Structural Polymorphism of Homopurine DNA Sequences. d(GGA)$_n$ and d(GGGA)$_n$ Repeats Form Intramolecular Hairpins Stabilized by Different Base-Pairing Interactions†

Dori Huertas and Fernando Azorín*

Departament de Biologia Molecular i Cel.lular, Centre d'Investigació i Desenvolupament-CSIC, Jordi Girona Salgado 18-26, 08034 Barcelona, Spain

Received April 29, 1996; Revised Manuscript Received August 2, 1996⊗

ABSTRACT: DNA sequences containing homopurine d($G_{1-3}A)_n$ tracts are known to be capable of adopting non-B-DNA conformations. The structural polymorphism of these sequences is a direct consequence of the structural properties of the homopurine d($G_{1-3}A)_n$ tracts. Depending on the conditions, d(GA)$_n$ DNA sequences can form antiparallel- and parallel-stranded homoduplexes, multistranded complexes, and ordered single-stranded conformations. On the other hand, much less is known about the structural properties of d(GGA)$_n$ and d(GGGA)$_n$ sequences. In this paper, we show that d(GGA)$_n$ and d(GGGA)$_n$ repeats form antiparallel-stranded, intramolecular hairpins. Under physiological salt and pH conditions, the thermal stability of these hairpin forms is high, showing, at 50 mM NaCl, melting temperatures in the range of 40−50 °C. The base-pairing interactions involved in the formation of the d(GGA)$_n$ and d(GGGA)$_n$ hairpins are different. G·A pairs importantly contribute to the stability of the d(GGA)$_n$ hairpins. On the other hand, the d(GGGA)$_n$ hairpins are stabilized by the formation of G·G and A·A, but not G·A pairs. Homopurine d($G_{1-3}A)_n$ tracts are frequently found at genomic locations performing specialized chromosomal functions (i.e. telomeres, centromeres, and recombination "hot-spots"). The molecular interactions described here are relevant for the understanding of the peculiar structural and biological properties of DNA sequences containing homopurine tracts.

A distinctive trait of eukaryotic genomic DNA is its high content of repetitive DNA sequences. A class of repetitive DNA sequences, the microsatellites, is formed by the tandem repetition of very simple monomer units, from one to eight nucleotides long. Homopurine tracts of the type d($G_{1-3}A)_n$ constitute a particularly abundant type of microsatellite DNA sequences (Birnboim *et al*., 1979; Manor *et al*., 1988). Repetitions of this type are found dispersed along the eukaryotic genome, being frequently located at gene regulatory regions as well as in genomic locations which are "hot-spots" for genetic recombination [reviewed in Wells *et al*. (1988) and Palecek (1991)]. Closely related sequences are also found at specialized chromosomal regions, such as centromeres and telomeres. Many telomeric repeats contain homopurine tracts of this type spaced by short runs of pyrimidines [reviewed in Blackburn and Szostak (1984) and Zakian (1989)]. In a few cases, such as in the mould *Dictyostelium discoideum*, the telomeric repeats do not contain pyrimidines (Emery & Weiner, 1981). Homopurine tracts are also found in centromeric satellites of different origins, including mammals, birds, insects, plants, etc. A few examples include the dodeca [d(GTACGGGACCAG) repeats] and AAGAG satellites of *Drosophila*, which are present at the centromeres of chromosomes 2 and 3, respectively (Abad *et al*., 1992; Lohe *et al*., 1993), or the human 5 bp satellite 3 [d(TGGAA)$_n$ repeats] which, accounting for around 5% of the total human genome, is found at the centromere of all human chromosomes (Grady *et al*., 1992).

DNA sequences containing homopurine d($G_{1-3}A)_n$ tracts are capable of forming non-B-DNA structures. Telomeric DNAs are known to form intramolecular fold-back structures as well as DNA tetraplexes [reviewed in Sundquist (1991)]. The formation of intramolecular hairpins has also been reported for some centromeric satellites containing homopurine d($G_{1-3}A)_n$ tracts (Grady *et al*., 1992; Catasti *et al*., 1994; Chou *et al*., 1994; Ferrer *et al*., 1995). The ability of telomeric DNAs to form intramolecular hairpins and/or tetraplexes is likely to be functionally relevant. Nuclear proteins exist which specifically recognize telomeric DNA and facilitates the formation of tetrastranded DNA (Fang & Cech, 1993a,b).

Purine−purine interactions are the basis of the characteristic structural behavior of DNA sequences containing homopurine tracts (Sundquist, 1991; Catasti *et al.*, 1994; Chou *et al.*, 1994; Ferrer *et al.*, 1995). Actually, simple repeating homopurine d($G_{1-3}A)_n$ sequences are by themselves structurally polymorphic. Alternating d(GA)$_n$ sequences are the best studied of these sequences. It has been reported that, depending on the environmental conditions, single-stranded d(GA)$_n$ sequences can self-associate to form antiparallel-stranded homoduplexes (Casasnovas *et al*., 1993; Huertas *et al*., 1993), parallel-stranded homoduplexes (Rippe *et al*., 1992; Evertsz *et al*., 1994; Ortiz-Lombardía *et al*., 1995), and DNA tetraplexes (Lee *et al*., 1980; Antao *et al*., 1988; Lee, 1990). Finally, when contained into negatively supercoiled DNA, alternating d(GA·TC)$_n$ sequences have

† This work was financed by grants from the Spanish DGICYT (PB93-102) and the CEC (BIO2-CT94-3069 and CHRX-CT94-0047). D.H. acknowledges receipt of a doctoral fellowship from the Spanish DGICYT. This work was carried out in the framework of the Centre de Referència en Biotecnologia of the CIRIT of the Generalitat de Catalunya.

* Corresponding author.

⊗ Abstract published in *Advance ACS Abstracts,* September 1, 1996.

S0006-2960(96)01020-3 CCC: $12.00

Table 1: Oligonucleotides Used in These Experiments[a]

oligo(GGA)
$(GGA)_{10}G$
oligoH[$(G_2A)_{10}$]
GCGATCGCGCG**$(GGA)_{10}$**CGCGCGATCGC
oligoH[$(G_2A)_{10}G$]
GCGATCGCGCG**$(GGA)_{10}G$**CGCGCGATCGC
oligoH[$(G_2A)_{10}GG$]
GCGATCGCGCG**$(GGA)_{10}GG$**CGCGCGATCGC
oligo(GGGA)
$A(GGGA)_8$
oligoH[$(G_3A)_8$]
GCGATCGCGCG**$(GGGA)_8$**CGCGCGATCGC
oligoH[$A(G_3A)_8$]
GCGATCGCGCG**$A(GGGA)_8$**CGCGCGATCGC
oligo[$H(G_3A)_7GGG$]
GCGATCGCGCG**$(GGGA)_7GGG$**CGCGCGATCGC
oligo52C
AGGATCCCGGGTCGCGACAGCTGTTTTTACAGCTGTCGCGACCCGGGATCCT

[a] The nucleotide sequences of the oligonucleotides used in this paper are shown. The homopurine tracts are indicated in boldface. Underlined bases in oligo52C correspond to the four thymine residues describing the loop of the hairpin.

been shown to form a variety of different conformations which include pyr·pu·pyr[1] and pu·pu·pyr intramolecular triplexes, as well as pu·pu intramolecular hairpins [reviewed in Mirkin and Frank-Kamenetskii (1994), Frank-Kamenetskii and Mirkin (1995), and Bernués and Azorín (1995)]. Much less is known about the structural properties of $d(GGA)_n$ and $d(GGGA)_n$ sequences. CD experiments have suggested that long $d(GGA)_n$ polynucleotide chains can form DNA tetraplexes (Lee *et al.*, 1980; Lee, 1990), and the formation of intramolecular complexes was reported in the cases of the $d(AG_7AGAG_6AG_6)$, $d(AGG)_8$, and $d(AGGGAGG)_8$ oligonucleotides (Henderson *et al.*, 1987; Muraiso *et al.*, 1992). However, the molecular interactions which determine the formation of these structures were not characterized.

MATERIAL AND METHODS

Oligonucleotides. All oligonucleotides used in these experiments were synthesized in an Applied Biosystems automatic synthesizer and purified by polyacrylamide gel electrophoresis. When required, oligonucleotides were radioactively labeled with [γ-^{32}P]ATP and polynucleotide kinase.

Conditions of DEPC Modification. The patterns of modification with DEPC were obtained, at an oligonucleotide concentration of 10^{-4} μg/μL, either in citric phosphate buffer (pH 7.0) at an ionic strength equivalent to 100 mM Na^+ or in 10 mM Tris-HCl and 50 mM NaCl (pH 7.0). The reaction was carried out with 2 μL of DEPC (Sigma) in a final volume of 20 μL for 3 h at 4 °C or 2 h at 20 °C. After modification, samples were ethanol precipitated, cleaved with 1 M piperidine, and analyzed on 20% polyacrylamide−7 M urea sequencing gels. Autoradiographs were recorded in Hyperfilm (Amersham). Scans were obtained in a Molecular Dynamics laser densitometer.

Conditions of DMS Modification. The patterns of DMS modification were obtained in 10 mM Tris-HCl and 50 mM NaCl (pH 7.0) at an oligonucleotide concentration of 10^{-4} μg/μL. The reaction was performed in the presence of 4 μg of carrier DNA with 1 μL of DMS (Sigma) in a final volume of 200 μL for 15 min at 4 °C or 3 min at 20 °C. After modification, samples were analyzed as indicated above. Similar results were obtained when the reaction was performed in citric phosphate buffer (pH 7.0).

Polyacrylamide Gel Electrophoresis. The electrophoretic behavior of the oligonucleotides was analyzed in 12% (w/v) native polyacrylamide−TBE [89 mM Tris, 89 mM boric acid, and 1 mM EDTA (pH 8.3)] gels. When indicated, 50 mM NaCl was added to both the gel and the running buffer. Before electrophoresis, oligonucleotides were treated at 90 °C for 2 min and, then, incubated for 30 min under the same buffer and temperature conditions at which the electrophoresis was to be carried out. All the experiments were carried out at a DNA concentration of 10^{-4} μg/μL, except when indicated. Electrophoresis was performed at 6.5 V/cm until the dye xylene cyanol had migrated around 10−12 cm. To maintain the temperature constant, electrophoresis was performed in a buffer-jacketed apparatus connected to a thermostat bath. After electrophoresis, gels were dried and autoradiographs were recorded in Agfa Curix RP2 film at −80 °C.

UV Melting Experiments. UV melting experiments were carried out at a DNA concentration of 6×10^{-3} μg/μL in TBE (pH 8.3) buffer, with a final volume of 800 μL, either in the absence or in the presence of 50 mM NaCl. Optical measurements were performed in a Shimadzu UV2101PC spectrophotometer with a Shimadzu SPR-8 temperature control unit. Absorbencies at 260 nm were recorded between 20 and 85 °C at a temperature increase rate of 0.5 °C/min, and the melting temperatures (T_m) were determined from the midpoints of the temperature-induced transitions.

RESULTS

Repeated $d(GGA)_n$ and $d(GGGA)_n$ DNA Sequences Form Fold-Back Structures. To investigate the conformational characteristics of homopurine $d(GGA)_n$ and $d(GGGA)_n$ DNAs, we have analyzed the structural properties of oligonucleotides containing such repeated DNA sequences. Table 1 summarizes the oligonucleotides used in these studies. Figure 1 shows the gel electrophoretic analysis of oligo(GGA), which contains a 31 b long d(GGA) sequence, and of oligo(GGGA), which contains a 33 b long d(GGGA) sequence. In the case of oligo(GGA), two distinct electro-

[1] Abbreviations: b, base; bp, base pair; CD, circular dichroism; DEPC, diethyl pyrocarbonate; DMS, dimethyl sulfate; *M*, molecular mass; pu, purine residue; pyr, pyrimidine residue; UV, ultraviolet light.

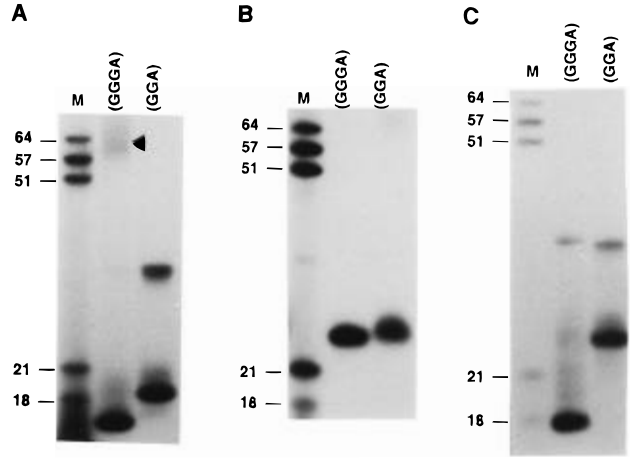

FIGURE 1: Electrophoretic behavior of oligo(GGA) and oligo(GGGA). Electrophoresis was performed at (A) 4 °C in 50 mM NaCl−TBE buffer, (B) 40 °C in 50 mM NaCl−TBE buffer, and (C) 4 °C in TBE buffer. Lanes M correspond to molecular mass markers, whose size in bp is indicated. The arrowhead in panel A indicates the position corresponding to the species with high *M* (see text).

phoretic species can be detected when the electrophoresis is carried out at 4 °C in the presence of 50 mM NaCl (Figure 1A). The slow-migrating species likely corresponds to the formation of an intermolecular homoduplex since it shows an apparent molecular mass (*M*) of around 36 bp, very close to the value expected for a bimolecular complex. Consistent with this interpretation, its relative abundance depends strongly on DNA concentration (not shown). At the DNA concentration shown in Figure 1A (2.5×10^{-5} μg/μL), it accounts for around 20−25% of the total population of molecules, but at higher DNA concentrations, it becomes predominant (i.e. at 2.5×10^{-4} μg/μL it accounts for around 70−75%) while at lower DNA concentrations (i.e. 5×10^{-6} μg/μL) it is undetectable. On the other hand, the fast-migrating species, which is the majority under the experimental conditions of Figure 1A, shows an apparent *M* very close to 18 bp, strongly indicating that it corresponds to an intramolecular fold-back form. Also in agreement with this interpretation, the fast-migrating species is detected over a broad range of DNA concentrations, from 5×10^{-6} to 2.5×10^{-4} μg/μL, being the only detectable species at 5×10^{-6} μg/μL (not shown). These two species are no longer observed when the electrophoresis is carried out at 40 °C. At this temperature, only one species is observed (Figure 1B), which corresponds to the single-stranded form of the oligonucleotide since it shows an apparent *M* of around 26 bp, which is very close to what is expected for the single-stranded form of oligo(GGA) (Maniatis *et al*., 1975).

Under the same experimental conditions, oligo(GGGA) exists principally as a fast-migrating species which shows an apparent *M* lower than 18 bp (Figure 1A), indicating that it corresponds to an intramolecular fold-back form. This fold-back form is also melted at 40 °C (Figure 1B). At this temperature, oligo(GGGA) shows an electrophoretic mobility slightly slower than that corresponding to oligo(GGA) (Figure 1B), as corresponds to its higher *M*. The formation of complexes of very low electrophoretic migration is also detected in Figure 1A. These forms, which must correspond to multistranded species, account for a very minor proportion of the total population. They show an apparent *M* of around 64 bp, suggesting that they are likely to correspond to

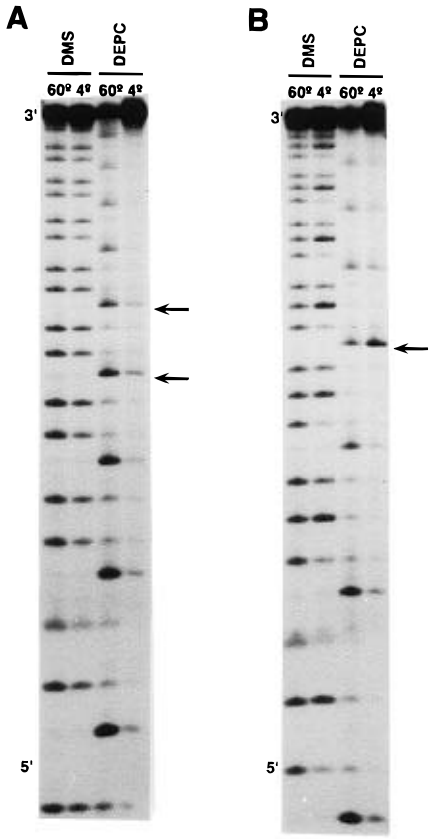

FIGURE 2: DEPC and DMS modification patterns of (A) oligo(GGA) and (B) oligo(GGGA). The patterns of modification were obtained at the temperature indicated. Arrows indicate adenines reactive with DEPC at low temperatures (see text). The 5′ to 3′ direction is also indicated.

tetrastranded complexes. At the DNA concentration of Figure 1A (2.5×10^{-5} μg/μL), there are no signs of the formation of the intermolecular homoduplex form which becomes detectable only at higher DNA concentrations (not shown). Similarly, increasing the DNA concentration also results in an increased proportion of multistranded complexes (not shown). However, in the case of oligo(GGGA), the intramolecular fold-back form was always the more abundant one, even at the highest DNA concentration tested (2.5×10^{-4} μg/μL).

The gel electrophoretic analyses described above indicate that homopurine d(GGA)$_n$ and d(GGGA)$_n$ DNA sequences have a high propensity to form intramolecular homoduplexes. The patterns of DEPC modification of oligo(GGA) and oligo(GGGA) are consistent with this interpretation (Figure 2). DEPC detects the presence of unpaired purines, principally of adenines [reviewed in Palecek (1991) and Lilley (1992)]. In the case of oligo(GGGA), at 4 °C, DEPC reactivity is constrained to the most central adenine residue (Figure 2B), indicating that, at this temperature, oligo(GGGA) exists principally as an intramolecular hairpin with the central adenine residue being located at the loop. The patterns of DMS modification of oligo(GGGA) are also in agreement with these results. At 4 °C, oligo(GGGA) shows a characteristic pattern of DMS modification in which the guanine residues of each GGG repeat show a rather unequal reactivity (see below). These patterns of reactivity are lost at high temperatures, when the oligonucleotide is single-stranded. Under these conditions, all guanine residues are equally reactive with DMS and all adenines are reactive with DEPC

(Figure 2B).

In the case of oligo(GGA), the patterns of DEPC modification also show clear signs of the formation of a hairpin form. At 4 °C, the central adenine residues show a significant DEPC reactivity consistent with the formation of an intramolecular hairpin (Figure 2A). However, in this case, DEPC reactivity is not constrained to a single adenine residue and the central residues appear to be less reactive than in the case of oligo(GGGA). This is likely to reflect the circumstance that, in the case of oligo(GGA), a significant proportion of the molecules form intermolecular homoduplexes which coexist with the intramolecular forms (Figure 1A). Under the same experimental conditions, oligo(GGGA) exists principally as intramolecular complexes (Figure 1A). The patterns of DEPC modification might also reflect the lower stability of the hairpin form of oligo(GGA) (see below). As in oligo(GGGA), all the adenines of oligo(GGA) become reactive with DEPC at higher temperatures (Figure 2A).

Antiparallel-Stranded d(GGA)$_n$ Homoduplexes Are Stabilized by the Formation of G·A and G·G Base Pairs. From the results shown above, we conclude that homopurine d(GGA)$_n$ and d(GGGA)$_n$ DNA sequences have a high tendency to form antiparallel-stranded homoduplexes. What are the molecular interactions which determine the stability of these structural forms? As shown in Figure 1, both types of homopurine sequences can also form intermolecular complexes. This is particularly evident in the case of oligo(GGA) that, under most experimental conditions, forms at least two different types of complexes. The coexistence of different structural forms constitutes a very serious limitation for the analysis of the molecular interactions involved in the formation of antiparallel-stranded d(GGA)$_n$ DNA. To overcome these limitations, we studied the structural behavior of oligonucleotides containing a stretch of repeated d(GGA)$_n$ DNA flanked by palindromic sequences (Table 1). These oligonucleotides have a very high tendency to form intramolecular hairpins. The extent to which the homopurine sequence contributes to the formation of such hairpins, as well as the base-pairing interactions involved in their formation, can then be determined from the patterns of chemical modification obtained with DEPC and/or DMS. This experimental approach was useful in the study of the structural properties of d(GA)$_n$ DNA sequences (Casasnovas *et al.*, 1993; Huertas *et al.*, 1993).

Figure 3A shows the patterns of DEPC modification corresponding to oligoH[(G$_2$A)$_{10}$] which contains a 30 b long d(GGA)$_{10}$ sequence (Table 1). When the reaction is carried out at 4 °C, only the residues located at the center of the homopurine sequence become reactive with DEPC, indicating that the oligonucleotide folds into an intramolecular hairpin which includes the homopurine sequence in the stem. This pattern of modification is very similar to that observed in the case of oligo(GGA), which contains no complementary bases flanking the homopurine d(GGA) sequence (Figure 2A). In oligoH[(G$_2$A)$_{10}$], hyperreactivity to DEPC extends from A15 to A18, indicating that these residues are in the loop of the hairpin (Figure 4A). Residue G14 also occurs at the loop since it shows a DEPC reactivity similar to those of residues G16 and G17, higher than those of the rest of guanines of the homopurine sequence (Figures 3A and 4A). The same central residues are also hyperreactive to DEPC at 20 °C, indicating that the hairpin is stable at this

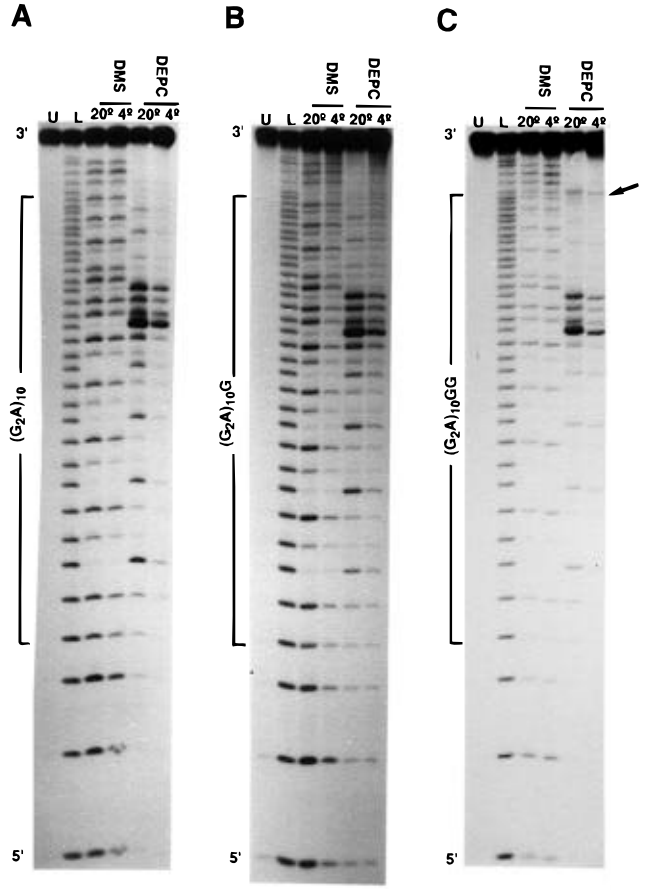

FIGURE 3: DEPC and DMS modification patterns of (A) oligoH[(G$_2$A)$_{10}$], (B) oligoH[(G$_2$A)$_{10}$G], and (C) oligoH[(G$_2$A)$_{10}$GG]. As indicated, the patterns of modification were obtained at 4 and 20 °C. Lanes L correspond to G+A sequencing ladders. Lanes U are controls, corresponding to the piperidine products obtained from the untreated oligonucleotides. The regions corresponding to the homopurine sequences are indicated. The 5′ to 3′ direction is also indicated. The arrow in panel C indicates the position corresponding to residue G32 (see text).

temperature. The rest of the adenines of the homopurine sequence become only moderately reactive with DEPC at 20 °C (Figure 3A).

The patterns of DMS modification of oligoH[(G$_2$A)$_{10}$] show some interesting features (Figure 3A). The 3′ guanine residue of each d(GGA) repeat shows a DMS reactivity very similar to that observed for the guanine residues occurring outside of the homopurine sequence. On the contrary, the 5′ guanine residue of each repeat shows a much lower reactivity (Figure 4A). Oligo(GGA) does not show the same alternation of reactive and unreactive guanines (Figure 2A), probably as a consequence of the coexistence of several structural forms in oligo(GGA) (Figure 1A). This characteristic pattern of DMS reactivity is observed in all the repeats of oligoH[(G$_2$A)$_{10}$], except those occurring at the loop region. In these cases, the reactivities of the two guanines are equally high as expected from their unpaired character. Notably, a similar situation is observed at the first d(GGA) repeat. Also in this case, the DMS reactivity of the 5′ guanine (residue G1) is as high as that of the 3′ guanine (residue G2), suggesting that residue G1 is most likely unpaired. According to these results, the hairpin form of oligoH[(G$_2$A)$_{10}$] would be stabilized by the formation of G·A and G·G base pairs (Figure 4B).

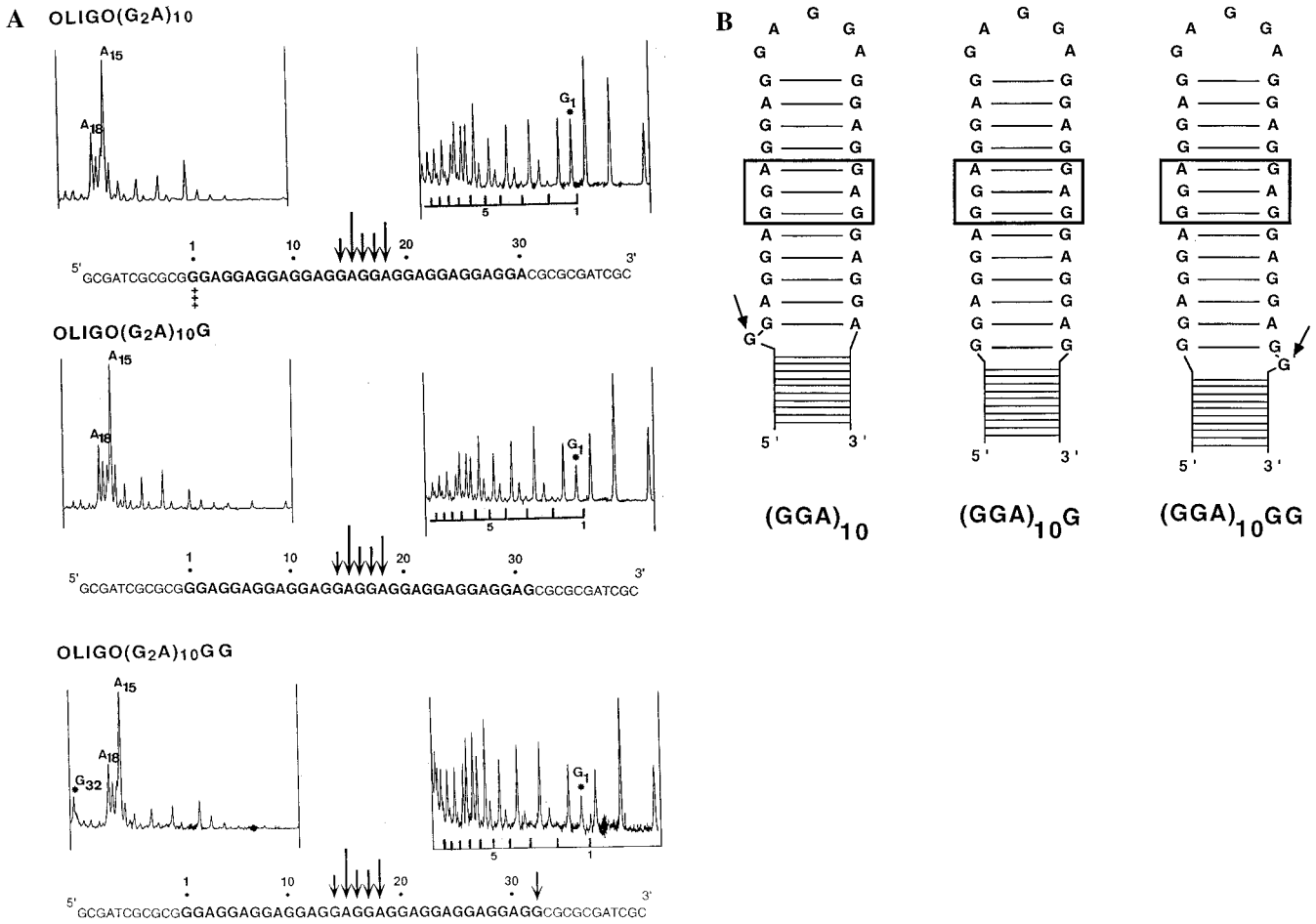

FIGURE 4: Quantitative analysis of the results shown in Figure 3. (A) Densitometer tracings corresponding to the patterns of DEPC (left) and DMS modification (right) obtained at 4 °C for oligoH[$(G_2A)_{10}$] (top), oligoH[$(G_2A)_{10}$G] (middle), and oligoH[$(G_2A)_{10}$GG] (bottom). The position of selected residues is indicated. The 5′ to 3′ direction is always right to left. For each oligonucleotide, a summary of the results is presented underneath the corresponding densitometer tracings. Bases hyperreactive to DEPC are indicated by arrows. Guanine residues showing anomalous DMS reactivity are indicated by +++. (B) Schematic representations of the conformations compatible with the patterns of chemical modification of oligoH[$(G_2A)_{10}$] (left), oligoH[$(G_2A)_{10}$G] (middle), and oligoH[$(G_2A)_{10}$GG] (right). Arrows indicated the position of unpaired residues located at the junction region between the homopurine sequence and the flanking B-DNA.

To further corroborate this hypothesis, the patterns of DEPC and DMS modification were obtained for oligoH[$(G_2A)_{10}$G] and oligoH[$(G_2A)_{10}$GG] which incorporate one or two additional guanine residues at the 3′ end of the d$(GGA)_{10}$ sequence (Table 1). Their patterns of chemical reactivity are very similar to those corresponding to oligoH[$(G_2A)_{10}$] (Figure 3B,C). The central residues (from G14 to A18) are hyperreactive to DEPC, and the patterns of DMS modification show the characteristic alternation of unreactive 5′ guanines and reactive 3′ guanines (Figure 4A). These results indicate that oligoH[$(G_2A)_{10}$G] and oligoH[$(G_2A)_{10}$GG] also form intramolecular hairpins which are stabilized through the same type of base-pairing interactions observed in the case of oligoH[$(G_2A)_{10}$]. However, the patterns of DMS modification of oligoH[$(G_2A)_{10}$G] and oligoH[$(G_2A)_{10}$GG] show some important differences with respect to oligoH[$(G_2A)_{10}$]. Residue G1, which was reactive with DMS in oligoH[$(G_2A)_{10}$], is significantly less reactive in oligoH[$(G_2A)_{10}$G] (Figures 3B and 4A), strongly suggesting that it becomes paired to the last guanine (G31) of the homopurine sequence which is not reactive with either DMS or DEPC (Figures 3B and 4A). Similarly, in oligoH[$(G_2A)_{10}$GG], the DMS reactivity of residue G1 is also low (Figures 3C and 4A) but the last guanine (G32) of the homopurine sequence is significantly reactive with DEPC, indicating that, in this case, residues G1 and G31 are paired but residue G32 is not. These results provide strong evidence in favor of the models shown in Figure 4B. OligoH[$(G_2A)_{10}$G] does not contain any unpaired bases at the junction between the homopurine sequence and the flanking B-DNA. On the other hand, oligoH[$(G_2A)_{10}$] and oligoH[$(G_2A)_{10}$GG] contain an unpaired guanine at either the 5′ end or the 3′ end of the homopurine sequence, respectively.

The electrophoretic behavior of oligoH[$(G_2A)_{10}$], oligoH[$(G_2A)_{10}$G], and oligoH[$(G_2A)_{10}$GG] is consistent with this interpretation (Figure 5). It is known that DNA molecules containing unpaired bases or bulges migrate slower than those molecules which are perfectly paired (Bhattacharyya & Lilley, 1989; Rice & Chrothers, 1989). The reason for the retarded migration of bulged DNAs must be found in the curvature that the unpaired residues introduce in the molecule. As shown in Figure 5A, when the electrophoresis is carried out at 4 °C, oligoH[$(G_2A)_{10}$G] shows a higher electrophoretic mobility than either oligoH[$(G_2A)_{10}$] or oligoH[$(G_2A)_{10}$GG]. At this temperature, the apparent *M* values of oligoH[$(G_2A)_{10}$G], oligoH[$(G_2A)_{10}$], and oligoH[$(G_2A)_{10}$GG] are 28.5, 29.5, and 30 bp, respectively. These differences in electrophoretic migration cannot be interpreted in terms of the actual differences in the *M* values of the oligonucleotides. In fact, despite its higher

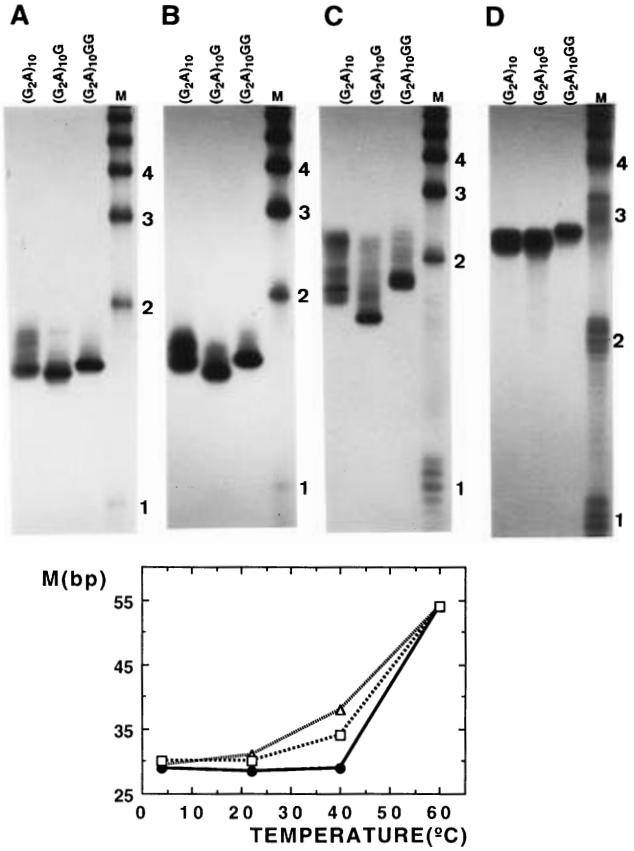

FIGURE 5: Electrophoretic behavior of oligoH[$(G_2A)_{10}$], oligoH-[$(G_2A)_{10}G$], and oligoH[$(G_2A)_{10}GG$]. Electrophoresis was performed at 4 °C (A), 20 °C (B), 40 °C (C), and 60 °C (D). Lanes M correspond to oligomers of an 18-mer oligonucleotide used as molecular mass markers. Numbers indicate the position corresponding to the monomer, dimer, trimer, and tetramer. The monomer melts at 40 °C, and at 60 °C, the dimer and trimer also melted. Quantitative analyses of the results are shown at the bottom. The apparent molecular mass (*M*) is presented as a function of the temperature at which the electrophoresis was performed: (△) oligoH[$(G_2A)_{10}$], (●) oligoH[$(G_2A)_{10}G$], and (□) oligoH[$(G_2A)_{10}$-GG]. The values of *M* were determined by interpolation or, when the corresponding markers were melted, from the distance to the first unmelted marker.

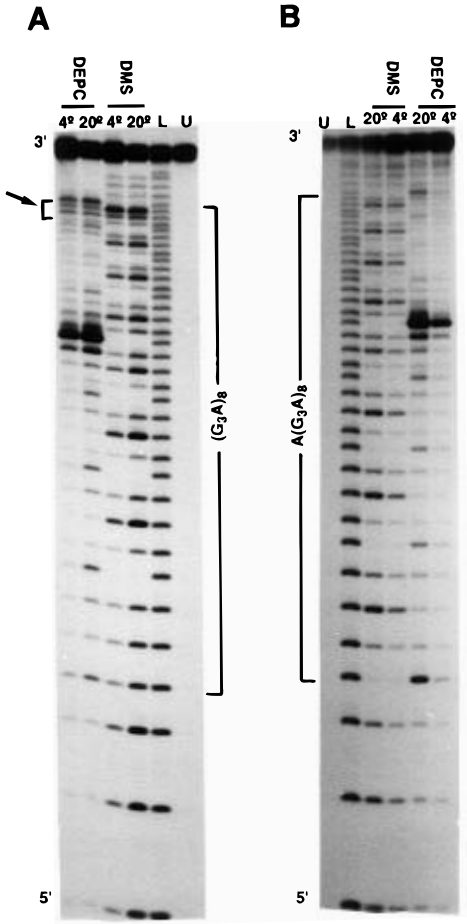

FIGURE 6: DEPC and DMS modification patterns of (A) oligoH-[$(G_3A)_8$] and (B) oligoH[$A(G_3A)_8$]. As indicated, the patterns of modification were obtained at 4 and 20 °C. Lanes L correspond to G+A sequencing ladders. Lanes U are controls, corresponding to the piperidine products obtained from the untreated oligonucleotides. The regions corresponding to the homopurine sequences are indicated. The 5′ to 3′ direction is also indicated. The arrow in panel A indicates the position corresponding to the 3′ region showing reactivity with DEPC (see text).

M, oligoH[$(G_2A)_{10}G$] migrates faster than oligoH[$(G_2A)_{10}$]. Therefore, the slower electrophoretic mobilities of oligoH-[$(G_2A)_{10}$] and oligoH[$(G_2A)_{10}GG$] are consistent with the presence of unpaired bases at their junction regions. OligoH-[$(G_2A)_{10}GG$] show an slightly lower electrophoretic migration than oligoH[$(G_2A)_{10}$], probably reflecting its higher *M*. Similar differences in electrophoretic mobility are detected at 20 °C (Figure 5B). Increasing further the temperature results in the melting of the intramolecular hairpins (Figure 5C,D). Partial melting of oligoH[$(G_2A)_{10}$] and oligoH[$(G_2A)_{10}$-GG] is detected at 40 °C as reflected by the formation of electrophoretic species with a slower mobility. Complete melting is observed only at 60 °C. On the other hand, oligoH[$(G_2A)_{10}G$] shows the same electrophoretic mobility at 40 °C as at 4 or 20 °C, indicating that its thermal stability is higher as was corroborated by UV melting experiments (not shown).

Antiparallel-Stranded d(GGGA)$_n$ Homoduplexes Are Stabilized by the Formation of G·G and A·A Pairs, but Not of G·A Pairs. A similar approach was used to investigate the molecular interactions determining the stability of antiparallel-stranded d(GGGA)$_n$ DNA. The oligonucleotides used in these experiments are also described in Table 1.

The patterns of chemical modification of oligoH[$(G_3A)_8$], carrying a 32 b long d$(GGGA)_8$ sequence, are very similar to those observed in oligo(GGGA) (compare Figures 6A and 2B). The pattern of DEPC modification of oligoH[$(G_3A)_8$] reveals unequivocally the formation of a hairpin. Most of the adenine residues of the d$(GGGA)_8$ sequence are not reactive with DEPC, but A16, which is located at the center of the repeated sequence, is highly reactive with DEPC. Residues G15 and G17 are also reactive (Figure 7A), indicating that they are also located at the loop of the hairpin. Guanines of each GGG trinucleotide repeat show a characteristic unequal DMS-reactivity (Figure 6A), the central guanine being highly reactive while the two flanking ones are significantly protected from DMS methylation, the 5′ guanine residue showing the lowest DMS reactivity (Figure 7A). This pattern of DMS modification is observed in all the GGG trinucleotide repeats except in those located either at the loop of the hairpin or at the junction between the homopurine sequence and the flanking B-DNA (Figures 6A and 7A). This junction region appears to be largely distorted. On one hand, the patterns of DMS reactivity of the first and last GGG trinucleotide repeats are clearly different from the

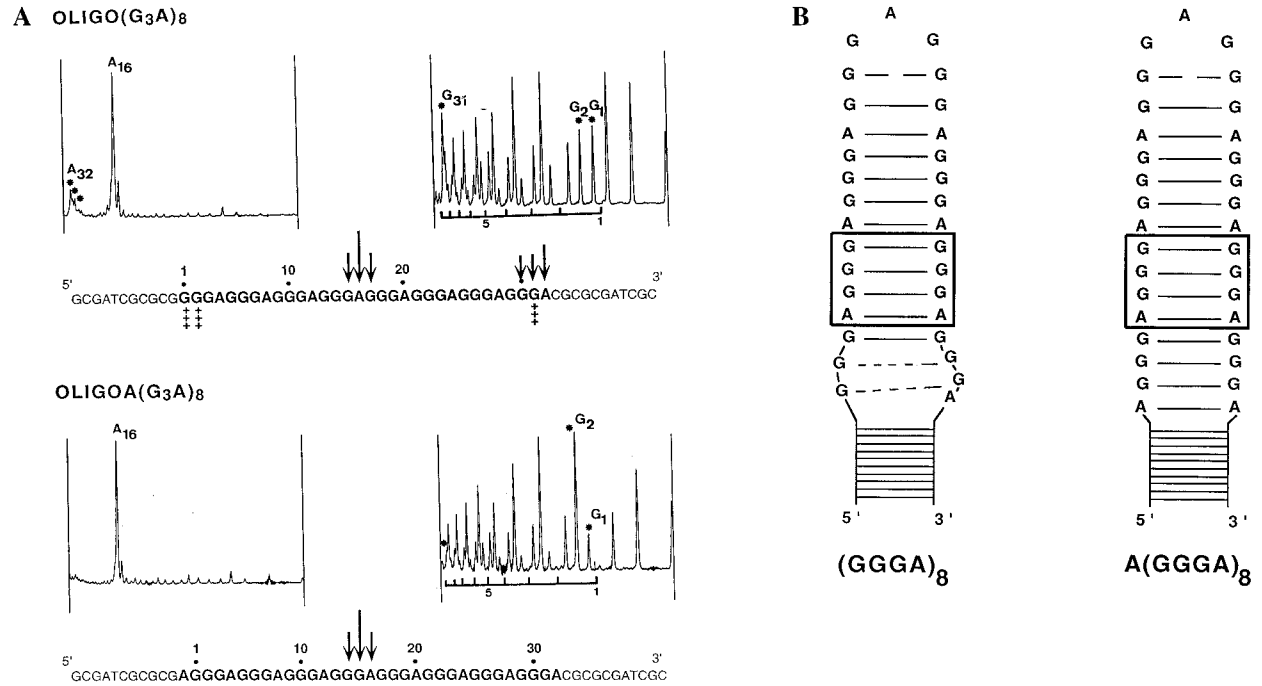

FIGURE 7: Quantitative analysis of the results shown in Figure 6. (A) Densitometer tracings corresponding to the patterns of DEPC (left) and DMS modification (right) obtained at 4 °C for oligoH[(G$_3$A)$_8$] (top) and oligoH[A(G$_3$A)$_8$] (bottom). The position of selected residues is indicated. The 5′ to 3′ direction is always right to left. For each oligonucleotide, a summary of the results is presented underneath the corresponding densitometer tracings. Bases hyperreactive to DEPC are indicated by arrows. Guanine residues showing anomalous DMS reactivity are indicated by +++. (B) Schematic representations of the conformations compatible with the patterns of chemical modification of oligoH[(G$_3$A)$_8$] (left) and oligoH[A(G$_3$A)$_8$] (right).

rest (Figure 7A). The first two guanines of the homopurine sequence show a rather similar DMS reactivity, and the 3′ guanine of the last GGG repeat is highly reactive with DMS. Second, the last three residues of the d(GGGA)$_8$ sequence, which includes A32 but also G31 and G30, are significantly reactive with DEPC (Figures 6A and 7A). These results indicate that, in oligoH[(G$_3$A)$_8$], the structural distortion of the junction involves, at least, the first two and the last three residues of the d(GGGA)$_8$ sequence (Figure 7B).

Interestingly, the junction region of oligoH[A(G$_3$A)$_8$], which contains an additional adenine at the beginning of the repeated sequence (Table 1), does not show any anomalous reactivity with either DEPC or DMS (Figures 6B and 7A). The first and last GGG trinucleotide repeats show the characteristic unequal DMS reactivity, and at 4 °C, DEPC reactivity is constrained to the central residues. These results indicate that, in contrast to what is observed in the case of oligoH[(G$_3$A)$_8$], the junction of oligoH[A(G$_3$A)$_8$] is not structurally distorted containing no unpaired residues (Figure 7B). At 20 °C, the first and last adenine residue of the homopurine sequence become moderately reactive with DEPC (Figure 6B).

The electrophoretic behavior of oligoH[A(G$_3$A)$_8$] is consistent with this interpretation (Figure 8). Despite its higher *M*, oligoH[A(G$_3$A)$_8$] shows, at 4 °C, a faster electrophoretic migration than oligoH[(G$_3$A)$_8$] which is consistent with the presence of unpaired residues in the latter but not in the former. A similar difference in electrophoretic migration is observed when the electrophoresis is performed at 20 and 40 °C (Figure 8B,C). At 60 °C, the oligonucleotides are melted (Figure 8D).

Therefore, as judged by chemical modification and gel electrophoretic experiments, the junction region of oligoH-[A(G$_3$A)$_8$] does not contain unpaired bases. These results indicate that the hairpin form of oligoH[A(G$_3$A)$_8$] is stabilized through the formation of G·G and A·A pairs, but not of G·A pairs (Figure 7B). Further evidence in favor of this interpretation comes from the electrophoretic behavior of oligoH[(G$_3$A)$_7$GGG] (Figure 8). This oligonucleotide, which starts and ends with a GGG trinucleotide (Table 1), has an electrophoretic migration similar to that of oligoH[A(G$_3$A)$_8$], indicating that its junction region does not contain either unpaired bases.

Stability of Antiparallel-Stranded d(GGA)$_n$ and d(GGGA)$_n$ DNAs. Under pH and ionic conditions close to physiological, the stability of antiparallel-stranded d(GGA)$_n$ and d(GGGA)$_n$ hairpins is high. In the presence of 50 mM NaCl, oligo(GGGA) shows a melting temperature (T_m) of 48 °C (Table 2). As discussed before, under these experimental conditions, oligo(GGGA) can form both intramolecular and intermolecular complexes. At the DNA concentration at which the UV melting experiments were performed (6×10^{-3} μg/μL), the intermolecular forms account for a significant proportion of the molecules, around 50% (not shown). Under these conditions, a single structural transition is detected, showing a relatively low cooperativity. Therefore, the melting temperature of oligo(GGGA) must correspond to the weighted average of the melting temperatures of the intramolecular and the intermolecular forms being, in this sense, only an indirect estimate of the actual thermal stability of antiparallel-stranded d(GGGA)$_n$ DNA. As shown above (Figure 6B), oligoH[A(G$_3$A)$_8$] exists exclusively as a hairpin containing around 15 bp of antiparallel-stranded d(GGGA) DNA together with an 11 bp long B-DNA duplex. As expected, oligoH[A(G$_3$A)$_8$] melts at a higher temperature (T_m = 63 °C) than oligo(GGGA) which is, however, lower than that corresponding to oligo52C (Table 2), a control hairpin of the same length as

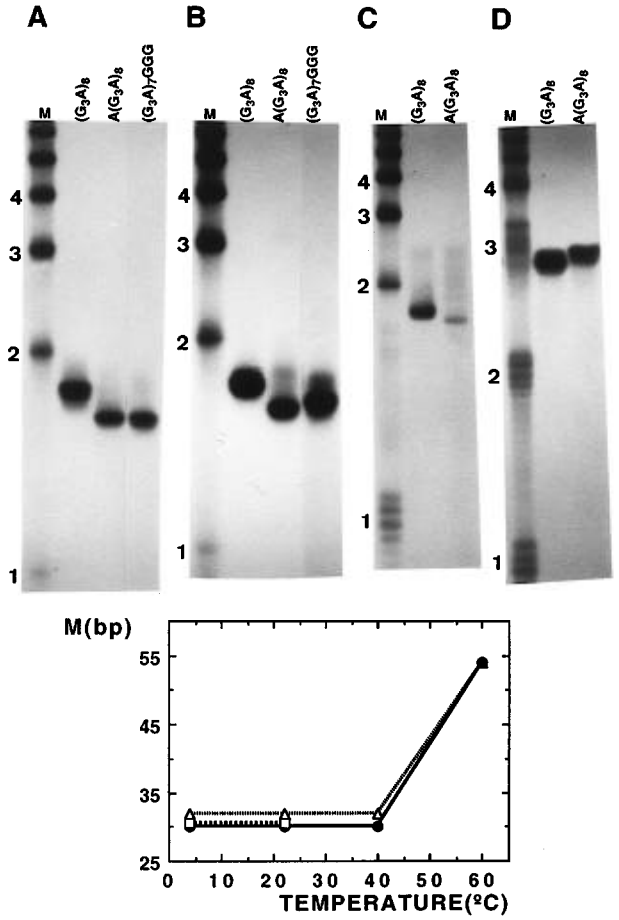

FIGURE 8: Electrophoretic behavior of oligoH[$(G_3A)_8$], oligoH-[$A(G_3A)_8$], and oligoH[$(G_3A)_7GGG$]. Electrophoresis was performed at 4 °C (A), 20 °C (B), 40 °C (C), and 60 °C (D). Lanes M correspond to oligomers of an 18-mer oligonucleotide used as molecular mass markers. Numbers indicate the position corresponding to the monomer, dimer, trimer, tetramer. The monomer melts at 40 °C, and at 60 °C, the dimer and trimer are also melted. Quantitative analyses of the results are shown at the bottom. The apparent molecular mass (*M*) is presented as a function of the temperature at which the electrophoresis was performed: (△) oligoH[$(G_3A)_8$], (●) oligoH[$A(G_3A)_8$], and (□) oligoH[$(G_3A)_7$-GGG]. The values of *M* were determined as indicated in Figure 5.

Table 2: UV Melting Experiments

	buffer	% hypercromicity	T_m (°C)
oligo(GGA)	50 mM NaCl−TBE	21	41
	TBE	12	−
oligoH[$(G_2A)_{10}G$]	50 mM NaCl−TBE	22	62
oligo(GGGA)	50 mM NaCl−TBE	22	48
	TBE	20	38
oligoH[$A(G_3A)_8$]	50 mM NaCl−TBE	22	63
oligo52C	50 mM NaCl−TBE	28	77
oligoH[$(GA)_{15}$][a]	50 mM NaCl−TBE	21	55
E. coli DNA[b]	50 mM NaCl−TBE	35	85

[a] OligoH[$(GA)_{15}$] contains a 30 b long homopurine $d(GA)_{15}$ sequence flanked also by complementary bases as in oligoH[$(G_2A)_{10}G$] and oligoH[$A(G_3A)_8$]. [b] The values corresponding to the melting of sheared *E .coli* DNA are presented for comparison.

oligoH[$A(G_3A)_8$] but formed exclusively by a B-DNA duplex (Table 1). Actually, the T_m of oligoH[$A(G_3A)_8$] corresponds roughly to the weighted average of the T_m values of oligo52C and of oligo(GGGA), suggesting that the actual T_m of the antiparallel-stranded $d(GGGA)_n$ homoduplex is close to the T_m of oligo(GGGA). As shown in Figure 1B, oligo(GGGA) behaves electrophoretically as single-stranded at 40 °C,

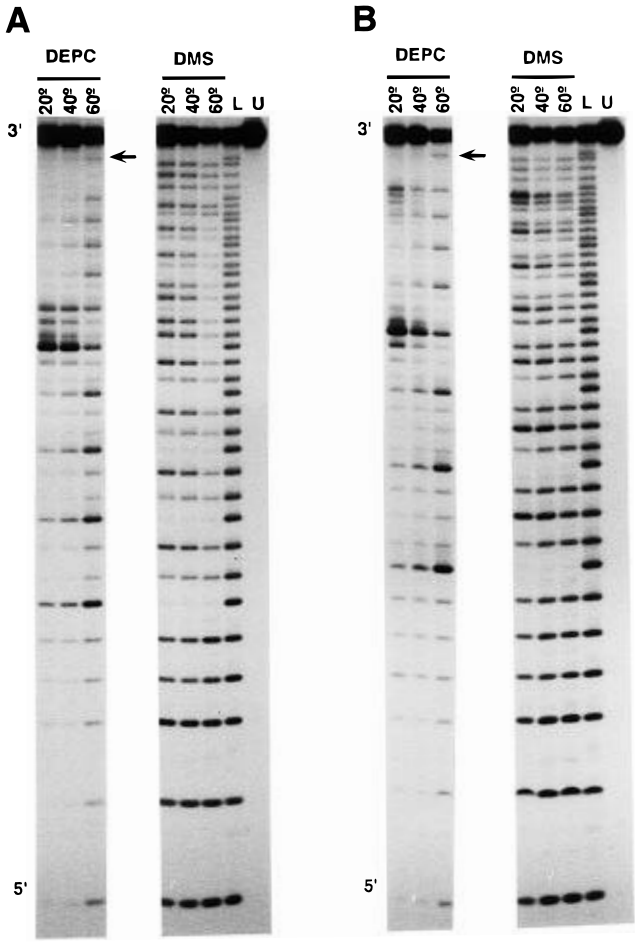

FIGURE 9: Patterns of DEPC and DMS modification of oligoH-[$(G_2A)_{10}$] (A) and oligoH[$(G_3A)_8$] (B), obtained at increasing temperatures. Lanes L correspond to G+A sequencing ladders. Lanes U are controls, corresponding to the piperidine products obtained from the untreated oligonucleotides. The 5′ to 3′ direction is indicated. The arrows indicate the position corresponding to the adenine residue located at the flanking B-DNA region which becomes reactive with DEPC at high temperatures (see text).

though, as determined by UV melting, its T_m is 48 °C. A similar discrepancy between the UV melting and gel electrophoresis results was also observed for oligo(GGA) (Figure 1B and Table 2). In general, all the oligonucleotides studied here showed a lower thermal stability under the electrophoretic conditions, likely as a consequence of changes in the experimental conditions (pH, ionic strength, etc.) occurring during electrophoresis. In good agreement with the UV melting results, the patterns of DEPC and DMS modification characteristic of the antiparallel-stranded $d(GGGA)_n$ homoduplex are still detected when the modification is performed at 40 °C (Figure 9B), indicating that this structural form is stable at this temperature. On the other hand, at 60 °C, all the guanines of the GGGA sequence are equally reactive with DMS and all the adenines are reactive with DEPC (Figure 9B). Notice that, at 60 °C, an adenine residue occurring at the B-DNA region flanking the GGGA sequence, which is totally unreactive at lower temperatures, becomes significantly reactive with DEPC (Figure 9B).

Antiparallel-stranded $d(GGA)_n$ DNA appears to be, under the same experimental conditions, less stable than antiparallel-stranded $d(GGGA)_n$ DNA. On one hand, at 50 mM NaCl, the T_m of oligo(GGA) is significantly lower than that of oligo(GGGA) (Table 2). Furthermore, in the absence of

NaCl, oligo(GGA) is not significantly structured while oligo(GGGA) is. In the absence of NaCl, the electrophoretic behavior of oligo(GGA) does not show any signs of becoming structured even at 4 °C (Figure 1C). Under these conditions, oligo(GGA) shows an apparent *M* of 26 bp which corresponds to the single-stranded form of the oligonucleotide (Figure 1B). In good agreement with these results, oligo(GGA) does not show any cooperative transition when the melting is performed in the absence of NaCl (Table 2). On the other hand, oligo(GGGA) is also structured in the absence of NaCl. Under these conditions, it melts at $T_m = 38$ °C (Table 2) and shows an electrophoretic apparent *M* close to 18 bp (Figure 1C), which is totally consistent with the formation of an intramolecular hairpin form. From the results shown in Table 2 and the patterns of DEPC and DMS modification of oligoH[$(G_2A)_{10}$] obtained at increasing temperature (Figure 9A), we estimated that, in the presence of 50 mM NaCl, the melting temperatures of antiparallel-stranded d(GGA)$_n$ DNA is around 40 °C.

DISCUSSION

As reported in this paper, homopurine d(GGA)$_n$ and d(GGGA)$_n$ sequences have a high tendency to form antiparallel-stranded homoduplexes. We have also detected the formation of intermolecular duplexes and, in the case of the d(GGGA)$_n$ sequences, of multistranded complexes. However, under the conditions explored in this paper, the intramolecular forms were always the majority. The proportion of molecules forming intermolecular duplexes is higher for the d(GGA)$_n$ sequences than for the d(GGGA)$_n$ sequences. Opposite to that of the intramolecular forms, which must necessarily be antiparallel-stranded, the strand orientation of the intermolecular duplexes could also be parallel. Interestingly, it has been recently reported that d(GGA)$_n$ sequences can also form parallel-stranded homoduplexes (Suda *et al*., 1995), providing a possible explanation for their higher tendency to form intermolecular duplexes. The stability of the antiparallel-stranded intramolecular forms of oligo(GGGA) is high, melting at around 48 °C in the presence of 50 mM NaCl. Formation of this structural form is also detected at low ionic strengths (i.e. in TBE buffer). On the other hand, oligo(GGA) is basically not structured at low ionic strengths. Others have also reported that a d(AGG)$_8$ oligonucleotide forms intramolecular complexes only at an ionic strength of 50 mM NaCl or higher (Muraiso *et al*., 1992). In the presence of 50 mM NaCl, oligo(GGA) melts at around 40 °C, confirming the lower stability of antiparallel-stranded d(GGA)$_n$ DNA. However, as shown in Table 2, the melting temperature of antiparallel-stranded d(GGA)$_n$ DNA is significantly higher than that of antiparallel-stranded d(GA)$_n$ DNA. The lower stability of antiparallel-stranded d(GGA)$_n$ is likely to account for the slower electrophoretic migration of oligo(GGA) when compared to that of oligo(GGGA) (Figure 1A).

Antiparallel-stranded d(GGA)$_n$ homoduplexes are stabilized by the formation of G·A and G·G pairs. This conclusion arises from the fact that, as judged by gel electrophoresis and chemical reactivity, oligoH[$(G_2A)_{10}$G] does not contain any unpaired bases at the junction between the homopurine sequence and the flanking B-DNA while both oligoH[$(G_2A)_{10}$] and oligoH[$(G_2A)_{10}$GG] contain an unpaired guanine residue at either the 5′ end or the 3′ end of the homopurine sequence, forming a bulge at the junction region (Figure 4B). Bulged DNAs are known to have a retarded electrophoretic migration (Bhattacharyya & Lilley, 1989; Rice & Chrothers, 1989), and in agreement with our interpretation, oligoH[$(G_2A)_{10}$] and oligoH[$(G_2A)_{10}$GG] show a slower electrophoretic migration than oligoH[$(G_2A)_{10}$G] (Figure 5). The patterns of DEPC and DMS modification are also consistent with this interpretation. The last guanine residue of the homopurine sequence of oligoH[$(G_2A)_{10}$GG] is reactive with DEPC (Figures 3C and 4A), which is consistent with it being unpaired. In the case of oligoH[$(G_2A)_{10}$], residue G1 does not show any significant DEPC reactivity, although, according to its high DMS reactivity and the electrophoretic behavior of the oligo, it is unpaired. Its unpaired character was corroborated by the decrease of its DMS reactivity observed in oligoH[$(G_2A)_{10}$G] and oligoH[$(G_2A)_{10}$GG]. The lack of DEPC reactivity of guanine G1 in oligoH[$(G_2A)_{10}$] might reflect a peculiar structural organization. As shown by NMR and other data, an extra purine on one strand of a DNA double helix has a high tendency to stack into the helix (Patel *et al*., 1987; Lomant & Fresco, 1973; Morden *et al*., 1983). Therefore, it is possible that residue G1, though unpaired, would still be maintained as helical through stacking interactions with the flanking guanine residues. Being helical is likely to diminish strongly its DEPC reactivity. Notice that purine residues in B-DNA are not reactive with DEPC, although their N7 groups are not directly involved in hydrogen-bonding interactions. According to this interpretation, the high DEPC reactivity of residue G32 in oligoH[$(G_2A)_{10}$GG] might reflect an extrahelical organization. Interestingly, in this case, the 3′-flanking residue is a pyrimidine and, in general, purine−pyrimidine stacks are less stable than purine−purine ones (Saenger, 1984).

Similar experiments indicate that antiparallel-stranded d(GGGA)$_n$ homoduplexes are stabilized by the formation of G·G and A·A pairs. Again, this conclusion arises from the absence of unpaired bases at the junction region of oligoH[A$(G_3A)_8$]. On the other hand, the junction region of oligoH[$(G_3A)_8$] appears to be largely distorted, extending for the first two and the last three residues of the homopurine sequence. The reason for the presence of such large structural distortion in oligoH[$(G_3A)_8$] is unclear. It is unlikely that the whole region would be unpaired since the thermal stability of oligoH[$(G_3A)_8$] is very similar to that of oligoH[A$(G_3A)_8$] (not shown). The altered chemical reactivity of oligoH[$(G_3A)_8$] might arise from the coexistence of various forms differing in the structural organization of the junction. For instance, given the significant stability of G·A pairs (Aboul-ela *et al*., 1985; Huertas *et al*., 1993), it is possible that, in a population of the molecules, the last adenine of the homopurine tract of oligoH[$(G_3A)_8$] will compete successfully with the guanine residues on top of it for pairing to the first guanine of the tract. In this way, the structural distortion could propagate in the 5′ direction for one or two residues. Alternatively, it might also be that the structure of the antiparallel-stranded d(GGGA)$_n$ homoduplex would deviate importantly from that of the right-handed B-DNA so that, in this case, the junction region is less stable and involves a larger number of bases than in the case of the d(GGA)$_n$ sequence.

A·A mismatches are less stable than either G·A or G·G ones (Aboul-ela *et al*., 1985). Therefore, purely on the basis of hydrogen-bonding interactions, the formation of A·A pairs

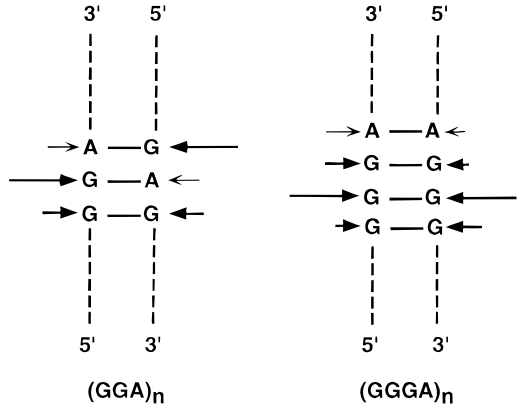

FIGURE 10: Summary of the chemical reactivity of each repeat in d(GGA)$_n$ DNA (left) and d(GGGA)$_n$ DNA (right). Arrows indicate the relative reactivity of guanine residues with DMS (thick) and of adenine residues with DEPC (thin). The length of the arrows is qualitatively proportional to the degree of reactivity.

in antiparallel-stranded d(GGGA)$_n$ DNA should result in a less stable structure. But, on the contrary, oligo(GGGA) shows a higher thermal stability than oligo(GGA). Stacking interactions, probably at the level of the GpG steps, are likely to contribute strongly to the higher stability of antiparallel-stranded d(GGGA)$_n$ DNA. Formation of A·A pairs has also been reported in parallel-stranded d(GA)$_n$ homoduplexes (Rippe *et al.*, 1992). Other A·A pairs have been proposed to occur upon protonation of the adenine residues (Saenger, 1984; Maskos *et al.*, 1993). Protonation of the adenines is unlikely to contribute to the stability of antiparallel-stranded d(GGGA)$_n$ DNA, since this structural form is stable at pH 8.3.

From the results reported here, it is difficult to unambiguously determine the precise type of base pairs being formed in antiparallel-stranded d(GGA)$_n$ and d(GGGA)$_n$ DNAs. However, the patterns of DEPC and DMS reactivity can provide information about the participation of the purine N7 groups in hydrogen-bonding interactions. Protection from DMS methylation of the guanines is generally interpreted as being indicative of the direct involvement of their N7 groups in hydrogen-bonding interactions. On the other hand, adenine residues in B-DNA are not reactive with DEPC, although their N7 groups are not hydrogen-bonded. Therefore, a lack of DEPC reactivity does not necessarily mean a direct involvement of the adenine N7 group on hydrogen bonding. As summarized in Figure 10, in the case of antiparallel-stranded d(GGA)$_n$ DNA, guanines of the G·A pairs are highly reactive with DMS, indicating that their N7 groups are not forming hydrogen bonds. Similarly, the adenine N7 groups are not likely to be involved either in hydrogen bonding since they are significantly reactive with DEPC, particularly at 20−40 °C (Figures 3 and 9A). Notice that adenines occurring at the B-DNA region become reactive only at higher temperatures, when the oligonucleotide is melted (Figure 9A). From the four different G·A pairs that have been observed experimentally (Saenger, 1984), only G(*anti*)·A(*anti*) does not involve either of the purine N7 groups in hydrogen bonding. Therefore, these results strongly suggest that the G·A pairs formed in antiparallel-stranded d(GGA)$_n$ DNA are of the G(*anti*)·A(*anti*) type. This G·A pair has been observed to occur both in crystals and in solution (Kan *et al.*, 1983; Privé *et al.*, 1987; Nikonowicz & Gorenstein, 1990). As judged from its reactivity with DMS, two different types of G·G pairs are formed in antiparallel-stranded d(GGA)$_n$ and d(GGGA)$_n$ DNAs. In antiparallel-stranded d(GGGA)$_n$ DNA, guanines of the central G·G pairs are strongly reactive with DMS, indicating that their N7 groups are not forming hydrogen bonds (Figure 10). This type of G·G pair, involving N2−N3 interactions, has been observed to occur experimentally (Wing *et al.*, 1980). In the rest of the G·G pairs formed in antiparallel-stranded d(GGA)$_n$ and d(GGGA)$_n$ DNAs, at least one of the guanines is significantly protected from DMS methylation (Figure 10), suggesting an involvement of its N7 group in hydrogen bonding. In antiparallel-stranded d(GGA)$_n$ DNA, both guanines of the G·G pairs are roughly equally protected from DMS methylation, suggesting that, likely as a consequence of the symmetric character of the oligonucleotides, there is an almost equal probability for the N7 group involved in hydrogen bonding to be provided by either of the two guanines of the pair. Notice that a G·G pair involving the N7 groups of both guanines is not stereochemically possible (Saenger, 1984). In contrast, in antiparallel-stranded d(GGGA)$_n$ DNA, there is a marked asymmetry in the degree of DMS reactivity of the guanines of each G·G pair, suggesting that, in this case, the 5′ guanine residue mainly provides the N7 group for hydrogen bonding. Two different types of G·G pairs, involving only one N7 group in the pairing, have been observed experimentally (Kim *et al.*, 1974; Robertus *et al.*, 1974; Catasti *et al.*, 1994). Our results cannot discriminate between these two G·G pairs.

These results are relevant in the context of telomeric DNA. The homopurine tracts of most telomeric repeats are formed by a combination of adenine and guanine residues containing, at least, three consecutive guanines. Telomeric DNA forms intramolecular fold-backs as well as tetrastranded complexes, the latter being stabilized by guanine tetrads [reviewed in Sundquist (1991)]. Our results indicate that the intramolecular fold-backs are likely to be stabilized by the formation of G·G and A·A pairs. The tetrastranded complexes could easily arise from the intramolecular or intermolecular association of these fold-back structures which could be considered as structural intermediates in the folding of telomeric DNA.

Homopurine sequences occur also at internal chromosomal locations (at some centromeric satellites or dispersed throughout the genome). Many of these homopurine tracts are not perfect repetitions of a simple unit, but they are rather combinations of adenine and guanine residues. From the results shown here and elsewhere (Casasnovas *et al.*, 1993; Huertas *et al.*, 1993), it can be predicted that most homopurine sequences, even if not perfectly repetitive, would have a high tendency to form intramolecular hairpins. The higher the G content of the homopurine tract, the more stable the hairpin would be. Formation of purine−purine hairpins at these internal sites, which requires local melting of the DNA duplex, can be facilitated during transcription or DNA replication. Actually, they are likely to account for the arrest of DNA replication observed at homopurine tracts both *in vitro* and *in vivo* (Rao *et al.*, 1988; Lapidot *et al.*, 1989; Samadashwily *et al.*, 1993). The existence of single-stranded DNA binding proteins which preferentially bind to pyrimidine-rich sequences (Yee *et al.*, 1991; Kolluri *et al.*, 1992; Muraiso *et al.*, 1992; Ito *et al.*, 1994) could also facilitate formation of these hairpins by maintaining the complementary pyrimidine strand as single-stranded.

REFERENCES

Abad, J. P., Carmena, M., Baars, S., Saunders, R. D. C., Glover, D. M., Ludeña, P., Sentis, C., Tyler-Smith, C., & Villasante, A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4663−4667.

Aboul-ela, F., Koh, D., & Tinoco, I., Jr. (1985) *Nucleic Acids Res. 13*, 4811−4824.

Antao, V. P., Gray, D. M., & Ratliff, R. L. (1988) *Nucleic Acids Res. 16*, 719−738.

Bernués, J., & Azorín, F. (1995) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 9, pp 1−21, Springer-Verlag, Berlin-Heidelberg.

Bhattacharyya, A., & Lilley, D. M. J. (1989) *Nucleic Acids Res. 17*, 6821−6840.

Birnboim, H. C., Sederoff, R. R., & Paterson, M. C. (1979) *Eur. J. Biochem. 98*, 301−307.

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

Casasnovas, J. M., Huertas, D., Ortiz-Lombardía, M., Kypr, J., & Azorín, F. (1993) *J. Mol. Biol. 233*, 671−681.

Catasti, P., Gupta, G., Garcia, A. E., Ratliff, R., Hong, L., Yau, P., Moyzis, R. K., & Bradbury, E. M. (1994) *Biochemistry 33*, 3819−3830.

Chou, S. H., Zhu, L., & Reid, B. R. (1994) *J. Mol. Biol. 244*, 259−268.

Emery, H. S., & Weiner, A. M. (1981) *Cell 26*, 191−201.

Evertsz, E. M., Rippe, K., & Jovin, T. M. (1994) *Nucleic Acids Res. 22*, 3293−3303.

Fang, G., & Cech, T. R. (1993a) *Cell 74*, 875−885.

Fang, G., & Cech, T. R. (1993b) *Biochemistry 32*, 11646−11657.

Ferrer, N., Azorín, F., Villasante, A., Gutiérrez, C., & Abad, J. P. (1995) *J. Mol. Biol. 245*, 8−21.

Frank-Kamenetskii, M. D., & Mirkin, S. M. (1995) *Annu. Rev. Biochem. 64*, 65−95.

Grady, D. L., Ratliff, R. L., Robinson, D. L., McCanlies, E. C., Meyne, J., & Moyzis, R. K. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1695−1699.

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

Huertas, D., Bellsolell, L., Casasnovas, J. M., Coll, M., & Azorín, F. (1993) *EMBO J. 12*, 4029−4038.

Ito, K., Sato, K., & Endo, H. (1994) *Nucleic Acids Res. 22*, 53−58.

Kan, L. S., Chandrasegaran, S., Pulford, S. M., & Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4263−4265.

Kim, S. H., Sussman, J. L., Suddath, F. L., Quigley, G. J., McPherson, A., Wang, A. H. J., Seeman, N. C., & Rich, A. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4970−4974.

Kolluri, R., Torrey, T. A., & Kinniburgh, A. J. (1992) *Nucleic Acids Res. 20*, 111−116.

Lapidot, A., Baran, N., & Manor, H. (1989) *Nucleic Acids Res. 17*, 883−900.

Lee, J. S. (1990) *Nucleic Acids Res. 18*, 6057−6060.

Lee, J. S., Evans, D. H., & Morgan, A. R. (1980) *Nucleic Acids Res. 8*, 4305−4320.

Lilley, D. M. J. (1992) *Methods Enzymol. 212*, 133−139.

Lohe, A. R., Hilliker, A. J., & Roberts, P. A. (1993) *Genetics 134*, 1149−1174.

Lomant, A. J., & Fresco, J. R. (1973) *J. Mol. Biol. 77*, 345−354.

Maniatis, T., Jeffrey, A., & van de Sande, H. (1975) *Biochemistry 14*, 3787−3794.

Manor, H., Rao, S., & Martin, R. G. (1988) *J. Mol. Evol. 27*, 96−101.

Maskos, K., Gunn, B. M., LeBlanc, D. A., & Morden, K. M. (1993) *Biochemistry 32*, 3583−3595.

Mirkin, M. S., & Frank-Kamenetskii, M. D. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 541−576.

Morden, K. M., Chu, Y. G., Martin, F. H., & Tinoco, I., Jr. (1983) *Biochemistry 22*, 5557−5563.

Muraiso, T., Nomoto, S., Yamasaki, H., Mishima, Y., & Kominami, R. (1992) *Nucleic Acids Res. 20*, 6631−6635.

Nikonowicz, E. P., & Gorenstein, D. G. (1990) *Biochemistry 29*, 8845−8858.

Ortiz-Lombardía, M., Eritja, R., Azorín, F., Kypr, J., Tejralova, I., & Vorlícková, M. (1995) *Biochemistry 34*, 14408−14415.

Palecek, E. (1991) *CRC Crit. Rev. Biochem. Mol. Biol. 26*, 151−226.

Patel, D. J., Shapiro, L., & Hare, D. (1987) *Q. Rev. Biophys. 20*, 35−112.

Privé, G. G., Heinemann, U., Chandrasegaran, S., Kan, L.-S., Kopka, M. L., & Dickerson, R. E. (1987) *Science 238*, 498−504.

Rao, S., Manor, H., & Martin, R. G. (1988) *Nucleic Acids Res. 16*, 8077−8094.

Rice, J. A., & Chrothers, D. M. (1989) *Biochemistry 28*, 4512−4516.

Rippe, K., Fritsch, V., Westhof, E., & Jovin, T. M. (1992) *EMBO J. 11*, 3777−3786.

Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C., & Klug, A. (1974) *Nature 250*, 546−551.

Saenger, W. (1984) *Principles of Nucleic Acids Structure*, Springer-Verlag, Berlin.

Samadashwily, G. M., Dayn, A., & Mirkin, S. (1993) *EMBO J. 12*, 4975−4983.

Suda, T., Mishima, Y., Asakura, H., & Kominami, R. (1995) *Nucleic Acids Res. 23*, 3771−3777.

Sundquist, W. I. (1991) *in Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.), Vol. 5, pp 1−24, Springer-Verlag, Berlin-Heidelberg.

Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., & Wohlrab, F. (1988) *FASEB J. 2*, 2939−2949.

Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature 287*, 755−758.

Yee, H. A., Wong, A. K. C., van de Sande, J. H., & Rattner, J. B. (1991) *Nucleic Acids Res. 19*, 949−953.

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

BI961020+