Articles

Structure of $d(T)_n \cdot d(A)_n \cdot d(T)_n$: The DNA Triple Helix Has B-Form Geometry with C2′-Endo Sugar Pucker

F. B. Howard,*,‡ H. Todd Miles,*,‡ Keliang Liu,‡ J. Frazier,‡ G. Raghunathan,§ and V. Sasisekharan*,‡

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, and Laboratory of Mathematical Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

Received May 14, 1992; Revised Manuscript Received August 28, 1992

ABSTRACT: The polynucleotide helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ is the only deoxypolynucleotide triple helix for which a structure has been published, and it is generally assumed as the structural basis for studies of DNA triplexes. The helix has been assigned to an A-form conformation with C3′-endo sugar pucker by Arnott and Selsing [1974; cf. Arnott et al. (1976)]. We show here by infrared spectroscopy in D_2O solution that the helix is instead B-form and that the sugar pucker is in the C2′-endo region. Distamycin A, which binds only to B-form and not to A-form helices, binds to the triple helix without displacement of the third strand, as demonstrated by CD spectroscopy and gel electrophoresis. Molecular modeling shows that a stereochemically satisfactory structure can be built using C2′-endo sugars and a displacement of the Watson–Crick base-pair center from the helix axis of 2.5 Å. Helical constraints of rise per residue (h = 3.26 Å) and residues per turn (n = 12) were taken from fiber diffraction experiments of Arnott and Selsing (1974). The conformational torsion angles are in the standard B-form range, and there are no short contacts. In contrast, we were unable to construct a stereochemically allowed model with A-form geometry and C3′-endo sugars. Arnott et al. (1976) observed that their model had short contacts (e.g., 2.3 Å between the phosphate-pendent oxygen on the A strand and O2 in the Hoogsteen-paired thymine strand) which are generally known to be outside the allowed range. We conclude that the triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ has a conformation of the B-form with sugar pucker near C2′-endo geometry. We note that since the two T strands are antiparallel, identical conformations of the two strands will result in a dyad relating the T strands.

Formation of nucleic acid triple helices by pairing of pyrimidines to the N7- and 6-substituent positions of purines in double helices was first established with high molecular weight polynucleotides and has been extensively studied in these molecules [cf. Felsenfeld et al. (1957), Felsenfeld and Rich (1957), Miles and Frazier (1964), Inman and Baldwin (1964), Chamberlin and Patterson (1965), Blake and Fresco (1965), Riley et al. (1966), Hattori et al. (1976), and Howard and Miles (1984)]. More recently, there has been a resurgence of interest in triple helices stimulated by their possible roles in biological processes [cf. Lyamichev et al. (1986), Htun and Dahlberg (1988), Mirkin et al. (1987), Frank-Kamenetskii (1990), and Hsieh et al. (1990)] and by the potential for recognition of specific sequences in duplex DNA or RNA. Applications of interest include site-specific chemical cleavage and possible targeting of drugs to specific DNA sequences [cf. Moser and Dervan (1987), Helene et al. (1989), Helene and Toulme (1990), Cooney et al. (1988), and Maher et al. (1989)]. RNA triple helices were reported to be of A-form as was the hybrid structure $r(U)_n \cdot d(A)_n \cdot r(U)_n$ (Arnott & Bond, 1973; Arnott et al., 1976). The only DNA triple helix for which a structure has been published is $d(T)_n \cdot d(A)_n \cdot d(T)_n$,[1] assigned by Arnott and co-workers to the A-form conformation with C3′-endo sugar pucker (Arnott & Selsing, 1974). The polynucleotide conformation of this A-form $d(T)_n \cdot d(A)_n \cdot d(T)_n$ model has been generally accepted and forms the conceptual structural basis for DNA triple helix research currently in progress. Helix conformation, in addition to its intrinsic significance, can be important for its control of specificity of binding of macromolecular and small ligands to DNA. Binding of the drugs distamycin and netropsin in solution, for example, is known to occur only to B- and not to A-form helices (Zimmer & Wahnert, 1986). Crystallographic studies also show that distamycin A preferentially binds to double-helical DNA in the B-form with C2′-endo sugar pucker and not in the A-form with C3′-endo sugar pucker (Kopka et al., 1985; Coll et al., 1989). This specificity is attributed to maximum contacts between the sugar phosphate backbone of the duplex DNA of B-form geometry and distamycin A (Coll et al., 1989; Dasgupta et al., 1990). In exploring the interaction of distamycin (Dst) with double-helical (Dasgupta et al., 1990) and triple-helical polynucleotides, we observed that distamycin binds to the triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ (see below). Since this triple helix has been characterized as A-form DNA, the unexpected binding result caused us to examine further the conformation of this helix.

Here, we show by infrared spectroscopy in D_2O solution that the triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ has C2′-endo sugar pucker and a backbone conformation of the B-form. Infrared bands at approximately 830 and at 810 cm^{-1} had been found to be characteristic of the backbone conformations of B- and A-forms, respectively, by spectroscopy of dry films of DNA and RNA [cf. Tsuboi (1969), Higuchi et al. (1969), and references there cited]. Subsequent Raman (solid state and

‡ National Institute of Diabetes and Digestive and Kidney Diseases.
§ National Cancer Institute.
[1] Following standard convention, the triple helix poly(dT)·poly(dA)·poly(dT) is abbreviated $d(T)_n \cdot d(A)_n \cdot d(T)_n$ throughout this paper.

solution: Ehrfurt et al., 1972; La Fleur et al., 1972; Brahms et al., 1974) and infrared (solid state: Nishimura et al., 1974; Liquier et al., 1991) studies have confirmed and extended these observations. These bands are backbone vibrations which reflect the sugar pucker and probably have also some phosphodiester stretch character. We note that sugar pucker and phosphodiester conformations are correlated and that the motion of one must affect the other. These are independent of base composition and are essentially the same in H_2O and D_2O. By use of short-path ZnSe cells and computer subtraction of solvent, we have obtained D_2O solution spectra of DNA and RNA polynucleotides in the region 750–1000 cm^{-1} (Figure 1). We observe the conformationally characteristic bands at about 810 and 830 cm^{-1} as well as strong IR bands at 865 cm^{-1} in A-form and at 970 cm^{-1} in B-form structures [these as well as other bands were reported in IR spectra of RNA and DNA films by Tsuboi (1969)]. These bands can also be observed in H_2O solution, but the spectral quality is lower because of the higher absorbance of the solvent in this region. The 970- and 865-cm^{-1} bands are usually the strongest bands below 1000 cm^{-1} in the infrared but are not observed with appreciable intensity in the Raman. This large difference in IR and Raman intensities suggests that the 865- and 970-cm^{-1} vibrations occur with large changes in dipole moment. We report here IR spectra measured in D_2O solution from 750 to 1000 cm^{-1} for the triple helix $d(T)_n{\cdot}d(A)_n{\cdot}d(T)_n$ and for a variety of double helices, which support the relevant correlation of solution IR spectra with helix conformation (we have previously reported solution spectra in this region using similar methods: Miles, 1980; Govil et al., 1981). Spectra in the well-documented double bond stretching region were also measured for all helices to establish the nature of the base-pairing interactions [cf. Miles and Frazier (1964), Howard et al. (1969), and Miles (1971)].

CD spectra and the electrophoresis experiments demonstrate binding of Dst to the triple helix and provide further support for a B-form structure for the triple helix $d(T)_n{\cdot}d(A)_n{\cdot}d(T)_n$. Computer modeling is used to show that a stereochemically satisfactory structure can be constructed with standard torsional angles without short contacts. Our model incorporates C2′-endo sugars and the values of rise per residue (h = 3.26 Å) and residues per turn (n = 12) obtained by fiber diffraction (Arnott & Selsing, 1974).

MATERIALS AND METHODS

Unless otherwise stated, all of the polynucleotides used were high molecular weight polymers obtained from Pharmacia. Poly[r(2NH_2A)] and poly[d(2NH_2A)] were prepared as described by Howard et al. (1976) and by Howard and Miles (1984), respectively. $d(T)_n{\cdot}d(A)_n{\cdot}d(T)_n$ was prepared by mixing $d(A)_n$ with $d(T)_n$ under conditions shown by Riley et al. (1966) to produce a triple helix. All experiments were carried out with high molecular weight polymers except the gel electrophoresis experiment on Dst A binding. $d(A)_{20}$ was obtained from Pharmacia. $d(C_3\text{-}T_{20})$ was synthesized using solid-state phosphoramidite chemistry with an ABI 380B DNA synthesizer. The oligomer was purified by preparative gel electrophoresis and was detected by UV shadowing over a Du Pont Cronex intensifying screen [cf. Maniatis et al. (1982)].

Infrared Spectroscopy. The IR spectra were measured in D_2O solution using ZnSe windows and path lengths in the 10–15-μm range [cf. Miles (1980) and Govil et al. (1981)]. The spectra measured with a Perkin-Elmer 580B spectrophotometer were digitized during the recording by the LDACS system as described previously (Howard & Miles, 1984) and processed by the Laboratory Analysis Package (LAP) written by John Powell and John Morris of the Computer Systems Laboratory of DCRT at NIH. The spectra were run versus air in the double-beam mode, and the solvent spectrum measured separately in the same cell at the same temperature was subtracted by computer (Miles & Frazier, 1978). Conditions are given in the legend of Figure 1.

Binding of Dst to the Triple Helix. CD spectra of the high molecular weight triple helix and of mixtures with Dst were measured with a Jasco J-500A spectropolarimeter and the data digitized and processed as described above. Experimental conditions are given in the legend of Figure 2. In order to demonstrate that the third strand is not displaced by Dst, a gel electrophoresis experiment was carried out. Oligomers containing 20 A or T residues rather than polydisperse high polymers were used in order to reveal clearly the different molecular species. The solutions were prepared in 0.2 M NaCl, 0.05 M $MgCl_2$, and 0.002 M sodium cacodylate, pH 7.0, and contained approximately 2.2×10^{-5} mmol of $d(A)_{20}$ and quantities of $d(C_3\text{-}T_{20})$ to give the ratios indicated in the legend of Figure 3. $d(C_3\text{-}T_{20})$, which had been synthesized for another purpose, was used instead of $d(T)_{20}$, but terminal $d(C)_3$ should not interfere with AT helix formation. The ratio of Dst to base was 1:1 in those samples containing Dst. After being freeze-dried in a vacuum centrifuge, the samples were dissolved in 30% glycerol and loaded on a 20% nondenaturing gel in a solvent containing 0.089 M Tris–borate buffer, pH 8.3, 2 mM EDTA, and 0.05 M $MgCl_2$. Electrophoresis was run at 350 V for 47 h at 5 °C. Buffer was recirculated by pumps to keep the pH constant.

Modeling. In the triple-helical structure proposed here for $d(T)_n{\cdot}d(A)_n{\cdot}d(T)_n$, one of the T strands is Watson–Crick hydrogen bonded and antiparallel to the A strand. The other T strand makes Hoogsteen-type hydrogen bonds with the A strand and is parallel to it. Therefore, the two T strands are antiparallel, and identical conformations of the two T backbones will lead to a dyad between the T strands as shown in Figure 4A. The A strand can be related to the antiparallel T strand by a pseudodyad. In a triple-stranded complex, however, there need not be a pseudodyad between the antiparallel T and A strands [see Arnott et al. (1976)]. As noted below, we could have generated structures with a backbone conformation for the A strand different from that of the T strand. In the structure presented here, the sugar phosphate backbones of all the three chains have identical conformations. Standard values for bond lengths and angles were used for generating the chains, and the sugar pucker used was in the C2′-endo region. Helical constraints of 12 residues per turn of the helix and 3.26 Å for the axial rise per residue, from fiber diffraction studies (Arnott & Selsing, 1974), were used, and the structures were obtained using the linked-atom least-squares procedure (Smith & Arnott, 1978).

RESULTS

Infrared Spectroscopy. The solution spectrum of the RNA double helix $r(A)_n{\cdot}r(U)_n$ has bands at 815 cm^{-1} (m) and 865 cm^{-1} (s) but no band near 830 or 970 cm^{-1}, indicating A-form geometry and C3′-endo sugar pucker. Similar spectra are observed for other ribopolynucleotide duplexes such as $r(A)_n{\cdot}r(T)_n$, $r(A\text{-}U)_n$, $r(2NH_2A)_n{\cdot}r(T)_n$, and $r(I)_n{\cdot}r(C)_n$ (Figure 1, panels B, C, D, and E; cf. Table I). In contrast, the DNA helices $d(A)_n{\cdot}d(T)_n$, $d(A\text{-}T)_n$, $d(G\text{-}C)_n$, and $d(G)_n{\cdot}d(C)_n$ have bands at 832–842 cm^{-1} (w–m) and 970 cm^{-1} (s) but no bands near 810 cm^{-1} and no strong bands near 865 cm^{-1} (Figure 1, panels F, G, M, and N), indicating that these have B-form

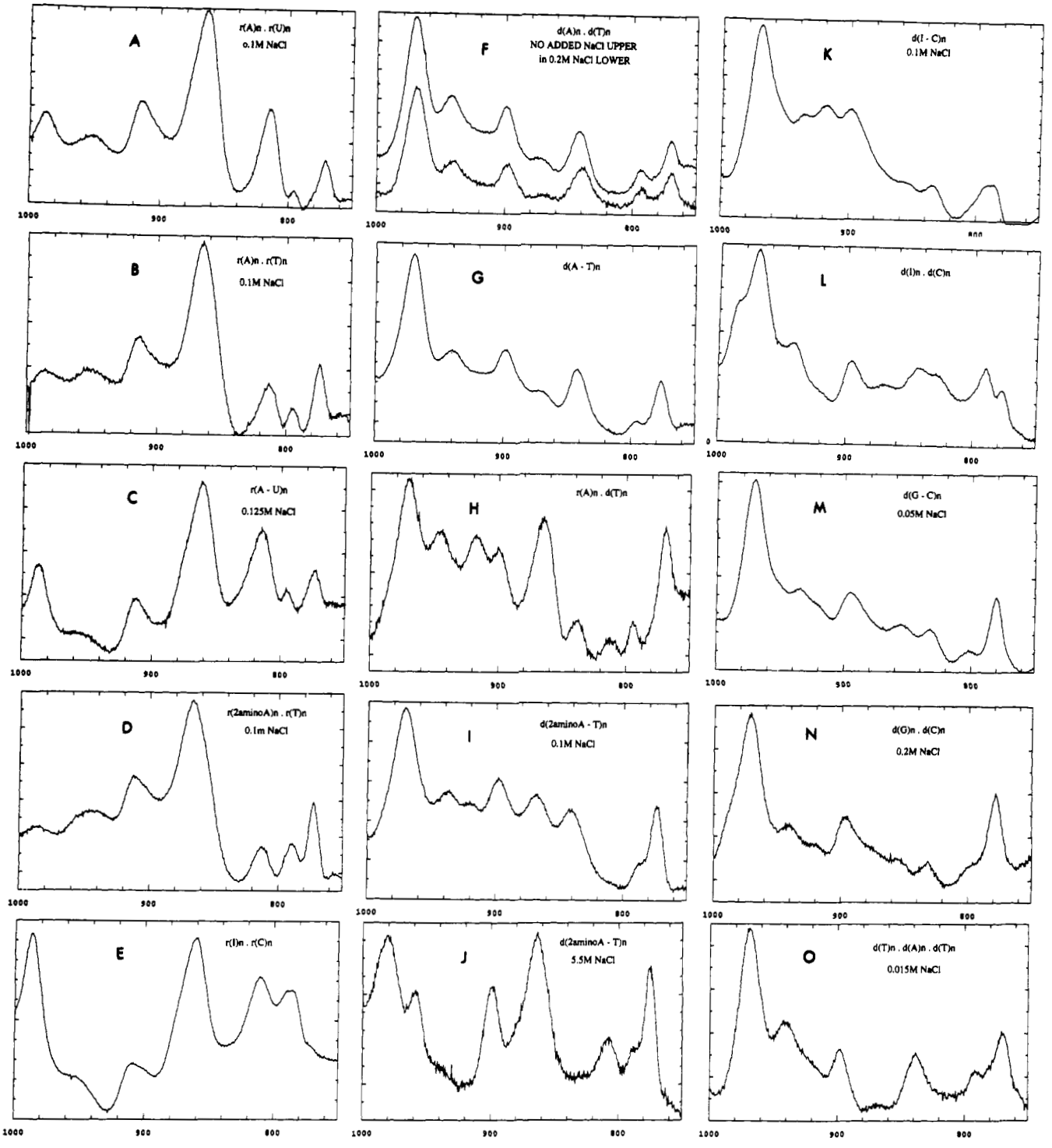

FIGURE 1: Infrared spectra of polynucleotide helices in D_2O solution (pD 7.9). These spectra were measured in ZnSe cells of 13.1-μm path length unless otherwise noted. Polymer concentrations are expressed as molarity in polymer phosphate. (A) Conditions: $r(A)_n$, 0.08 M; $r(U)_n$, 0.08 M; Na_2HPO_4, 0.02 M; NaCl, 0.1 M; 27 °C. (B) Conditions: $r(A)_n$, 0.08 M; $r(T)_n$, 0.08 M; Na_2HPO_4, 0.01 M; NaCl, 0.1 M; 27 °C. (C) Conditions: $r(A\text{-}U)_n$, 0.10 M; Na_2HPO_4, 0.025 M; NaCl, 0.125 M; 5 °C; path length, 13.5 μm. (D) Conditions: $r(2NH_2A)_n$, 0.08 M; $r(T)_n$, 0.08 M; Na_2HPO_4, 0.01 M; NaCl, 0.1 M; 28 °C. (E) Conditions: $r(I)_n$, 0.1 M; $r(C)_n$, 0.1 M; Na_2HPO_4, 0.1 M; no added NaCl; 27 °C. (F, upper curve) Conditions: $d(A)_n$, 0.05 M; $d(T)_n$, 0.05 M; Na_2HPO_4, 0.01 M; no added NaCl; 27 °C. (F, lower curve) Conditions: $d(A)_n$, 0.03 M; $d(T)_n$, 0.03 M; Na_2HPO_4, 0.005 M; NaCl, 0.2 M; 32 °C. (G) Conditions: $d(A\text{-}T)_n$, 0.13 M; Na_2HPO_4, 0.01 M; no added NaCl; 27 °C. (H) Conditions: $r(A)_n$, 0.04 M; $d(T)_n$, 0.04 M; Na_2HPO_4, 0.0075 M; no added NaCl; 27 °C. (I) Conditions: $d(2NH_2A\text{-}T)_n$, 0.16 M; Na_2HPO_4, 0.02 M; NaCl, 0.1 M; 27 °C. (J) Conditions: $d(2NH_2A\text{-}T)_n$, 0.14 M; Na_2HPO_4, 0.018 M; NaCl, 4.5 M; 27 °C. (K) Conditions: $d(I\text{-}C)_n$, 0.14 M; sodium pyrophosphate, 0.05 M; NaCl, 0.20 M; 27 °C. (L) Conditions: $d(I)_n$, 0.08 M; $d(C)_n$, 0.08 M; Na_2HPO_4, 0.1 M; no added NaCl; 27 °C. (M) Conditions: $d(G\text{-}C)_n$, 0.08 M; Na_2HPO_4, 0.01 M; NaCl, 0.05 M; 27 °C. (N) Conditions: $d(G)_n$, 0.05 M; $d(C)_n$, 0.05 M; Na_2HPO_4, 0.009 M; NaCl, 0.19 M; 27 °C. (O) Conditions: $d(T)_n$, 0.03 M; $d(A)_n$, 0.015 M; Na_2HPO_4, 0.005 M; NaCl, 0.015 M; 20 °C; path length, 21.6 μm.

conformations with C2′-endo sugar pucker. Similar spectra are observed for $d(2NH_2A\text{-}T)_n$, $d(I\text{-}C)_n$, and $d(I)_n{\cdot}d(C)_n$ (Figure 1, panels I, K, and L; cf. Table I). For the alternating copolymer $d(2NH_2A\text{-}T)_n$, which is known to undergo a structural transition with inversion of CD bands in high salt (Howard & Miles, 1983), the A rather than the Z conformation has been found in 4 M NaCl by 2D NMR (Borah et al., 1985). The IR spectrum of this polymer in high salt (Figure 1, panel J) shows A-form bands at 815 and 863 cm^{-1} and disappearance of the B-form bands at 842 and 970 cm^{-1}, in

Table I: Infrared Frequencies in cm^{-1} (D_2O Solution)[a]

(A) $(A)_n$·r(U)$_n$	(B) r(A)$_n$·r(T)$_n$	(C) r(A-U)$_n$	(D) r(2NH$_2$A)$_n$·r(T)$_n$	(E) r(I)$_n$·r(C)$_n$	(F) d(A)$_n$·d(T)$_n$ (upper curve)	(F) d(A)$_n$·d(T)$_n$ (lower curve)	(G) d(A-T)$_n$
988	986	988	986.5	987	970.5	969	970.5
952	953.5	952	945	955	942.5	940	940.5
914.5	913.5	914	912.5	909	899.5	897	898.5
864	865	875 (sh)	867	860.5	875	870	870
814.5	813.5	863.5	813	810.5	842.5	839	842.5
795.5	795	815	789.5	786	793.5	792.5	796
771	773.5	794.5	772.5		770	769	777
		773.5					

(H) r(A)$_n$·d(T)$_n$	(I) d(2NH$_2$A-T)$_n$	(J) d(2NH$_2$A-T)$_n$	(K) d(I-C)$_n$	(L) d(I)$_n$·d(C)$_n$	(M) d(G-C)$_n$	(N) d(G)$_n$·d(C)$_n$	(O) d(T)$_n$·d(A)$_n$
970.5	971	980.5	970	985	971	972	972.5
946.5	937	959	936	970	935	941	945.5
918	920	940 (sh)	918.5	941.5	895.5	920	899.5
900.5	898	899.5	899.5	895.5	855.5	896.5	870
865	868.5	865	855	872	832.5	~857 (sh)	840.5
838.5	842.5	808	834.5	843.5	801.5	831.5	792
813	788 (sh)	788 (sh)	786.5	830	780.5	~800 (sh)	770
794	773.5	775.5		790			
769				777.5			

[a] Conditions are given in the legend to Figure 1.

agreement with the NMR results. The hybrid r(A)$_n$·d(T)$_n$ helix has previously been shown to have, under highly hydrated conditions, B DNA-like helix parameters with, however, the r(A) strand having C3′-endo sugar pucker and an A-form backbone conformation (Zimmerman & Pheiffer 1981), a conclusion supported by ^{31}P NMR (Shindo & Matsumoto, 1984). The IR spectrum of r(A)$_n$·d(T)$_n$ (Figure 1, panel H) has clearly resolved bands at 813 and 865 cm^{-1}, characteristic of A-form, and likewise at 838 and 970 cm^{-1}, indicating a B conformation, in agreement with the conclusions of Zimmerman and Pfeiffer. From the above results it is clear that these IR spectra in D_2O solution can identify and distinguish A and B backbone helical conformations, either alone or in combination. Finally, for the triple helix d(T)$_n$·d(A)$_n$·d(T)$_n$, at 20 °C we observe bands at 840 and 973 cm^{-1} and none near 810 or 865 cm^{-1}, identifying the structure as B-form with sugar pucker near C2′-endo. Recently, the IR spectrum of a film of d(T)$_n$·d(A)$_n$·d(T)$_n$ has also been reported (Liquier et al., 1991) to have a band at 840 cm^{-1}, consistent with the solution spectra in the present work. Development of models to conform with these experimental data is described below.

Distamycin Binding. Distamycin (Dst) is known to bind in the minor groove of DNA and to be highly selective for conformation and for base composition. Binding occurs only to B-form and not to A-form structures (Zimmer & Wahnert, 1986). We see in Figure 2 typical CD spectra characteristic of Dst binding to DNA. Since Dst is achiral, it has no CD spectrum unless bound and, therefore, has none in the presence of RNA or A DNA. The spectrum thus indicates that the triple helix is B-form or, possibly, that the ligand has displaced the third strand (T), allowing reversion to double helix with binding of Dst. Since Dst A is known to cover a space of three to four base pairs along the minor groove (Zimmer & Wahnert, 1986), an experimental ratio of one Dst per phosphate constitutes a large excess and should ensure that displacement of the third strand would be complete if it occurs at all. We demonstrate, however, that displacement is ruled out by the gel electrophoresis experiment shown in Figure 3. Double and triple helices formed from stoichiometric amounts of d(A)$_{20}$ and d(C$_3$-T$_{20}$) migrate at distinct rates in lanes 1 and 5, respectively. When excess d(C$_3$-T$_{20}$) is added, it is readily detected moving far ahead of the triple helix in lane 7. Addition

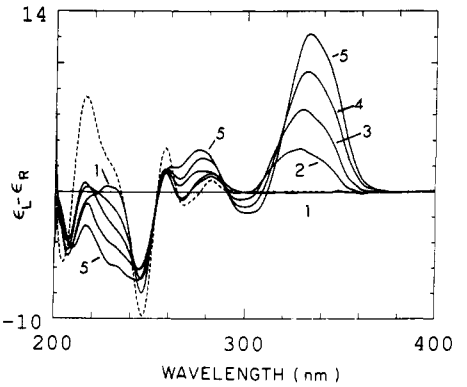

FIGURE 2: Titration of d(A)$_n$·2d(T)$_n$ with distamycin. Development of a new CD band near 330 nm accompanies the addition of 1.53×10^{-4} M distamycin to 2.24×10^{-4} M d(A)$_n$·2d(T)$_n$ (polymer P), the [Dst]/[polymer P] ratio (r) being equal to 0 (curve 1), 0.025 (curve 2), 0.051 (curve 3), 0.101 (curve 4), and 0.222 (curve 5). The dashed curve is the spectrum of d(A)$_n$·d(T)$_n$. Conditions: 0.002 M sodium cacodylate, pH 7.0, 0.2 M NaCl, 20.0 °C. Each CD spectrum is the average of nine scans and has been baseline corrected and smoothed by a computer.

of Dst to the double helix causes it to move at a faster rate (lanes 2 and 4). The complex formed by Dst and the triple helix also moves faster than the parent helix (lane 6), but at a rate distinctly slower than the complex of Dst with the duplex (lane 2), demonstrating that the complex is not that of a duplex formed by displacement of a third strand. There is, moreover, no d(C$_3$-T$_{20}$) visible, as there would be if it had been displaced from the triple helix. When an excess of d(C$_3$-T$_{20}$) is added, it is clearly visible, moving at its usual position far ahead of both the triple helix and the Dst–triple helix complex (lanes 7 and 8, respectively). Confirmation of the electrophoresis result is provided by the CD spectra. The CD spectrum of the double helix d(A)$_n$·d(T)$_n$ has a strong positive band at 216 nm, whereas the triple helix has no positive band near this wavelength [cf. Greve et al. (1977) and Gray et al. (1978)]. If displacement of the third strand had occurred, then the resulting double helix would have a band at 216 nm in the presence of distamycin, but none is observed in these experiments. We conclude that the third strand is not displaced by Dst and that binding of Dst to the triple helix confirms that it has a B rather than an A conformation. A preliminary

FIGURE 3: Electrophoresis of oligonucleotide helices and their complexes with Dst on a nondenaturing gel. Lanes (from left to right): 1, pd(A)$_{20}$·d(C$_3$-T$_{20}$) (T/A = 1.03); 2, pd(A)$_{20}$·d(C$_3$-T$_{20}$) + Dst (T/A = 1.02, Dst/base = 1.00); 3, pd(A)$_{20}$·1.5d(C$_3$-T$_{20}$) (T/A = 1.51); 4, pd(A)$_{20}$·1.5d(C$_3$-T$_{20}$) + Dst (T/A = 1.58, Dst/base = 1.03); 5, pd(A)$_{20}$·2d(C$_3$-T$_{20}$) (T/A = 2.02); 6, pd(A)$_{20}$·d(C$_3$-T$_{20}$) + Dst (T/A = 2.07, Dst/base = 1.00); 7, pd(A)$_{20}$·3d(C$_3$-T$_{20}$) (T/A = 3.03); 8, pd(A)$_{20}$·3d(C$_3$-T$_{20}$) + Dst (T/A = 3.05, Dst/base = 1.02); outermost lanes, marker dyes. Conditions: 0.1 M Tris–borate, pH 8.3, 0.2 mM EDTA, 0.05 $MgCl_2$, 350–550 V, 47 h, 5 °C. Each sample contained about 2.2×10^{-5} mmol (base) of pd(A)$_{20}$. As indicated in the text d(C$_3$-T$_{20}$) is not displaced from pd(A)$_{20}$·2d(C$_3$-T$_{20}$) upon binding of distamycin.

modeling study indicates that there is no difficulty accommodating Dst within the minor groove of the triple helix proposed in this paper.

Modeling Studies. While it has not previously been recognized that a dyad symmetry element could be present in a triple helix [see, for example, Arnott and Selsing (1974) and Arnott et al. (1976)], we note that such a dyad axis exists between the two antiparallel T strands in the d(T)$_n$·d(A)$_n$·d(T)$_n$ triple helix. As shown in Figure 4A, the dyad is perpendicular to the helix axis. Several structures with sugar pucker in the C2′-endo region and helical constraints of n = 12 and h = 3.26 Å were analyzed. The structures were examined for allowed backbone torsional angles, absence of unfavorable nonbonded short distances, and hydrogen-bonding geometries of all interstrand base pairs. Our studies indicate that only a very small region of conformational space is available for any triple helix on the basis of the experimentally determined constraints (rise per residue, residues per turn, and sugar pucker). The principal reason for this is that accommodation of the third strand, with its additional sets of hydrogen bonds and potential nonbonded interactions, introduces far more constraints than are present in a double helix. The modeling clearly shows that the value of displacement D must be no less than 2.5 Å in order to make the Hoogsteen-type hydrogen bonds, and values larger than 2.6 or 2.7 Å cause unallowed short contacts. Panels B and C of Figure 4 give the top and side views of the triple-stranded structure for d(T)$_n$·d(A)$_n$·d(T)$_n$. The torsional angles for the monomer are given in Table II.

DISCUSSION

The structure for the d(T)$_n$·d(A)$_n$·d(T)$_n$ triple helix presented here has sugar puckers in the C2′-endo region, 12 residues per turn of the helix, and an axial rise per residue of 3.26 Å. The structure is consistent with the infrared spectral results (see above), as well as CD and gel electrophoresis experiments that point to a C2′-endo sugar pucker and a B-form backbone geometry. It is also consistent with a recent NMR study (Macaya et al., 1992) carried out using the approach of phase-sensitive correlated spectroscopy (P.COSY). It was concluded that the majority of sugars in their short intramolecular triple helix have C2′-endo sugar conformations. In an earlier NMR study of a hexamer, Umemoto et al. (1985) concluded that sugars in their hexamer complex have C3′-endo conformations on the basis of NOESY and MINSY estimates of a single interproton distance.

Triple helices with C3′-endo sugars result in short nonbonded interatomic distances (Arnott & Selsing, 1974; Arnott et al., 1976). This is due to the large displacement of the bases from the helix axis and a resulting shorter chain separation between the A strand and the parallel Hoogsteen T strand. We find, however, that even with C2′-endo sugar geometry, similar large displacements (D > 2.7 Å) of the bases from the helix axis always result in sterically unfavorable close distances between the two chains. Inasmuch as C3′-endo sugar pucker is generally associated with a larger displacement of the bases from the helix axis, close contacts will persist in models having C3′-endo sugar puckers and h = 3.26 Å (with consequent small or zero tilt of the bases); hence, no further attempt was made to construct models with C3′-endo sugar puckers. There are no short nonbonded distances in our proposed model. There is a 5° tilt of the bases with respect to the helical axis; however, structures with zero tilt are also feasible.

The structure presented here was generated while maintaining identical backbone conformations for all three strands. Structures in which the conformation of the A strand is not identical to that of the T strands, however, can easily be obtained. Since this would increase the number of adjustable parameters, it would be easier to attain a stereochemically satisfactory structure that fits the helical parameters.

Arnott and co-workers based their structure for d(T)$_n$·d(A)$_n$·d(T)$_n$ primarily on stereochemical considerations rather than on conventional fiber diffraction methods, since quantifiable diffraction data were meager (Arnott & Selsing, 1974). Later, using additional stereochemical criteria, Arnott et al. (1976) proposed a second, similar structure. In both structures, however, as they have pointed out, there are short contacts between nonbonded atoms (e.g., 2.3 Å between the phosphate-pendent oxygen on the A strand and O2 in the Hoogsteen-paired T strand). Using the helical parameters n (number of residues per turn) and h (axial rise per residue) obtained from X-ray diffraction, they proposed an A genus helix with C3′-endo sugar pucker and with bases positioned more than 3 Å from the helix axis. Since values of rise per residue and number of residues per turn can be reliably obtained even from poor-quality fiber diffraction data, we have taken their values of h = 3.26 Å and n = 12. Using these values and the computational methods described above, we have obtained a structure with C2′-endo sugar pucker, a displacement of 2.5 Å of the bases from the helix axis, and a 5° tilt of the base with respect to the helix axis.

In the literature [cf. Saenger (1984)] DNA has been classified into essentially two major forms, B-form with C2′-endo sugar pucker and A-form with C3′-endo sugar pucker. Primarily on the basis of X-ray crystallographic and fiber diffraction data, we note that four structural parameters have been used to distinguish the A and B conformations: (1) sugar pucker, (2) axial rise per residue, h (this value is directly related to the tilt of the planes of the base pairs with respect to the helix axis) [see Saenger et al. (1984)], (3) displacement of the bases from the helix axis (D, defined as the distance between the midpoint of the line joining C8 and C6 and the helix axis), and (4) the number of residues per turn, n. The

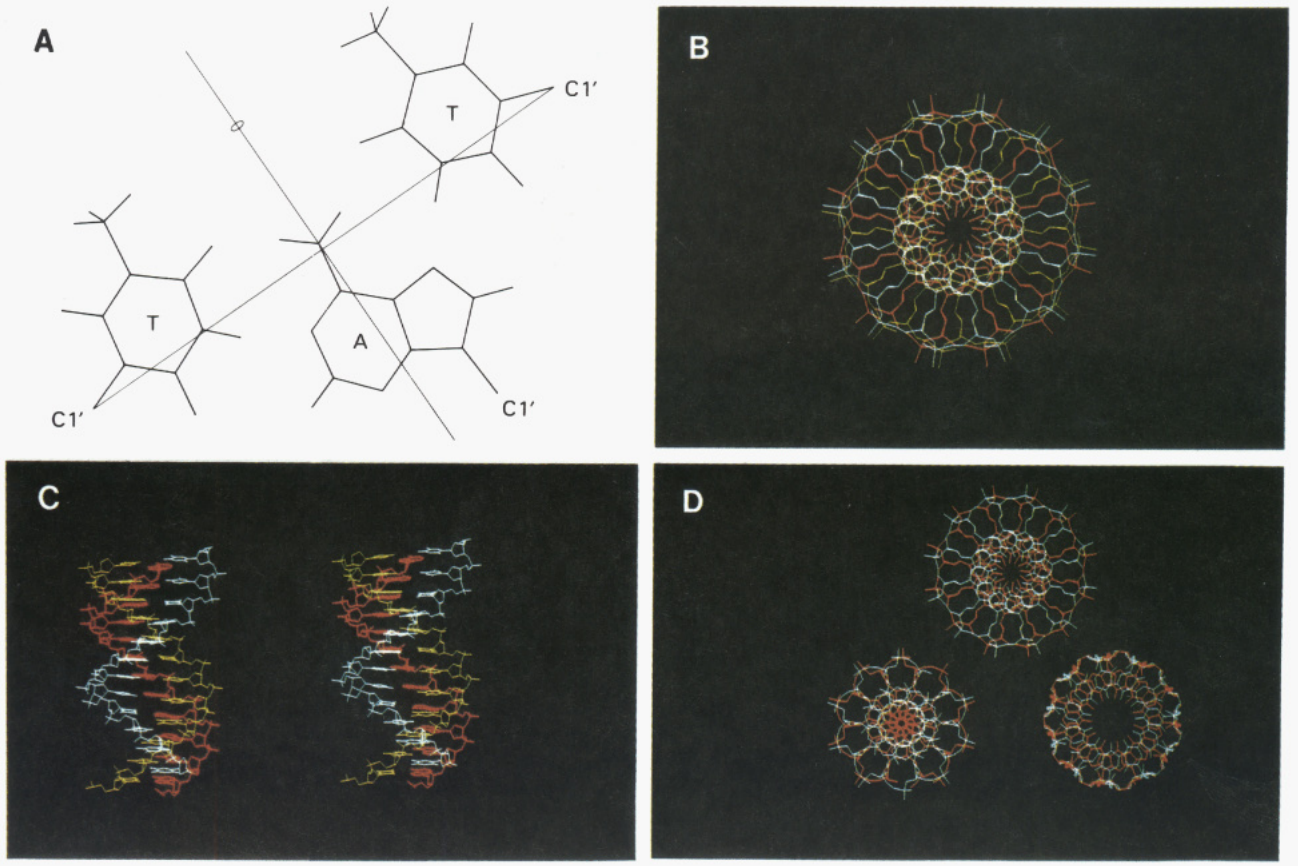

FIGURE 4: (A) Schematic diagram perpendicular to the helix axis showing Watson–Crick and Hoogsteen base pairing schemes for the $d(T)_n$·$d(A)_n$·$d(T)_n$ triple helix and the dyad symmetry between the two antiparallel T strands. (B and C) Top and side views of the $d(T)_n$·$d(A)_n$·$d(T)_n$ triple helix. The adenine strand (shown in red) is Watson–Crick hydrogen bonded to the thymine strand (shown in cyan) and makes Hoogsteen-type hydrogen bonds to the thymine strand (shown in yellow). There are 12 units per turn, and the rise per residue is 3.26 Å. (D) Comparison of the antiparallel thymine and adenine strands of the $d(T)_n$·$d(A)_n$·$d(T)_n$ triple helix proposed in this paper (top) with B-DNA (Arnott & Hukins, 1972) (bottom left) and A-DNA (Arnott et al., 1972) (bottom right). The sugar phosphate backbone and the disposition of the bases in the groove for the triple helix are more similar to B-DNA than to A-DNA.

Table II: Torsional Angles (in deg) for the Monomer of the $d(T)_n$·$d(A)_n$·$d(T)_n$ Triple Helix[a]

α (P–O5′)	–36.6 (g^-)	ϵ (C3′–O3′)	–163.5 (t)
β (O5′–C5′)	157.7 (t)	ζ (O3′–P)	–125.7 (t)
γ (C5′–C4′)	36.1 (g^+)	χ (C1′–N9)	52.5 (anti)
δ (C4′–C3′)	130.2 (E^2)		

[a] The symbols used to denote the corresponding conformations are given in parentheses. In the conformational space, the B-form is characterized by ($g^- t g^+ E^2 t t$) and the A-form by ($g^- t g^+ E^3 t g^-$). The conformation of the triple helix is characterized by ($g^- t g^+ E^2 t t$), which is the same as the B-form.

widths and depths of the major and minor grooves are also very different in the two forms; these values, however, are not independent but are determined by the foregoing parameters [cf. Saenger (1984)]. Double-helical DNA in the B family usually has the following ranges of parameters: sugar pucker in the C2′-endo region, an axial rise per residue of 3.03–3.37 Å, displacement in the range of –0.2 to 1.5 Å, and 10 residues per turn of the helix (in solution this value increases to 10.3–10.6, and in the C and D conformations, which are also considered members of the B family, it has values as low as 9.3 and 8). Similarly, double-helical DNA in the A-form usually has the following ranges of parameters: sugar pucker in the C3′-endo region, an axial rise per residue for double helices in the range 2.56–3.03 Å (with a consequent larger tilt of the bases), a displacement in the range 3.6–4.5 Å, and 11 or 12 residues per turn of the helix. Arnott and co-workers have assigned triple helices, both RNA and DNA, to the A family with C3′-endo sugar pucker. We conclude that for the DNA triple helix $d(T)_n$·$d(A)_n$·$d(T)_n$ the assignment to A-form with C3′-endo sugars is incorrect because of the contrary experimental evidence presented here and because of disallowed short contacts present in their model. We show here that the DNA triple helix $d(T)_n$·$d(A)_n$·$d(T)_n$ belongs to the B family with C2′-endo sugar pucker and not to the A family with C3′-endo sugar pucker. Molecular modeling indicates that only a very small region of the conformational space is available for any triple helix. The principal reason for this is that accommodation of the third strand with its additional sets of hydrogen bonds and potential nonbonded contacts introduces far more constraints than are present in a double helix. For example, the value of D must be no less than 2.5 Å in order to make Hoogsteen-type hydrogen bonds. Further, we found from geometry alone, for a given backbone conformation, that both D and n are related. The larger the value of D, the smaller the twist angle subtended at the helix axis by the projection of the chord between identical atoms of adjacent nucleotides. At D = 2.5 Å, the twist angle turns out to be 30°, and hence n must be 12. Moreover, values of D = 2.5 Å and n = 12, as given for our model, have been reported to be energetically satisfactory for B DNA double helices by Zhurkin et al. (1978). Thus all four structural parameters for our model lie in the B range, whereas three of the four parameters are not in the A range. A comparison of our structure with the standard A and B structures is shown in Figure 4D. Further, of the six torsional angles which describe the backbone conformation, four are common for A-

and B-forms of DNA, and two are different. Our model has conformations for these two torsional angles, namely, C4'–C3' and O3'–P, which are the same as those for the B-form (E^2 and trans) and different from those for the A-form (E^3 and gauche$^-$). We therefore conclude that the triple helix $d(T)_n{\cdot}d(A)_n{\cdot}d(T)_n$ is a member of the B rather than the A family of helices.

ACKNOWLEDGMENT

We thank Drs. R. L. Jernigan of the National Cancer Institute and S. B. Zimmerman of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH, for discussions and critical reading of the manuscript.

REFERENCES

Arnott, S., & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun. 47*, 1504–1509.
Arnott, S., & Bond, P. J. (1973) *Science 181*, 68–69.
Arnott, S., & Selsing, E. (1974) *J. Mol. Biol. 88*, 509–521.
Arnott, S., Hukins, D. W. L., & Dover, S. D. (1972) *Biochem. Biophys. Res. Commun. 48*, 1392–1399.
Arnott, S., Bond, P. J., Selsing, E., & Smith, P. J. C. (1976) *Nucleic Acids Res. 3*, 2459–2470.
Blake, R. D., & Fresco, J. R. (1966) *J. Mol. Biol. 19*, 145–160.
Borah, B., Cohen, J. S., Howard, F. B., & Miles, H. T. (1985) *Biochemistry 24*, 7456–7462.
Brahms, S., Brahms, J., & Pilet, J. (1974) *Isr. J. Chem. 12*, 153–163.
Brown, E. B., & Peticolas, W. L. (1975) *Biopolymers 14*, 1259–1271.
Chamberlin, M. J., & Patterson, D. L. (1965) *J. Mol. Biol. 12*, 410–428.
Coll, M., Aymami, J., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1989) *Biochemistry 28*, 310–320.
Cooney, M., Czernoszewicz, G., Postel, E., Flint, S. J., & Hogan, M. E. (1988) *Science 245*, 456–459.
Dasgupta, D., Howard, F. B., Sasisekharan, V., & Miles, H. T. (1990) *Biopolymers 30*, 223–227.
Erfurth, S. C., Kiser, E. J., & Peticolas, W. L. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 938–941.
Felsenfeld, G., & Rich, A. (1957) *Biochim. Biophys. Acta 26*, 457–468.
Felsenfeld, G., Davies, D. R., & Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023–2024.
Frank-Kamenetskii, M. D. (1990) in *DNA Topology and its Biological Effects* (Cozarelli, N. R., & Wang, J. C., Eds.) pp 185–215, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Goodwin, D. D., & Brahms, J. (1978) *Nucleic Acids Res. 5*, 835–850.
Govil, G., Fisk, C. L., Howard, F. B., & Miles, H. T. (1981) *Biopolymers 20*, 573–603.
Gray, D. M., Morgan, A. R., & Ratliff, R. L. (1978) *Nucleic Acids Res. 5*, 3679–3695.
Greve, J., Maestre, M. F., & Levin, A. (1977) *Biopolymers 16*, 1489–1504.
Hattori, M., Frazier, J., & Miles, H. T. (1976) *Biopolymers 15*, 523–531.
Helene, C., & Toulme, J. J. (1990) *Biochim. Biophys. Acta 1049*, 99–125.
Helene, C., Thuong, N. T., Saison-Behmoaras, T., & Francois, J. C. (1989) *Trends Biotechnol. 7*, 310–315.
Higuchi, S., Tsuboi, M., & Iitaka, Y. (1969) *Biopolymers 7*, 909–916.
Howard, F. B., & Miles, H. T. (1983) in *Nucleic Acids: The vectors of life* (Pullman, B., & Jortner, J., Eds.) pp 511–520, D. Reidel Publishing Co., Dordrecht, The Netherlands.
Howard, F. B., & Miles, H. T. (1984) *Biochemistry 23*, 6723–6732.
Howard, F. B., Frazier, J., & Miles, H. T. (1969) *Proc. Natl. Acad. Sci. U.S.A. 64*, 451–458.
Howard, F. B., Frazier, J., & Miles, H. T. (1976) *Biochemistry 15*, 3783–3795.
Hsieh, P., Camerini-Otero, C. S., & Camerini-Otero, R. D. (1990) *Genes Dev. 4*, 1951–1963.
Htun, H., & Dahlberg, J. E. (1988) *Science 241*, 1791–1796.
Inmam, R. B., & Baldwin, R. L. (1964) *J. Mol. Biol. 8*, 452–469.
Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376–1380.
La Fleur, L., Rice, J., & Thomas, G. J., Jr. (1972) *Biopolymers 11*, 2423–2437.
Liquier, J., Coffinier, P., Firon, M., & Taillandier, E. (1991) *J. Biomol. Struct. Dyn. 9*, 437–445.
Lyamichev, V. I., Mirkin, S. M., & Frank-Kamenetskii, M. D. (1986) *J. Biomol. Struct. Dyn. 3*, 667–669.
Macaya, R., Wang, E., Schultze, P., Sklenar, V., & Feigon, J. (1992) *J. Mol. Biol. 225*, 755–773.
Maher, L. J., III, Wood, B., & Dervan, P. B. (1989) *Science 245*, 725–730.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning*, pp 173–178, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Miles, H. T. (1971) *Proced. Nucleic Acid Res. 2*, 205–232.
Miles, H. T. (1980) *Biomol. Struct. Evol. 2*, 251–265.
Miles, H. T., & Frazier, J. (1964) *Biochem. Biophys. Res. Commun. 14*, 21–28, 129–136.
Miles, H. T., & Frazier, J. (1978) *Biochemistry 17*, 2920–2927.
Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filipov, S. A., & Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495–497.
Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645–650.
Nishimura, Y., Morikawa, K., & Tsuboi, M. (1974) *Bull. Chem. Soc. Jpn. 47*, 1034–1044.
Riley, M., Maling, B., & Chamberlin, M. J. (1966) *J. Mol. Biol. 20*, 359–389.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, Berlin.
Shindo, H., & Matsumoto, U. (1984) *J. Biol. Chem. 259*, 8682–8684.
Smith, P. J. C., & Arnott, S. (1978) *Acta Crystallogr. A34*, 3–11.
Thomas, G. A., & Peticolas, W. L. (1983) *J. Am. Chem. Soc. 105*, 986–992, 993–996.
Tsuboi, M. (1969) *Appl. Spectrosc. Rev. 3*, 54–55.
Tsuboi, M., Matsuo, K., Shimanouchi, T., & Kyogoku, Y. (1963) *Spectrochim. Acta 19*, 1617–1618.
Umemoto, K., Sarma, M. H., Gupta, G., Luo, J., & Sarma, R. H. (1990) *J. Am. Chem. Soc. 112*, 4539–4545.
Zhurkin, V. B., Lysov, Y. P., & Ivanov, V. I. (1978) *Biopolymers 17*, 377–412.
Zimmer, Ch., & Wahnert, U. (1986) *Prog. Biophys. Mol. Biol. 41*, 31–112.
Zimmerman, S. B., & Pheiffer, B. H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 78–82.