Symmetry and Molecular Structure of a DNA Triple Helix: $d(T)_n \cdot d(A)_n \cdot d(T)_n$

G. Raghunathan,*,‡ H. Todd Miles,*,§ and V. Sasisekharan*,§

Laboratory of Mathematical Biology, Room 217, 6010 Executive Boulevard, National Cancer Institute, National Institutes of Health, and Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Building 2, Room 201, National Institutes of Health, Bethesda, Maryland 20892

Received September 11, 1992; Revised Manuscript Received November 4, 1992

ABSTRACT: A structure for the triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ consistent with recent infrared spectral data is proposed, and its salient features are discussed. The present structure preserves the pseudodyad between the Watson–Crick base-paired adenine and thymine strands and in addition has a pseudorotational symmetry relating the Hoogsteen-paired adenine and thymine strands. The simultaneous presence of these two symmetries gives rise to a dyad between the two thymine polynucleotides. These symmetries result in identical backbone conformations for all three strands, unlike any previously proposed model for a triple helix. The proposed structure has an axial rise per residue of 3.26 Å and 12 residues per turn obtained from X-ray fiber diffraction [Arnott S., & Selsing, E. (1974) *J. Mol. Biol. 88*, 509–521]. The present structure is structurally and conformationally similar to double helical B-form DNA and has sugar pucker in the C2′-endo region. This structure is fundamentally different from the one proposed by Arnott and co-workers, which was based on structural and conformational features similar to double helical A-form DNA with C3′-endo sugar pucker. It is stereochemically satisfactory, and it does not have the disallowed nonbonded distances present in the earlier model of Arnott and co-workers. It is energetically much more favorable than their structure. Coordinates of the present structure are given.

In the last few years, there has been a revival of interest in triple helices both due to their role in some of the molecular processes of biology and also because of their increasing potential as therapeutic agents (Postel et al., 1991; Orson et al., 1991; Helene et al., 1989; Maher et al., 1990). Triple helices are potential inhibitors of DNA replication (Birg et al., 1990) and transcription (Cooney et al., 1988; Maher et al., 1990) and could arrest DNA synthesis (Baran et al., 1991). The third strand can also block DNA recognition by sequence-specific DNA binding proteins such as restriction endonucleases (Maher et al., 1989). Their selective binding properties have also been utilized to achieve cuts and excisions at specific sequences by attaching DNA cleaving agents (Moser & Dervan, 1987, Helene & Touleme, 1990). In all these systems, the pyrimidine oligomer is positioned parallel to the purine strand of the Watson–Crick duplex.

Although there has been a resurgence of activity on these molecules recently, the formation of the triple helix was first observed more than 3 decades ago in synthetic $r(U)_n \cdot r(A)_n \cdot r(U)_n$ (Felsenfeld et al., 1957). Later, it was shown that triple helices involving G and C could be formed by protonation of poly (C) residues to form Hoogsteen hydrogen bonds (Howard et al., 1964) and in alternating sequence of the high polymer $d(CT)_n \cdot d(GA)_n \cdot r(C^+U)_n$ (Morgan & Wells, 1968). In the above polymers, the pyrimidine strand runs parallel to the purine strand of the Watson–Crick double helix and is positioned to form Hoogsteen-type hydrogen bonds (Hoogsteen, 1963) with the purine. A similar Hoogsteen triple helix was proposed for $d(TC)_n \cdot d(GA)_n \cdot d(TC^+)_n$ at acidic pH (Lee et al., 1979). This type of triple helix was also suggested for a purine–pyrimidine tract in the human thyroglobulin gene from S1-nuclease digestion patterns (Cristophe et al., 1985). A similar structural element was indicated for the superhelical DNA with a sequence $d(GA)_{16} \cdot d(TC^+)_{16}$ (Lyamichev et al., 1986) and later for supercoiled plasmids (Htun & Dahlberg, 1988; Hanvey et al., 1988; Johnston, 1988; Kohwi & Kohwi-Shigematsu, 1988). Another type of triple helix with sequence specificity was suggested to be a stable intermediate during homologous recombination mediated by rec A protein (Hsieh et al., 1990; Zhurkin et al., 1991).

Despite the intense interest and activity on triple helices, however, there is very little precise structural information currently available for these complexes. The only DNA triple helical structure in atomic detail is that of Arnott and co-workers for $d(T)_n \cdot d(A)_n \cdot d(T)_n$ (Arnott & Selsing, 1974; Arnott et al., 1976). They used molecular modeling and limited fiber diffraction data to conclude that the structure is A-form DNA and has C3′-endo sugar pucker. This model has been generally accepted and has formed the structural basis for most subsequent work on DNA triple helices. Recent solution infrared spectral studies on the polynucleotide helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$, however, have clearly shown that it has B-form structure and sugar pucker in the C2′-endo region (Howard et al., 1992). Solid-state films of this polymeric helix (Liquier et al., 1991) also have bands consistent with this conclusion, and NMR studies of different oligonucleotide model systems (Macaya et al., 1992) indicate that many of the sugar residues have puckers near C2′-endo. Thus, the model of Arnott and co-workers for DNA triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ with an A-form structure and C3′-endo sugar pucker is inconsistent with these new experimental findings. In addition, the structure proposed by Arnott and co-workers is stereochemically unsatisfactory (see Results and Discussion). In this paper we present and discuss a structure for the $d(T)_n \cdot d(A)_n \cdot d(T)_n$ triple helix that is fundamentally different from that of Arnott and co-workers, and its energy is significantly more favorable than the structure proposed by them. Further, we show that the three strands of the triple helix can be related by symmetry so that the backbone conformations of all three strands are identical.

‡ National Cancer Institute.
§ National Institute of Diabetes and Digestive and Kidney Diseases.

0006-2960/93/0432-455$04.00/0

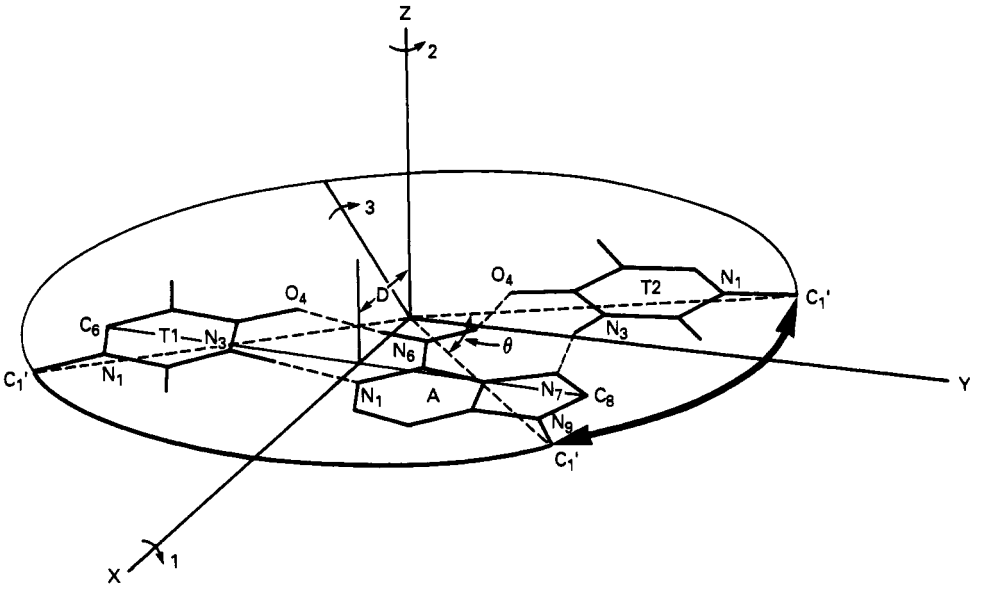

FIGURE 1: Schematic diagram showing the base pairing scheme for the triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ and symmetry elements used in the construction of the helix. Watson–Crick-paired T1 and A are antiparallel to each other and related by a pseudodyad about the *x*-axis (marked as 1). Hoogsteen-paired A and T2 are related by a pseudorotational symmetry by an angle θ about the *z*-axis (marked as 2). Application of the pseudodyad and the pseudorotational symmetry leads to a dyad symmetry between the two identical and antiparallel strands formed by T1 and T2 about an axis as shown, perpendicular to the *Z*-axis (in the *xy* plane and marked as 3). For the present case of the triple helix, substitution of the value $\theta = 69.5°$ (Table II) in the matrix 3 (Appendix) shows that the angle between the axes 3 and 1 in the *xy* plane is 145.3°. *D* is the displacement of the center of Watson–Crick base pair from the helix axis.

METHODS

The triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ presented here can be considered to be composed of two double helices: an antiparallel Watson–Crick helix formed between T1 and A strands and a parallel Hoogsteen helix formed between A and T2 strands, the A strand being common to both double helices (Figure 1). In arriving at the symmetry in the triple helix, one therefore can utilize the symmetries of the component double helices. The pseudodyad symmetry of the Watson–Crick duplex is well known and can exist also between the T1 and A strands of the triplex. Though less familiar, Hoogsteeen duplexes are also capable of independent existence. Double-stranded Hoogsteen ribonucleotide helices with blocking groups at C2 of modified A residues were described a number of years ago (Ishikawa et al., 1972; Ikehara et al., 1972; Hattori et al., 1974). We have also recently obtained DNA Hoogsteen duplexes of normal unmodified bases with parallel strands (K. Liu, V. Sasisekharan, and H. T. Miles, unpublished experiments). It has not been previously recognized that there can be a pseudorotational symmetry axis in a duplex involving A (or G) and T (or C^+) strands with Hoogsteen hydrogen bonds. We observe, however, that the backbone strands of a Hoogsteen duplex are related by a pseudorotational symmetry, just as a pseudodyad symmetry relates the strands of a Watson–Crick duplex. In the latter case, the helix axis is perpendicular to the pseudodyad axis and can lie anywhere on the pseudodyad. In a Hoogsteen duplex, on the other hand, the helix axis coincides with the pseudorotational axis. From well-known principles of symmetry operations [see, for example, McKie and McKie (1986)], simultaneous application of the pseudodyad (*x*-axis in Figure 1) for the T1 and A (Watson–Crick) strands of the triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$ and the pseudorotational symmetry between the Hoogsteen paired A and T2 strands about the helical axis (*z*-axis in Figure 1) leads to an exact dyad symmetry between the two thymine strands, T1 and T2, about an axis perpendicular to the helix axis (*z*-axis), as shown in Figure 1. These symmetry elements (see Appendix) are constraints in the molecular modeling procedure adopted.

A modeling study was undertaken by observing the effect on Hoogsteen hydrogen bonds [N7 (A)···N1 (T2) and N6 (A)···O4 (T2)] of varying simultaneously the following parameters: displacement (*D*), tilt of the bases, and rotational angle, θ between the Hoogsteen-paired A and T2 strands about the helix axis (see Figure 1). We observed that for a given *D* and θ, tilt had negligible influence on the two hydrogen bond lengths. The variations of these two hydrogen bond lengths with respect to *D* and θ for a 0° tilt is shown in Figure 2. The length of the N7···N1 hydrogen bond varies more rapidly than the N6···O4 hydrogen bond for a given increment of variation in θ. Hydrogen bond lengths, in the range of 2.8–3.0 Å for N7···N1 and in the range of 2.7–2.9 Å for N6···O4 are realizable for *D* values between 2.4 and 2.65 Å and θ values between 68° and 71°. Outside these narrow limits of *D* and θ, the hydrogen bond lengths assume unacceptable values. In addition, larger values of *D* bring the Hoogsteen-paired adenine and thymine strands too close, resulting in prohibitively short nonbonded distances. Thus, only an extremely small range of *D* was found to be allowed for the triple helix (see Figure 2), while maintaining symmetry among all three strands—a dyad perpendicular to the helix axis between the identical thymine strands, a pseudodyad between the Watson–Crick hydrogen-bonded thymine and adenine strands, and a consequent pseudorotational symmetry between the Hoogsteen-paired thymine and adenine strands (see Figure 1). Helical parameters, number of residues per turn ($n = 12$), axial rise per residue ($h = 3.26$ Å) as obtained from X-ray fiber diffraction (Arnott & Selsing, 1974), and sugar pucker in the C2′-endo region as shown by infrared spectroscopy were used; structures were obtained using the linked-atom least-squares procedure (Smith & Arnott, 1978). In the structure discussed in this paper, all three strands have identical sugar–phosphate backbone conformations. However, structures in which the conformation of the adenine strand is not identical to those of the thymine strands can easily be obtained by preserving only the dyad between the identical thymine strands. Since this different conformation of A would increase the number of adjustable parameters in the optimi-

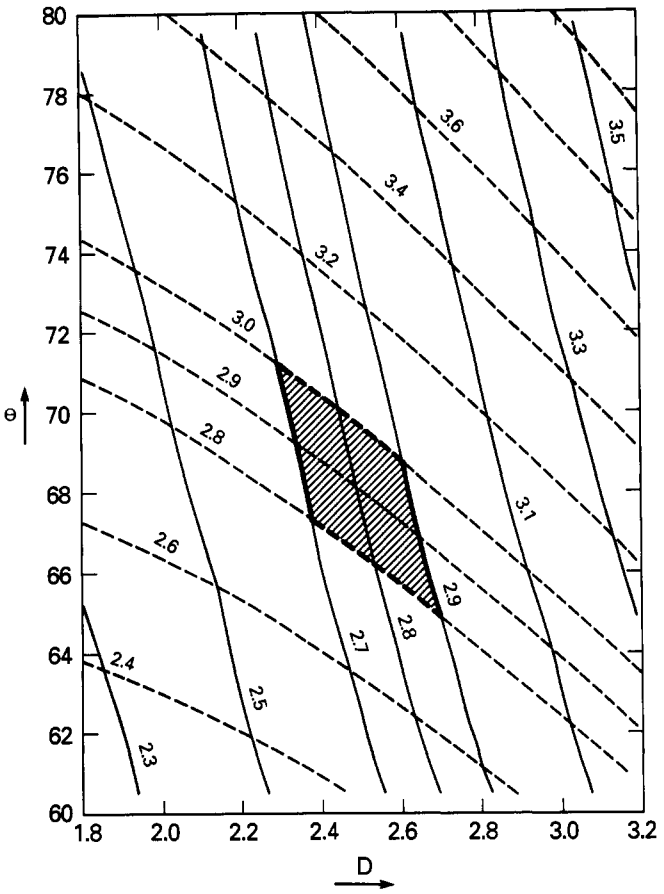

FIGURE 2: Variation of the hydrogen bond lengths N7···N3 (shown in broken lines) and N6···O4 (shown in continuous lines) between the Hoogsteen-paired adenine and thymine strands as a function of displacement (D) of the Watson–Crick base pair from the helix axis and pseudorotational symmetry angle (θ) between the A and T2. The region corresponding to hydrogen bond lengths for N7···N3 between 2.8 and 3.0 Å and N6···O4 between 2.7 and 2.9 Å is shown in the hatched area.

zation procedure, it would be easy to obtain structures that fit the helical parameters. This alternative approach was not pursued since a structure with identical backbone conformations for all three strands was achieved. Energy computations for the final refined structure were done using the CHARMm force field parameters (Brooks et al., 1983) as incorporated in the QUANTA molecular modeling software.

RESULTS AND DISCUSSION

Several structures were generated for the experimentally determined helical parameters of 12 residues per turn (n), 3.26 Å for the axial rise per residue (h), and sugar puckers in the C2′-endo region. Stereochemistry of each structure was critically examined for allowed backbone torsional angles, absence of disallowed nonbonded distances, and proper hydrogen bond lengths and angles in the Watson–Crick and Hoogsteen base pairs. A 2.5-Å displacement (D) of the bases from the helix axis and an associated value for the rotational angle between the Hoogsteen-paired adenine and thymine strands are necessary to obtain a stereochemically satisfactory structure for the $d(T)_n$·$d(A)_n$·$d(T)_n$ triple helix while maintaining symmetry among all three strands. This is accomplished by requiring C2′-endo sugar puckers with bases nearly normal to the helix axis. The structure for $d(T)_n$·$d(A)_n$·$d(T)_n$ for one turn of the helix is shown in Figure 3. The torsional angles and coordinates of this structure are presented in Tables I and II. All torsional angles are well within the allowed ranges for B-form DNA (Kennard & Hunter, 1990).

Table I: Torsional Angles (in degrees) for Monomer in Each Strand of $d(T)_n$·$d(A)_n$·$d(T)_n$ Triple Helix

α (P–O5′)	−36.8
β (O5′–C5′)	168.7
γ (C5′–C4′)	34.6
δ (C4′–C3′)	128.2
ϵ (C3′–O3′)	−170.5
ζ (O3′–P)	−112.6
χ (C1′–N9)	65.8

Table II: Cartesian and Cylindrical Polar Coordinates of Atoms for One Unit of Triple Helix $d(T)_n$·$d(A)_n$·$d(T)_n$[a]

	x	y	z	R	Φ
		Phosphate			
O3′	2.6799	9.8827	2.3216	10.2396	74.83
P	2.9767	10.2929	0.8038	10.7147	73.87
O5′	4.0181	9.1604	0.3648	10.0029	66.32
O1P	1.7167	10.2613	0.0281	10.4040	80.50
O2P	3.5460	11.6585	0.7634	12.1858	73.08
		Sugar			
C5′	4.9741	8.7096	1.3281	10.0299	60.27
C4′	5.6837	7.4635	0.8394	9.3813	52.71
C3′	5.8895	7.4781	−0.6676	9.5188	51.78
C2′	4.9608	6.3620	−1.1274	8.0675	52.05
C1′	4.9351	5.3473	0.0000	7.2766	47.30
O1′	4.9942	6.2221	1.1170	7.9785	51.25
		Thymine			
N1	3.7427	4.4809	0.0000	5.8383	50.13
C2	3.9385	3.1208	0.0000	5.0250	38.39
N3	2.7896	2.3690	0.0000	3.6598	40.34
C4	1.4884	2.8374	0.0000	3.2041	62.32
C5	1.3654	4.2722	0.0000	4.4850	72.28
C6	2.4733	5.0223	0.0000	5.5983	63.78
C5M	−0.0020	4.9361	0.0000	4.9361	90.02
O4	0.5526	2.0438	0.0000	2.1172	74.87
O2	5.0441	2.6099	0.0000	5.6793	27.36
		Adenine			
N9	3.7427	4.4809	0.0000	5.8383	50.13
C8	2.4260	4.8485	0.0000	5.4216	63.42
N7	1.6043	3.8257	0.0000	4.1485	67.25
C6	2.1901	1.3383	0.0000	2.5666	31.43
C5	2.4439	2.7243	0.0000	3.6599	48.10
C4	3.7676	3.1173	0.0000	4.8900	39.60
N3	4.8452	2.3175	0.0000	5.3710	25.56
C2	4.4830	1.0357	0.0000	4.6011	13.01
N1	3.2534	0.5082	0.0000	3.2928	8.88
N6	0.9595	0.8155	0.0000	1.2593	40.36

[a] x, y, z, and R are given in angstroms, and Φ is given in degrees. The basic unit of the triple helix can be generated as follows: Coordinates of the adenosine are retained as they are given. Watson–Crick hydrogen-bonded thymine strand is obtained by a twofold rotation of the given coordinates about the x-axis. Hoogsteen-paired thymine strand is generated by a rotation of 69.5° about the z-axis. Successive units can be generated by adding multiples of 30.0° to Φ and the same multiples of 3.26 Å to z.

The model presented in this paper for the $d(T)_n$·$d(A)_n$·$d(T)_n$ triple helix has a dyad symmetry between the two identical antiparallel T strands. Such a dyad was not recognized earlier (Arnott & Selsing, 1974; Arnott et al., 1976). Moreover, the models for triple helix proposed by Arnott and co-workers do not possess any symmetry among the strands, even though the structure proposed by them was based on stereochemical considerations and not on conventional diffraction data. We have used the same values of the helical parameters, namely, the number of residues per turn (n) and the axial rise per residue (h). The structure of Arnott and co-workers was constructed for C3′-endo sugar with 8.5° tilt of the bases with respect to the helix axis and with the Watson–Crick base pair of the triple helix displaced (D) by more than 3 Å from the helix axis. In contrast, our structure has C2′-endo sugars, 0° tilt of the bases, and a D of 2.5 Å. A structure with C2′-endo sugar pucker and a small tilt (5°) can also be

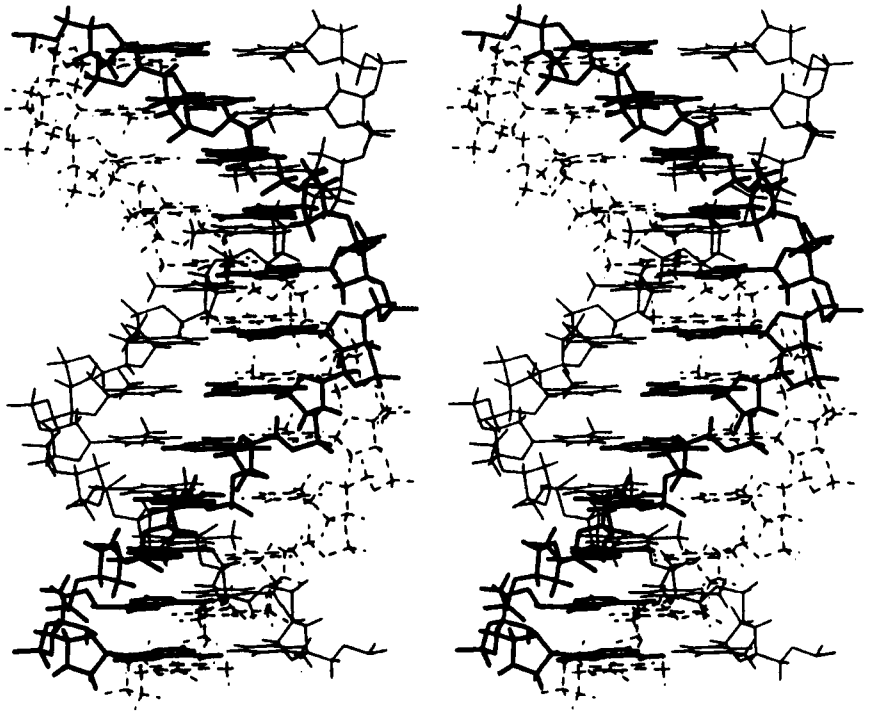

FIGURE 3: Stereo projection of one turn of the triple helix for $d(T)_n \cdot d(A)_n \cdot d(T)_n$. The adenine strand is shown in thick lines, the Watson–Crick base-paired thymine strand is shown in thin lines, and the Hoogsteen base-paired thymine strand is shown in broken lines. The Watson–Crick base pairs are displaced (*D*) by 2.5 Å.

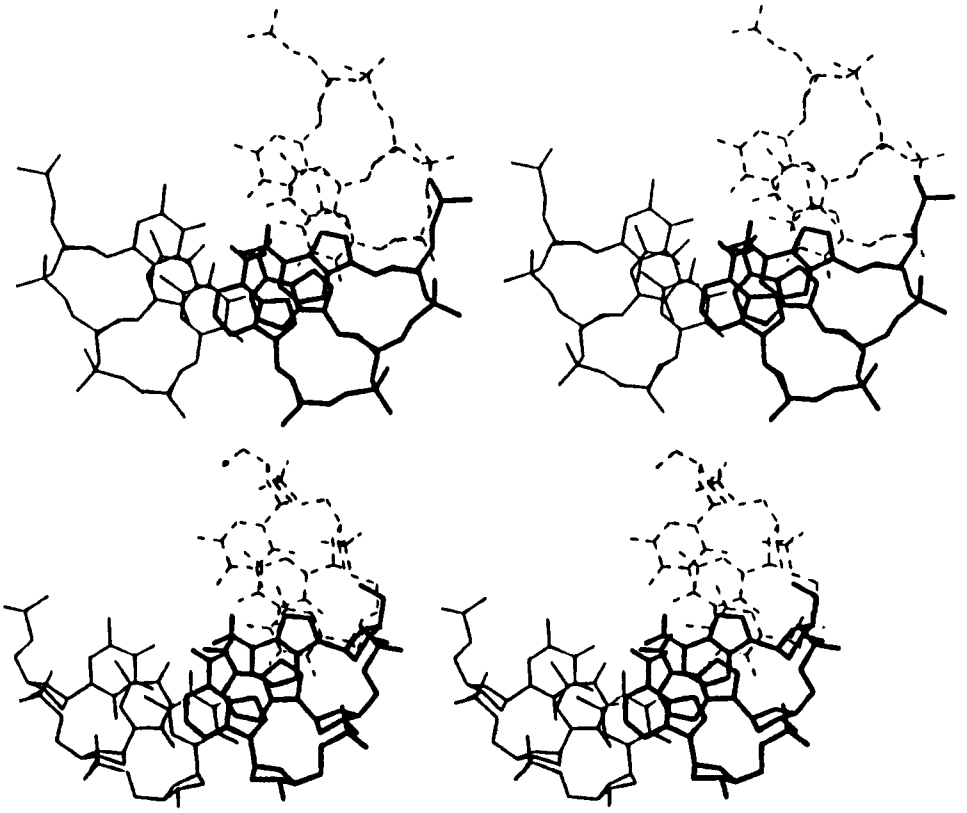

FIGURE 4: Stereoviews of projections down the helix axis of a three base triplet segment for our structure (top) and that of Arnott and co-workers (bottom). Base stacking is similar in the two cases. In our structure (top), the sugar puckers are C2′-endo and similar to a B-form DNA structure. The diameter of our structure is 21.4 Å. The sugar puckers in their structure are C3′-endo, and the structure has an average diameter of 18.5 Å. The adenine strand is shown in thick lines, the Watson–Crick base-paired thymine strand is shown in thin lines and the Hoogsteen base-paired thymine strand is shown in broken lines.

obtained (Howard et al., 1992).

Stereodiagrams for two projections of a three base pair segment of both the present structure and the structure of Arnott and co-workers are given in Figures 4 and 5. The stacking patterns of the two structures are very similar. In both structures, the glycosidic torsions are anti and the torsional angles about the bonds C4′–C5′, C5′–O5′, and O3′–C3′ are gauche$^+$, trans, and trans, respectively. The phosphodiester conformation in our structure is (trans, gauche$^-$) and is coupled to the C2′-endo sugar. The phosphodiester conformation in the structure of Arnott and co-workers is (gauche$^-$,gauche$^-$) and is associated with the C3′-endo sugar pucker. The

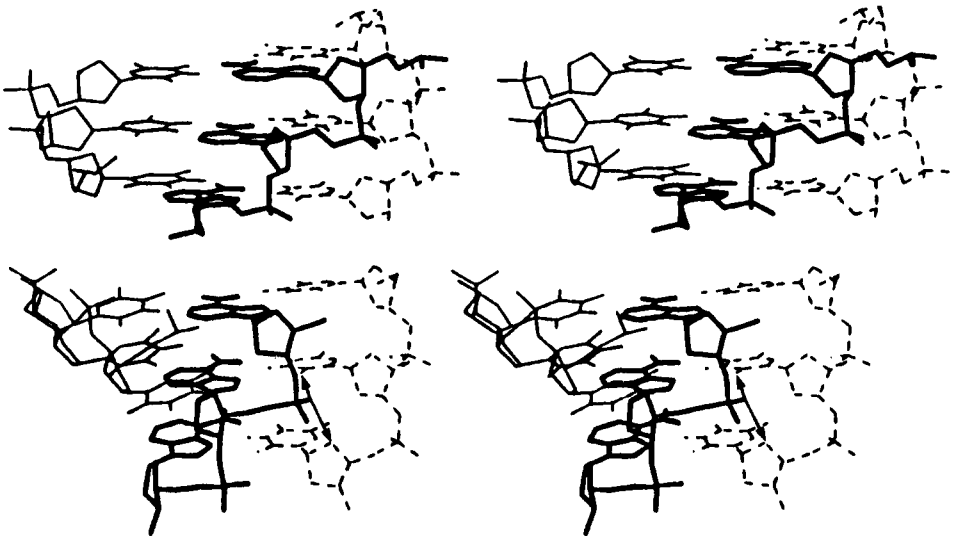

FIGURE 5: Stereoviews from the minor groove of our structure (top) and that of Arnott and co-workers (bottom). The adenine strand is shown in thick lines, the Watson–Crick base-paired thymine strand is shown in thin lines, and the Hoogsteen base-paired thymine strand is shown in broken lines. The short nonbonded distances in their structure (bottom) are shown by arrows. The pendant oxygen of the phosphate in the adenine strand is at a distance of 2.32 Å from the O2 of the base on the Hoogsteen-paired thymine strand and at a distance of 2.67 Å from the sugar O4 on the same strand. The oxygen is sandwiched between these two atoms. There are no such short distances in our structure (top), and the corresponding distances in our structure are 5.20 and 3.63 Å, respectively.

coordinates of the phosphorous atom are generally used to define the diameter of the helix. Accordingly, the diameter of the triple helix in our structure is 21.4 Å. Since the corresponding phosphorus atoms of the three chains in their structure are not equivalent, the average diameter of their structure is 18.5 Å. The increase in diameter of the triple helix does not pose any difficulty in the packing of these molecules in the unit cell obtained by Arnott and Selsing (1974) from fiber diffraction data. In their structure for $d(T)_n$·$d(A)_n$·$d(T)_n$, the pendant oxygen of the phosphate in the adenine strand is too close to the Hoogsteen base-paired thymine strand: a much too short distance of 2.32 Å from the O2 of the thymine and a short distance of 2.67 Å from the sugar O4′ (shown by arrows in Figure 5, bottom). The corresponding distances in our structure are 5.20 and 3.63 Å (see Figures 4 and 5). The conformational aspects of the model described above have contributed to the disallowed nonbonded distances in their model. The sugar–phosphate backbone is farther away from the base pairs in our structure. This feature and the larger diameter of our structure is a consequence of C2′-endo sugars and a shorter displacement of 2.5 Å of the Watson–Crick base pair from the helix axis.

The usual procedure for deciding between two structures would be the comparison of their Fourier transforms with that obtained from X-ray diffraction of the fiber. As Arnott et al. have noted (1976), however, there is a scarcity of quantifiable diffraction data caused by limited ordering in the samples studied. Since no such comparison could be made, we have not computed the transforms, but as an alternative we have calculated the energies of the two structures.

A comparison of the calculated interstrand nonbonded and electrostatic energies among the three strands and the total energy per turn of the triple helix are given in Table III. The total energy per turn of the helix for our structure is 255 kcal lower than that for the structure of Arnott and co-workers. As would be anticipated, there is negligible interaction energy between the two thymine strands. The major differences is in the nonbonded interaction energy between the Hoogsteen-paired A and T strands, which is more favorable by 123 kcal per turn of the helix for our structure. The electrostatic energy between these two strands for our structure is lower by 67 kcal. The energy terms for the Watson–Crick-paired adenine and thymine strands are also lower for our structure. Although there are uncertainties associated with energy calculations, such large differences are significant.

Table III: Interstrand Nonbonded (Nbond) and Electrostatic (elect) Energies (in kcal) for One Turn of Triple Helix $d(T)_n$·$d(A)_n$·$d(T)_n$ Proposed in This Paper and Corresponding Values for Structure Proposed by Arnott and Co-workers (Arnott et al., 1976)[a]

	Values for the Structure in This Paper			
	T1-A-T2	T1-A	A-T2	T1-T2
Nbond	−59	0	−54	−4
elect	−205	−173	−49	16
sum	−264	−173	−103	12
	Values for the Structure (Arnott et al., 1976)			
	T1-A-T2	T1-A	A-T2	T1-T2
Nbond	114	46	69	−1
elect	−122	−156	18	16
sum	−8	−110	87	15

[a] T1 and A refer to the Watson–Crick hydrogen-bonded thymine and adenine strands, and T2 refers to Hoogsteen hydrogen-bonded thymine strand.

Our model for the $d(T)_n$·$d(A)_n$·$d(T)_n$ triple helix is consistent with the results of solution and solid-state film studies. As discussed in detail in an earlier paper (Howard et al., 1992), solution infrared spectra of DNA duplexes have bands in the range of 830–840 (medium, m) and 970 cm^{-1} (strong, s), characteristic of B-form structure and C2′-endo sugar pucker. Structures of either RNA or A-form DNA double helices, in contrast, have characteristic bands near 810 (m) and 865 cm^{-1} (s). The triple helix $d(T)_n$·$d(A)_n$·$d(T)_n$ has bands at 840 and 973 cm^{-1} characteristic of a B-form helix and a C2′-endo sugar pucker and none near 810 and 865 cm^{-1}. Furthermore, the IR spectrum of the solid-state film of the triple helix $d(T)_n$·$d(A)_n$·$d(T)_n$ (Liquier et al., 1991) also has a characteristic band at 840 cm^{-1}, confirming the B-form helix with C2′-endo sugar pucker for the triple helix in the solid-state also. The stacking pattern of the base pairs, the sugar puckers, the phosphodiester conformation of the backbone, and the displacement of the Watson–Crick base pairs from the helix axis all resemble the double helical B-form DNA structure. It is thus not surprising that UV and gel electrophoresis studies show binding of distamycin [a drug that binds preferentially to the minor groove of double helical B-form DNA and not

to A-form DNA, cf. Zimmer and Wahnert (1986)] to the triple helix $d(T)_n \cdot d(A)_n \cdot d(T)_n$. CD spectroscopy and gel electrophoresis demonstrate that distamycin does not displace the third strand and that DNA retains its triple helical structure in the bound complex (Howard et al., 1992).

An NMR study of an intermolecular triple helix formed by $d(TC)_4 \cdot d(GA)_4 \cdot d(TC)_4$ (Rajagopal & Feigon, 1989a,b) has been interpreted in terms of the occurrence of C2′-endo and significant deviation from C3′-endo puckers for the sugars. In a more recent analysis of an intramolecular triple helix, Macaya et al. (1992) have concluded that a great majority of the sugars of all three strands have predominantly C2′-endo sugar puckers.

Since the model proposed by Arnott and co-workers has been the only available structure with atomic detail for a DNA triple helix, this structure has been used as the starting conformation for several structural calculations and simulations. As we would have anticipated, the short nonbonded distances mentioned above in the structure of Arnott and co-workers resulted in repuckering of sugars in the course of simulation. Molecular dynamics calculations for a canonical triple helix (Laughton & Neidle, 1992), and $d(T)_{10} \cdot d(A)_{10} \cdot d(T)_{10}$ (Hausheer et al., 1990), both starting from the coordinates of Arnott and co-workers (Arnott et al., 1976), resulted in considerable deviations from the starting conformation, including sugar pucker variations that ranged from the starting C3′-endo to the C2′-endo region. While neither of these studies gave an explanation for the results of their molecular dynamics simulations, it is quite likely that the overly short nonbonded distances in the starting structures contributed to their final results. Another molecular modeling study also reported repuckering of sugars to O1′-endo from the starting C3′-endo for the purine strand (Sun et al., 1991).

One of the above simulations (Laughton & Neidle, 1992) reported reorientations of the phosphate groups of the adenine strands. This phosphate group is too close to the base and sugar atoms of the Hoogsteen-paired thymine strand in their starting structure (see Figures 4 and 5, bottom, and marked by arrows in Figure 5, bottom). The orientation of the phosphate group with respect to an average helical axis is a backbone feature that characterizes sugar pucker (Sasisekharan et al., 1983). Thus, the reorientation of the phosphates provides further support for departure from the starting C3′-endo sugar pucker. Structural transitions, such as repuckering of sugars, in this case from C3′-endo to a mixture of C3′-endo and C2′-endo sugar puckers, could manifest themselves in several local and global structural variations. Although their 20-ps simulation on the system as large as $d(TC)_5 \cdot d(GA)_5 \cdot d(C^+T)_5$ with 3530 water molecules and 22 counterions is too short to reach definitive conclusions about the average structural parameters for the triple helix, the tendency for significant deviations from the starting C3′-endo sugar conformation is clear. The conformations of the individual nucleotides in the three strands are different during molecular dynamics simulation, and the helical constraints of rise per residue and the number of residues per turn of the helix observed in the fiber diffraction were not maintained during these simulations. If these helical parameters were imposed, they would serve only to facilitate changes from the starting structure. However, we would expect that molecular simulation studies carried out with our structure as the starting conformation would retain sugar puckers in the C2′-endo region during simulation since only a very limited region in the conformational space is allowed for the triple helix (see Figure 2).

CONCLUSION

The structure for the $d(T)_n \cdot d(A)_n \cdot d(T)_n$ triple helix presented in this paper is fundamentally different from the structure proposed by Arnott and co-workers and is energetically much more favorable than their structure. While Arnott and co-workers suggested an A-form structure with C3′-endo sugar puckers, our structure belongs to B-form DNA and has C2′-endo sugar puckers. The structure has 12 residues per turn of the helix and a pitch of 3.26 Å, as obtained from fiber diffraction studies (Arnott & Selsing, 1974). Further, all three strands of the present structure have identical sugar–phosphate backbone conformations, unlike previously proposed structures for any triple helix. In the present structure, there is a pseudorotational symmetry between the Hoogsteen-paired adenine and thymine strands, and this is in addition to the pseudodyad that relates the Watson–Crick-paired thymine and adenine strands. As a consequence, the two thymine strands are related by a dyad perpendicular to the helix axis. Detailed modeling studies indicate that only a very small region of conformational space becomes available for any triple helix when all three strands are identical and symmetry related. We conclude that the triple helix structure for $d(T)_n \cdot d(A)_n \cdot d(T)_n$ is similar to the B-form DNA and has C2′-endo sugar pucker, displacement of 2.5 Å for the bases from the helix axis, and bases nearly normal to the helical axis. Further details of the structure can only be established by refinement against experimental data, in particular X-ray diffraction by fibers or single crystals. Deviations from the model may well occur, but because of the triple helical constraints discussed above, we would expect them to be small.

ACKNOWLEDGMENT

We would like to thank Dr. R. L. Jernigan of the National Cancer Institute and Drs. D. R. Davies and S. B. Zimmerman of the National Institute of Diabetes and Digestive and Kidney Diseases of NIH for discussions and critical reading of the manuscript.

APPENDIX

To show that a rotation by 180° about the x-axis followed by rotation by an angle θ about the z-axis is equivalent to a rotation by 180° about an axis which lies in the xy plane at an angle $[180–(\theta/2)]°$ from the x-axis (see Figure 1).

The general form of matrix (Jeffreys & Jeffreys, 1956) for rotation by an angle α about a unit vector

$$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

is

$$\begin{bmatrix} \cos\alpha + V_x^2(1-\cos\alpha) & V_xV_y(1-\cos\alpha) - V_z\sin\alpha & V_zV_x(1-\cos\alpha) + V_y\sin\alpha \\ V_xV_y(1-\cos\alpha) + V_z\sin\alpha & \cos\alpha + V_y^2(1-\cos\alpha) & V_yV_z(1-\cos\alpha) - V_x\sin\alpha \\ V_zV_x(1-\cos\alpha) - V_y\sin\alpha & V_yV_z(1-\cos\alpha) + V_x\sin\alpha & \cos\alpha + V_z^2(1-\cos\alpha) \end{bmatrix} \quad (1)$$

For a rotation by $\alpha = 180°$, the above matrix becomes

$$\begin{bmatrix} -1 + 2V_x^2 & 2V_xV_y & 2V_zV_x \\ 2V_xV_y & -1 + 2V_y^2 & 2V_yV_z \\ 2V_zV_x & 2V_yV_z & -1 + 2V_z^2 \end{bmatrix} \quad (2)$$

Matrix *M1* for rotation by 180° about the *x*-axis is given by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Matrix *M2* for rotation by an angle θ about *z*-axis is given by

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$M1 \cdot M2 = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ -\sin\theta & -\cos\theta & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (3)$$

To show that $M1 \cdot M2 = M3$, the matrix *M3* is a rotation by an angle 180° about an axis which lies in the *xy* plane at an angle $[180 - (\theta/2)]°$ from the *x*-axis. The column vector for this axis is

$$\begin{bmatrix} \cos[180 - (\theta/2)] \\ \sin[180 - (\theta/2)] \\ 0 \end{bmatrix}$$

Substituting the above values of V_x, V_y, and V_z in the matrix 2, one gets the following matrix

$$\begin{bmatrix} -1 + 2\cos^2[180 - (\theta/2)] & 2\cos[180 - (\theta/2)]\sin[180 - (\theta/2)] & 0 \\ 2\cos[180 - (\theta/2)]\sin[180 - (\theta/2)] & -1 + 2\sin^2[180 - (\theta/2)] & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (4)$$

This becomes

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ -\sin\theta & -\cos\theta & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

which is the same as the right-hand side of eq 3. Hence, $M1 \cdot M2 = M3$.

$\theta = 180°$ is a special case of this general formulation. In this situation, simultaneous presence of dyads about two mutually orthogonal axes, namely, *x*- and *z*, results in a dyad about an axis perpendicular to both *x* and *z*, namely, *Y*.

REFERENCES

Arnott, S., & Selsing, E. (1974) *J. Mol. Biol. 88*, 509–521.

Arnott, S., Bond, P. J., Selsing, E., & Smith, P. J. C. (1976) *Nucleic Acids Res. 3*, 2459–2470.

Baran, N., Lapidot, A., & Manor, H. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 507–511.

Birg, F., Praseuth, D., Zenal, A., Thoung, N., Asseline, U., LeDoan, T., & Helene, C. (1990) *Nucleic Acids Res. 18*, 2901–2908.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Cooney, M., Czernoszewicz, G., Postel, E., Flint, S. J., & Hogan, M. E. (1988) *Science 245*, 456–459.

Cristophe, D., Cabrer, B., Bacolla, A., Targovnik, H., Pohl, P., & Vassart, G. (1985) *Nucleic Acids Res. 13*, 5127–5144.

Felsenfeld, G., Davies, D. R., & Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023–2024.

Hanvey, J. C., Shimizu, M., & Wells, R. D. (1988) *J. Mol. Biol. 85*, 6292–6296.

Hattori, M., Ikehara, M., & Miles, H. T. (1974) *Biochemistry 13*, 2754–2761.

Hausheer, F. H., Singh, U. C., Saxe, J. D., Colvin, O. M., & T'so, P. O. P. (1990) *Anti-Cancer Drug Res. 5*, 159–167.

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.

Hoogsteen, K. (1963) *Acta Crystallogr. 16*, 907–916.

Howard, F. B., Frazier, J., Lipsett, M. N., & Miles, H. T. (1964) *Biochem. Biophys. Res. Commun. 17*, 93–97.

Howard, F. B., Miles, H. T., Liu, K., Frazier, J., Raghunathan, G., & Sasisekharan, V. (1992) *Biochemistry 31*, 10671–10677.

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.

Ikehara, M., Hattori, M., & Fikui, T. (1972) *Eur. J. Biochem. 31*, 329–334.

Ishikawa, F., Frazier, J., Howard, F. B., & Miles, H. T. (1972) *J. Mol. Biol. 70*, 475–490.

Jeffreys, H., & Jeffreys, B. S. (1956) in *Methods of Mathematical Physics*, p 122, Cambridge University Press, Cambridge.

Johnston, B. H. (1988) *Science 241*, 1800–1804.

Kennard, O., & Hunter, W. N. (1990) in *Numerical Data and functions in Science and Technology* (Saenger, W., Ed.) Vol. VII/1a, Subvolume a, pp 255–359, Springer Verlag, Berlin.

Kohwi, Y., & Kohwi-Shigematsu, T. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3781–3785.
Laughton, C. A., & Neidle, S. (1992) *J. Mol. Biol. 223*, 519–529.
Lee, J. S., Johnson, D. A., & Morgan, A. R. (1979) *Nucleic Acids Res. 6*, 3073–3091.
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. F., 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.
Maher, L. J., III, Dervan, P. B., & Wood, B. (1990) *Biochemistry 29*, 8820–8826.
McKie, D., & McKie, C. (1986) in *Essentials of Crystallography*, Blackewell Scientific Publications, Oxford.
Morgan, A. R., & Wells, R. D. (1968) *J. Mol. Biol. 37*, 63–80.
Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645–650.
Orson, F. M., Thomas, D., McShan, W. M., Kessler, D. J., & Hogan, M. E. (1991) *Nucleic Acids Res. 19*, 3435–3441.
Postel, E., Flint, S. J., Kessler, D. J., & Hogan, M. E. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 8227–8231.
Rajagopal, P., & Feigon, J. (1989a) *Nature 339*, 637–640.
Rajagopal, P., & Feigon, J. (1989b) *Biochemistry 28*, 7859–7870.
Sasisekharan, V., Bansal, M., & Gupta, G. (1983) in *Nucleic Acids: Vectors of Life* (Pullman, B., & Jornter, J., Eds.) pp 101–111, D. Reidel Publishing Co., Dordrecht.
Smith, P. J. C., & Arnott, S. (1978) *Acta Crystallogr. A34*, 3–11.
Sun, J., Mergny, J., Lavery, R., Montenay-Garestier, T., & Helene, C. (1991) *J. Biomol. Struct. Dyn. 9*, 411–423.
Zhurkin, V. B., Raghunathan, G., Camerini-Otero, R. D., & Jernigan, R. L. (1991) *J. Biomol. Struct. Dyn. 8*, a268.
Zimmer, Ch., & Wahnert, U. (1986) *Prog. Biophys. Mol. Biol. 47*, 31–112.