and Gilbert, W. (1988) *Nature*, *334*, 364−366.
45. Robinson, H., van Boom, J. H., and Wang, A. H. (1994) *J. Am. Chem. Soc. 116*, 1565−1566.
46. Robinson, H. van der Marel, G. A., van Boom, J. H., and Wang, A. H. (1992) *Biochemistry 31*, 10510−10517.

BI9726842