Refinement of the Solution Structure of a Branched DNA Three-Way Junction

Igor V. Ouporov and Neocles B. Leontis
Chemistry Department, Bowling Green State University, Bowling Green, Ohio 43403-0213 USA

ABSTRACT We have refined the structure of the DNA Three-Way Junction complex, TWJ-TC, described in the companion paper by quantitative analysis of two 2D NOESY spectra (mixing times 60 and 200 ms) obtained in D_2O solution. NOESY crosspeak intensities extracted from these spectra were used in two kinds of refinement procedure: 1) distance-restrained energy minimization (EM) and molecular dynamics (MD) and 2) full relaxation matrix back calculation refinement. The global geometry of the refined model is very similar to that of a published, preliminary model (Leontis, 1993). Two of the helical arms of the junction are stacked. These are Helix 1, defined by basepairs S1-G1/S3-C12 through S1-C5/S3-G8 and Helix 2, which comprises basepairs S1-C6/S2-G5 through S1-G10/S2-G1. The third helical arm (Helix 3), comprised of basepairs S2-C6/S3-G5 through S2-C10/S3-G1 extends almost perpendicularly from the axis defined by Helices 1 and 2. The bases S1-C5 and S1-C6 of Strand 1 are continuously stacked across the junction region. The conformation of this strand is close to that of B-form DNA along its entire length, including the S1-C5 to S1-C6 dinucleotide step at the junction. The two unpaired bases S3-T6 and S3-C7 lie outside of the junction along the minor groove of Helix 1 and largely exposed to solvent. Analysis of the refined structure reveals that the glycosidic bond of S3-T6 exists in the syn conformation, allowing the methyl group of this residue to contact the hydrophobic surface of the minor groove of Helix 1, at S3-G11. The helical parameters of the three helical arms of the structure exhibit only weak deviations from typical values for right-handed B-form DNA. Unusual dihedral angles are only observed for the sugar-phosphate backbone joining the "hinge" residues, S2-G5 and S2-C6, and S3-G5 through S3-G8. The glycosidic bond of S3-G8 also lies within the syn range, allowing favorable Watson-Crick base-pairing interactions with S1-C5. The stability of this structure was checked in 39 ps molecular dynamic simulation at 330 K in water. The structure of TWJ-TC retained the geometrical features mentioned above at the end of the simulation period. The final R(1/6)-factor of the refined structure is 5%.

INTRODUCTION

Three-way junctions (TWJ) are the simplest and most commonly occurring branched structures in biologically active, single-stranded nucleic acids. TWJ lacking unpaired bases appear to be conformationally flexible, according to studies employing native gel electrophoresis (Duckett and Lilley, 1990), chemical and enzymatic probing (Guo et al., 1990), and directed ligation (Ma et al., 1986). They do not assemble to form homogeneous complexes amenable to NMR study. Using optically monitored, thermal denaturation, and native gel electrophoresis, we have shown that two unpaired nucleotides inserted in one strand at the junction site suffice to stabilize TWJ (Leontis et al., 1991). In the companion paper, we demonstrated that TWJ complexes with two unpaired pyrimidines form homogeneous structures in solution that give rise to ^1H NMR spectra that can be assigned and analyzed. In previous work, we presented a qualitative model for the three-dimensional structure of the TWJ complex TWJ-TT (Leontis et al., 1993). This model was derived from composite NMR and chemical probing data on TWJ-TT and the closely related junction TWJ-TC. The sequences and secondary structures of these complexes are shown schematically in Fig. 1 of the companion paper. Because the TWJ-TC sample fortuitously gave rise to better resolved spectra exhibiting sharper line widths, we chose it for quantitative analysis. But it should be noted that all available data indicate that the structures of the two junctions are very similar.

The global features of the preliminary, qualitative model (referred to henceforth as the "initial structure" for the purposes of the refinement reported here) are preserved in the final structure. These features include: the stacking of two of the helices, the extra-helical disposition of the unpaired bases, and the folding of the phospho-diester backbone in a fairly tight hairpin-like loop at the unpaired bases. NMR studies of another DNA TWJ complex, which also comprises two unpaired pyrimidines, were recently reported by Rosen and Patel (1993). The similarities and differences between that structure and the present one were compared in the companion paper. For the present work, we used the well established methods of full matrix relaxation refinement to refine our preliminary model of the solution structure of TWJ-TC. We then carried out an analysis of the conformations of the residues and of the global geometry of the refined structure.

MATERIALS AND METHODS

DNA oligonucleotides

All oligonucleotides used in NMR studies were purchased in purified form from the Midland Reagent Company (Midland, TX). The sequences are: S1,

Received for publication 2 August 1994 and in final form 20 October 1994.

Address reprint requests to Dr. Neocles B. Leontis, Department of Chemistry, Bowling Green State University, Bowling Green, OH 43403. Tel.: 419-372-8663; Fax: 419-372-9809; Email: neocles@rosalind.bgsu.edu.

Abbreviations used: NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; TWJ, three-way junction; EM, energy minimization; MD, molecular dynamics.

0006-3495/95/01/266/09 $2.00

5′-GCT GCC ACC G; S2, 5′-CGG TGC GTC C; and S3TC, 5′-GGA CGT CGC AGC. The preparation of NMR samples is described in the companion paper.

NMR spectroscopy

Proton NMR spectra were acquired on the Varian 600 MHz Unity Plus Spectrometer at Akron University (Akron, OH). Two matched 2D NOESY spectra were acquired in phase-sensitive mode at mixing times of 60 and 200 ms. The spectra were acquired in D_2O solution with water presaturation on the same day and temperature (27°C). The spectral width was 6000 Hz in F1 and F2, the data size was 2048 points in F2 and 512 increments in F1, acquisition time was set equal 0.171 s, relaxation delay was 3 s, and 16 scans were averaged per t1 value. The FID sets were processed using Felix 2.10 software (Biosym, Inc., San Diego, CA). The FIDs were apodized using the sine-squared bell, 90°-shifted function in both the F1 and F2 dimensions to keep the integrated peak volumes unchanged. Chemical shifts of nonexchangeable and exchangeable protons for TWJ-TC are presented in Table 2 of the companion paper. Volumes of assigned peaks were integrated using the Felix 2.10 Peak Volume features and were input as NOE intensities in further calculations. The number of measured NOE intensities per residue is presented in Fig. 1. Overall, 398 NOE peak intensities were measured from the 200 ms spectrum and 253 from the 60 ms spectrum. These integrated intensities were used for restricted molecular dynamics calculations and matrix relaxation back calculation refinement. The X-PLOR 3.1 software package (Brünger, 1993) was employed for structure refinement.

Initial distance restraints

We chose the 60 ms mixing time spectrum for extracting approximate distance restraints between assigned protons. We calibrated mean distances between protons using the fixed distance between cytosine H5 and H6 protons. Because of spin diffusion effects, even for the relatively short 60 ms mixing time, we chose to omit all intra-sugar crosspeaks and only to take into account intra-residue and inter-residue aromatic-sugar crosspeaks as well as inter-residue aromatic-aromatic crosspeaks. NOEs involving exchangeable protons, obtained from separate experiments carried out in H_2O solution, were handled in a qualitative manner (see below). Overall, 174 restraints were extracted from the 60 ms NOESY spectrum. The square-well function was chosen as the potential function for the distance restraint term in calculating the global potential energy of the system. The scale factor for this potential was set equal to 25 kcal/mol Å^2. The flat-well width for each proton pair was calculated as $D = D_- + D_+$, where $D_- = (0.1)\, d_{exp}$ and $D_+ = (0.2)\, d_{exp}$, and d_{exp} is the distance between the relevant protons as estimated from the 60 ms NOESY spectrum using the calibration procedure. This choice corresponds to the quality of our spectra and to the accuracy of the calibration procedure.

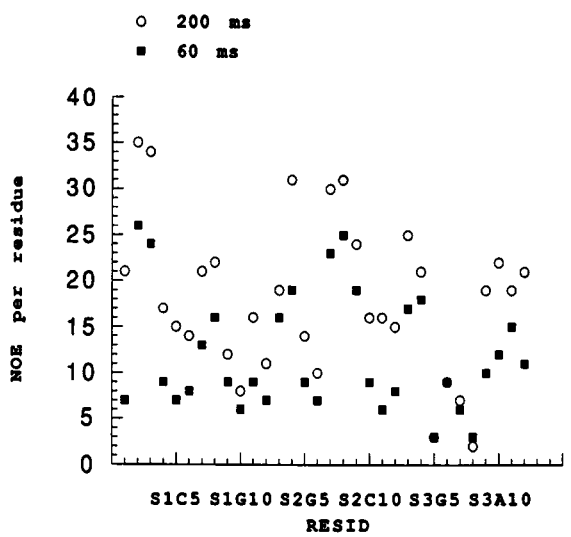

FIGURE 1 The number of measured NOE integrals per residue for the two 2D NOESY spectra analyzed in this work (mixing times 60 and 200 ms).

Further restraints

To maintain the proper base-pairing, base-stacking, and hydrogen-bonding in the TWJ complex, an additional set of restraints was applied to reinforce the hydrogen- bonding energy of adjacent bases and to model NOE peaks observed in NOESY spectra acquired in H_2O among exchangeable imino and amino protons. The overall number of such qualitative restraints was 70, of which 42 mimic the hydrogen- bonding in the complex and 28 model the NOEs observed between exchangeable protons. To model these restraints, we also used the square-well potential with the same force constant 25 kcal/mol Å^2. The parameters for modeling hydrogen bond constraints in Watson-Crick basepairs were adopted from published work (Mujeeb et al., 1993): for G-C basepairs the distances were set to 2.81–3.01 Å (G-O6 to C-N4), 2.85–3.05 Å (G-N1 to C-N3), and 2.76–2.96 Å (G-N2 to C-O_2); for A-T basepairs the bounds were set to 2.72–2.92 Å (A-N1 to T-N3) and 2.85–3.05 Å (A-N6 to T-O4). The bounds on inter-proton distances corresponding to NOE crosspeaks observed in 2D NOE spectra in H_2O (Leontis et al., 1993) were set as follows: 1.8–3.0 Å for observed intra-basepair imino-amino and imino-adenosine H2 crosspeaks; 2.0–4.0 Å for observed imino-imino crosspeaks involving adjacent base pairs. These distance constraints were used in all calculations presented in this work. No direct constraints on dihedral angles were used.

Relaxation matrix refinement

The back calculation relaxation matrix refinement of the TWJ structure was carried out using the tools provided in the X-PLOR 3.1 package. We made two series of otherwise identical calculations using each of the two different force fields provided with X-PLOR 3.1 (Brünger, 1993), namely PARALLHDG.DNA, the special force field provided for NMR structure refinement of DNA (Kuszewski et al., 1992), and CHARMM 22 (Brünger, 1993). Nonbonded interactions were modeled during simulation using the switched van der Waals function in combination with the switched electrostatic function with radius dependent dielectric constant. The value of the cutoff radius at which the switching function forces the nonbonded energy to zero was set equal to 11 Å. The characteristic of refinement quality was the R-factor with $n = 1/6$. The uniform error parameter (D) was set to 0.1 for the 60 ms spectrum and to 0.075 for the 200 ms spectrum. These values were chosen to correspond to the ratio of signal-to-noise in our spectra. The constant Weight = 1000 was used in the matrix relaxation refinement (Brünger, 1993). We also used distance restraints that mimic Watson-Crick hydrogen-bonding between paired nucleotides and distance restraints between exchangeable protons that were observed in 2D NOE spectra in water (Leontis et al., 1993), as discussed above. When using the PARALLHDG.DNA force field, we found it necessary to apply the additional planarity constraints to maintain the mutual base planarity of the terminal basepairs in each helix. In calculations carried out without such constraints, we observed large propeller and buckling motions of terminal basepairs during energy minimization and molecular dynamics. When using the CHARMM22 force field, planarity constraints were imposed on each individual aromatic base, forcing each heavy atom that belongs to the given base as well as C1′ carbon of the corresponding sugar ring to lie in a plane. This avoids puckering that is otherwise seen during simulations. It was not necessary to apply the additional planarity constraints to terminal adjacent base pairs because in our simulation with the CHARMM 22 force field the propeller and buckle parameters did not deviate as strongly as in calculations with the PARALLHDG.DNA force field. The overall rotational relaxation time was set to 8 ns. This value was found empirically to produce the lowest R-factor for the initial structure. The initial model for TWJ-TT (Leontis et al., 1993) was modified using QUANTA/CHARMm 3.3

software (Molecular Simulation Inc., Waltham, MA) to replace S3-T7 and S3-G12 with cytosines and S1-C1 by guanine to give the correct sequence for the TWJ-TC junction.

The refinement protocol

Our initial structure (Leontis et al., 1993) consists of the central bulged bases of Strand 3 and the three double-helical arms. The arguments in favor of this structure are discussed in detail in the companion paper. The pattern of NOE connectivities and peak intensities demonstrates that the helical arms exist in conformations very similar to B-form DNA. Thus, the main goal of our refinement was to elucidate and investigate the structure of the bulged region and the deviations from B-form of the basepairs flanking the junction region. From the initial structure, one observed that the residues in the junction region have sufficient freedom to move to achieve more favorable conformations consistent with observed NOE intensities. Therefore, it was sufficient to conduct "gentle" refinement to satisfy all NOE requirements. For this reason, we omitted the initial steps typically used in refinement procedures, namely, distance geometry calculation and high temperature simulated annealing. We limited our calculations to the final step: molecular dynamics at room temperature and energy minimization including the relaxation matrix term in the potential function for the system. This approach is valid when one has a reasonable initial structure that is consistent with NMR data and other independent evidence that is sensitive to the general geometry of the system under consideration. A similar refinement method was recently applied to study the structural characteristics of an intramolecular DNA triplex (Macaya et al., 1992). In support of this choice, we refer to the fact that for both the 200 and the 60 ms spectra, the R-factor for the initial structure was calculated to be 10%. These values indicate that the initial structure is compatible with NOE restraints on interproton distances.

For distance-restraints refinement, the initial structure was subjected to 3000 steps of restrained energy minimization (EM) followed by restrained molecular dynamics (MD) at 300 K for 5 ps. The temperature-coupling method was used for temperature control (Berendsen et al., 1984). The average structure for the last picosecond was subjected to another 3000 steps of restrained EM. During this calculation, the value of the distance restraints term in the potential energy of the system was decreased by a factor of five and the value of the mean-square violation of distance restraints decreased from 0.37 to 0.14 Å. The final and starting structures are similar with respect to the global geometry of the complex. The main changes in the course of this calculation occurred in the junction region of our complex, but other parts of the complex remained the same, having the typical geometry of B-form DNA.

The relaxation matrix back calculation protocol was applied to refinement using NOE volumes from the 60 ms spectrum and from the 200 ms spectrum separately. The initial structure was subjected to 500 steps of energy minimization, then 3 ps of molecular dynamics at 300 K with temperature-coupling (Berendsen et al., 1984). The average structure for the last picosecond was again energy-minimized (300 steps). In Table 1 we present the energy characteristics of the structure of the TWJ complex at different stages of calculation, the values of RMSD from the initial structure, the R-factor values, and the values of the glycosidic angle of the S3-T6 nucleotide. As mentioned above, the value of the R-factor was equal to 10% for the initial structure. The R-factor decreased by a factor of 2 during the first step of matrix relaxation back calculation, 500 steps of EM. Further MD calculations slightly decreased the R-factor and lowered the potential energy of system. Such small decreases were accompanied by undesirable changes in the conformation of the arms, namely, fraying of the terminal basepairs and partial unwinding of the helixes. We noted above that these distortions of the helical geometry were more prominent when using the PARALLHDG.DNA force field than with the CHARMM 22 force field. This indicates that for the final steps of structural refinement the CHARMM 22 force field is preferable to PARALLHDG.DNA. The main changes in the conformation of the bulged region of the junction occurred during the first 500 steps of EM. The base of S3-T6 settles into the syn glycosidic conformation range during EM and remains there during MD. The locations of the bases of nucleotides S3-T6 and S3-C7 in the minor groove of Helix 1 changed little during the MD calculation. The structures obtained after 500 steps of matrix relaxation back calculation EM from the 60 ms and from the 200 ms spectra are very similar; the RMSD between the structures is equal to 0.74 Å. Therefore, we have decided to present as the "final refined structure" the one resulting from the initial 500 steps of EM obtained in calculation with the CHARMM 22 force field using intensities from the 200 ms NOESY spectrum. In Fig. 2 we show the final refined structures (from the 60 and 200 ms data) superimposed upon the initial structure, as defined at the beginning of this paper. From this picture, one can see that the backbone conformations of the arms in the structures are very similar except for the conformations of residues in the terminal regions of the arms. A stereoview of the final structure (from 200 ms data) is presented in Fig. 3.

RESULTS AND DISCUSSION

The distance constraints and NOE intensities for protons belonging to the junction region that were used in the refinement are as follows. 1) The NOE cross-peak between the imino protons of S3-G8 and S2-G5 observed in the 2D NOESY spectra taken in H_2O (Leontis et al., 1993). The distance used was 3 ± 1 Å. 2) From the same spectra, the NOE crosspeak between the imino proton of S3-G8 and the amino proton of S1-C5. This peak was used to constrain the distance between S3-G8 H1 and S1-C5 N3 to 2.4 ± 0.6

TABLE 1 The results of calculations

Force Field	Spectrum	Structure	RMSD (A)	E-rel	E-total	E-potential	R-factor (%)	Khi (S3T6)
CHARMM	60 ms	Initial		1030	751	-529.1	10	90
		500 st. EM	0.92	210	-1463	-1675	5.3	21
		3ps MD+ 300 st. EM	1.93	157.3	-1612	-1771	5.1	11
	200 ms	Initial		1876.6	1597.1	-529.1	10	90
		500 st. EM	0.87	438.9	-1143.7	-1584.5	5.1	-3.5
		3ps MD+ 300 st. EM	1.86	366.8	-1329.6	-1698.3	5	-0.2
PARALLHDG.DNA	60 ms	Initial		1030	8729	7010.7	10	90
		500 st. EM	0.76	386.1	737.5	302.3	6.6	10
		3ps MD+ 300 st. EM	2.26	300.4	556.3	231	6.2	4
	200 ms	Initial		1876.6	9576	7010.7	10	90
		500 st. EM	0.74	720.8	1165	389.6	6.5	37
		3ps MD+ 300 st. EM	2.82	581	900	295.8	6	5

RMSD is the root mean-square deviation between the specified and the initial structure. E-rel is the value of the relaxation matrix term in the potential energy of the structure. E-total is the total value of the potential energy for the specified structure. E-potential is the value of the generic potential energy of the system, E-potential = $E_{bond} + E_{angle} + E_{dihedral} + E_{improper} + E_{vdw} + E_{elec}$ (all energies are in kcal/mol). The R-factor is the value of R(1/6) for the structure. Khi (χ) is the value of the glycosidic dihedral angle of the S3-T6 nucleotide.

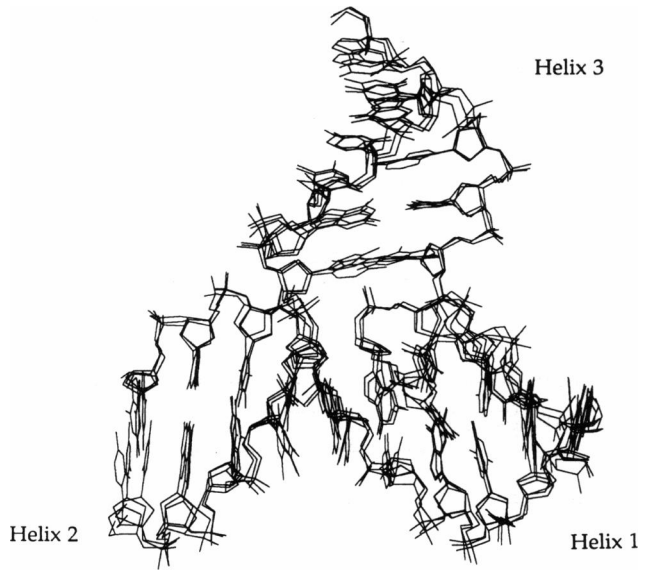

FIGURE 2 The superposition of three structures: (1) the initial structure, (2) the structure after backcalculation EM with NOE intensities from the 60 ms spectrum, and (3) the structure after backcalculation EM with NOE intensities from the 200 ms spectrum. Note that large changes are only observed in residues S3-T6 through S3-G8 and in the terminal residues.

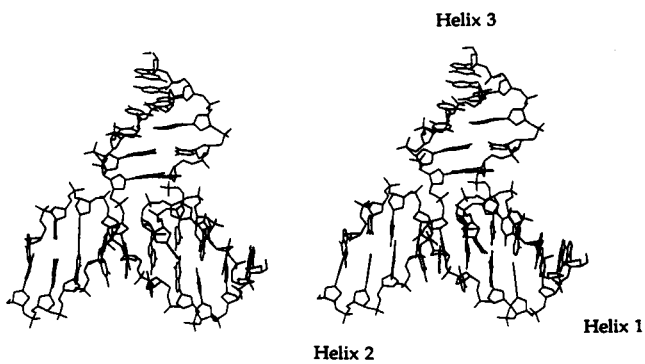

FIGURE 3 Stereoview of the refined structure of the TWJ-TC complex corresponding to structure (3) of Fig. 2. Only bonds between heavy atoms are shown.

Å. 3) The conformation of the S3-T6 base was determined primarily by the intense NOE between the methyl group of S3-T6 and H4′ of S3-G11. This crucial NOE peak is well resolved in all our spectra. 4) The NOE peak between the H5 proton of S3-C7 and H6 of S3-T6. This NOE connectivity is important for determination of the conformation of second base in the bulged region, S3-C7. 5) For S3-T6 and S3-C7, we also observed NOEs between H6 and all sugar protons in the same nucleotide (except H5′/H5″). 6) For S3-G8 and S3-G5, the NOE between H8 and H1′. For S2-G5, S1-C6, and S1-C5, the observed NOE patterns are very similar to those of nucleotide in B-form DNA helix, as described in the companion paper.

In Fig. 4 and Table 2, the values of the dihedral angles of the refined structure are plotted. These were calculated using the analysis feature of X-PLOR 3.1. Table 3 presents the helical parameters of each of the three arms of the TWJ, treated as separate helices, calculated using the "Dials and Windows" program (Ravishanker et al., 1989; Lavery and Sklenar, 1988). In addition, Strand 1 was analyzed using the same program to characterize the helical parameters of the dinucleotide step S1-C5 to S1-C6 that spans the junction.

Global structure after refinement

As in the preliminary model, the three helical arms form two domains. Two of the helices are stacked forming one continuous helical domain; these are Helices 1 and 2 formed by Strand 1. The other helical domain (Helix 3) is formed by the 5′-half of S3TC and the 3′-half of S2. The angle between Helix 3 and the Helix 1/Helix 2 domain equals approximately 80°. This value was determined by comparing angles between lines drawn normal to basepair planes at the ends of each helix. By comparison, Rosen and Patel determined a more acute angle of about 60° for J3CC (Rosen and Patel, 1993). In work devoted to the modeling of Holliday four-way junctions, the value of the crossing angle between the stacked arms was predicted to be close to 60° (von Kitzing et al., 1990). Experimental measurements in four-way junctions employing time-resolved fluorescence resonance energy transfer revealed that this angle can experience large fluctuations and is likely to deviate significantly from 60° (Eis and Millar, 1993).

Deviations from B-form DNA

The continuously stacked strand, S1, conforms to B-form DNA along its entire length (Fig. 3). Even at the junction, which is spanned by the S1-C5 to S1-C6 dinucleotide step, the dihedral angles of the backbone and the helical parameters fall within the range expected for B-DNA (see Fig. 4 and Tables 2 and 3). Deviation is observed, however, between S1-C5 and S1-C6 in the value of the parameter TWIST, which is 13° as compared with the usual 36° for B-form DNA. The backbone dihedral angles of the residues comprising the helical arms (S2-C1 to S2-G5, S2-C6 to S2-C10, S3-G1 to S3-G5, and S3-G8 to S3-C12) fall broadly within the range of values expected for B-form DNA (Berman, 1981). The nucleotides connecting Helix 3 and the continuously stacked Helix 1/Helix 2 domain are S2-G5 and S2-C6, and S3-G5 through S3-G8. It should be noted that no constraints were applied to the dihedral angles in any of the refinement or simulation protocols we used. Therefore, it is not surprising that a fairly large range of values is observed for most of the dihedral angles. Nonetheless, significant deviations occur primarily for backbone dihedral angles at the interface between the two domains, namely between S2-G5 and S2-C6 and between S3-G5 and S3-G8. The dihedral angles follow the standard convention, namely P-α-O5′-β-C5′-γ-C4′-δ-C3′-ϵ-O3′-ζ-P.

The dihedral angles at phosphorus, α and ζ, both rotate with low potential energy barriers. Many combinations of these angles are possible and have been observed in crystal

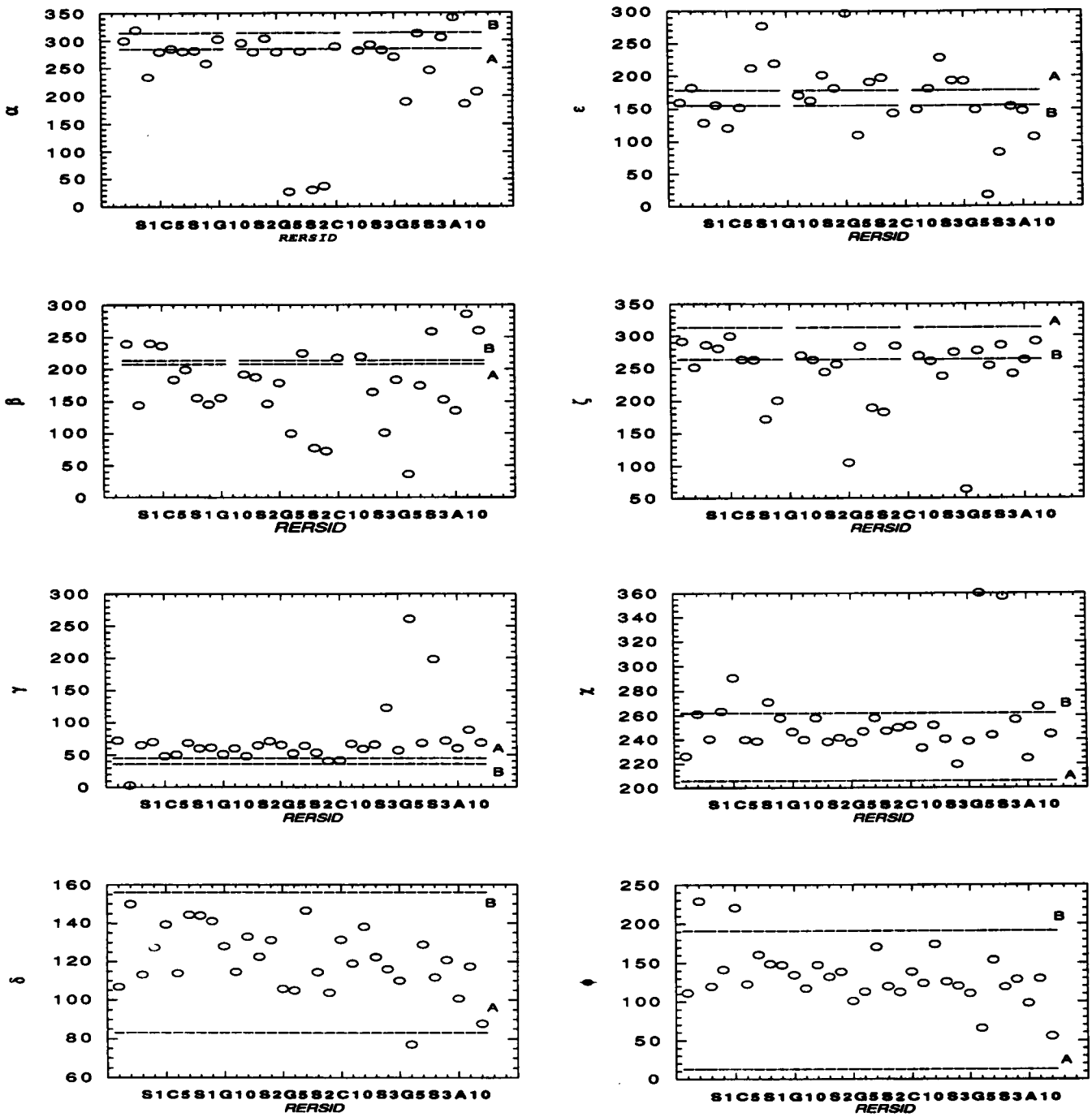

FIGURE 4 Backbone torsional angles calculated for the refined structure of TWJ-TC. The dashed lines indicates typical values of dihedral angles for B-form and A-form DNA duplexes (Ravishanker et al., 1989).

structures of nucleic acids (Berman, 1981). In the helical regions of the refined structure, we observe primarily the usual $\zeta^-\alpha^-$ combination, with some residues in $\zeta^t\alpha^-$, $\zeta^t\alpha^+$, and $\zeta^-\alpha^t$. For the S2-G5/S2-C6 step, which links Helix 2 to Helix 3, the unusual $\zeta^+\alpha^+$ conformation is observed. This is the conformation reported for the 180° turn between the third T and the A in the 5′-TTTA-3′ DNA hairpin studied by Blommers and co-workers (Blommers et al., 1991). For the S3-G5/S3-T6 step the conformation is $\zeta^+\alpha^t$.

The dihedral angle β is usually found in the *trans* configuration, with a large range of variation (Berman, 1981). This is also the case in the refined TWJ-TC structure. The largest deviation is found for the S3-G5 to S3-T6 step, where β is in the *gauche*$^+$ range.

The dihedral angle γ has been observed in all three potential energy minima (Berman, 1981), but crystallographically, it usually is found in the *gauche* domain. The only deviations from γ^+ found in the refined TWJ-TC structure

TABLE 2 Values of the dihedral angles of the refined structure of the TWJ-TC complex

RERSID	α	β	γ	δ	ε	ζ	χ	φ
S1 1 GUA			72.0	106.7	158.9	291.1	226.2	110.9
S1 2 CYT	299.2	239.3	2.3	149.8	181.9	251.4	260.9	228.7
S1 3 THY	318.5	143.8	65.0	113.0	128.2	286.0	240.2	119.6
S1 4 GUA	233.0	239.7	69.7	127.3	155.3	280.3	263.1	141.0
S1 5 CYT	279.0	236.1	47.8	139.1	120.6	299.3	290.3	220.3
S1 6 CYT	284.7	183.7	49.9	113.7	151.2	263.3	239.8	122.5
S1 7 ADE	279.7	199.0	68.1	144.4	212.3	262.3	238.5	160.4
S1 8 CYT	281.5	154.9	59.6	143.9	276.5	171.6	270.8	148.5
S1 9 CYT	257.9	145.1	60.8	141.0	219.0	200.1	257.8	147.2
S1 10 GUA	301.9	154.9	50.4	127.9			246.2	134.3
S2 1 CYT			59.3	114.5	170.6	269.7	239.7	116.7
S2 2 GUA	295.3	191.5	47.7	133.0	162.3	262.9	257.6	146.8
S2 3 GUA	278.7	186.8	64.2	122.4	201.1	244.1	238.2	131.6
S2 4 THY	303.0	145.8	70.7	131.0	181.2	256.0	241.4	138.4
S2 5 GUA	278.6	178.4	65.0	105.6	296.2	104.9	237.3	100.5
S2 6 CYT	26.6	99.5	51.9	104.8	109.9	283.5	246.7	112.3
S2 7 GUA	279.5	224.2	63.7	146.5	190.9	188.8	257.7	170.4
S2 8 THY	30.0	76.6	52.9	114.3	197.3	182.7	247.1	119.6
S2 9 CYT	36.9	71.9	40.2	103.5	143.1	284.2	249.6	111.9
S2 10 CYT	288.7	217.4	41.3	131.3			251.4	138.2
S3 1 GUA			66.6	118.8	148.9	269.2	233.1	123.2
S3 2 GUA	281.1	219.4	58.6	137.9	180.4	260.8	251.8	173.5
S3 3 ADE	292.0	163.9	65.3	122.1	228.2	237.6	240.2	125.4
S3 4 CYT	281.7	100.7	122.6	115.9	193.1	274.3	219.5	119.8
S3 5 GUA	269.5	182.7	56.4	109.8	192.1	64.0	238.5	110.4
S3 6 THY	189.2	35.8	260.8	76.8	148.3	277.1	359.8	65.6
S3 7 CYT	311.9	174.1	67.6	128.6	17.4	253.6	243.6	153.5
S3 8 GUA	245.4	258.0	197.8	111.4	83.3	285.3	357.2	118.6
S3 9 CYT	305.1	152.3	71.4	120.5	153.5	240.9	256.5	128.2
S3 10 ADE	340.5	134.9	59.6	100.6	146.6	262.4	224.6	97.7
S3 11 GUA	185.6	284.8	88.3	117.3	106.7	291.0	267.2	129.5
S3 12 CYT	207.3	259.7	68.2	87.7			244.3	55.4

are for the S3-G5 to S3-T6 dinucleotide step and the S3-C7/S3-G8 step, both of which are in the *trans* range.

Only the *trans* and *gauche*$^-$ conformations are allowed for the dihedral angle ε. In oligonucleotide crystals, tRNA, and fibrous polynucleotides ε is found primarily in the *trans* range, between 180° to 270° (Berman, 1981). In the von Kitzing model of the four-way junction, the ε angle at the strand cross-over point changes from the *trans* to the *gauche*$^-$ conformation (von Kitzing et al., 1990). In models of the J3CC three-way junction studied by Rosen and Patel, the dihedral angle ε in the C20/T21 dinucleotide step (which spans the junction region) also migrates toward the *gauche*$^-$ conformation. The corresponding dinucleotide step in TWJ-TC is S2-G5/S2-C6. The ε dihedral angle at this dinucleotide step in the refined model of TWJ-TC is −64°, which is also in the *gauche*$^-$ conformational range as is the epsilon angle between the third and fourth residues in the 5′-TTTA-3′ hairpin loop characterized by Blommers et al. (1991).

We also present the values of the pseudo-rotation angle, φ, for each sugar ring in the TWJ. These values fall within the range typically observed for B-form DNA, with the exception of the terminal residue, S3-C12. The pseudo-rotation angles are presented with the caveat that reliable determination of φ angles for each sugar will require determining sugar ring dihedral angles by conducting 2QF-COSY experiments and measuring the coupling constants between individual sugar protons.

Conformation of S3-T6 and S3-G8

Unusual values are only observed for the glycosidic angles of S3-T6 and S3-G8. Both of these are in the syn conformational range in the refined structure. In the case of S3-T6, this is necessary to allow the hydrophobic methyl group of the unpaired thymidine base to interact with the hydrophobic surface of the minor groove of the adjacent helix. The syn conformation allows the S3-G8 base to pair with S1-C5, given the unusual conformation of the sugar-phosphate backbone in the tight junction loop that joins the two helical domains of the TWJ complex.

We observe 9 NOE crosspeaks for protons of S3-T6 and 6 crosspeaks for protons from S3-C7. The NOE connectivity that confirmed the conformation of S3-T6 is between protons of the methyl group and aromatic proton H6 of this residue and the H4′ proton of S3-G11. Under the influence of these NOEs, the glycosidic angle changes from 90° in the starting structure to values around 0° (right in the syn range) after relaxation matrix refinement with volumes from the NOESY data. The S3-T6 base has some freedom to rotate around the glycosidic bond. We have also carried out the calculation with NOE constraints starting with this angle set to 180° (anti) to ascertain the influence of the initial value of this angle. In this case, the final value of the glycosidic angle was 242°, which is still within the anti range, and the final distance between the methyl carbon of S3-T6 and H4′ of S3-G11 is 4.8 Å, which is 1 Å more than obtained when the

TABLE 3 The helical parameters of the arms of the TWJ-TC complex determined using the "Dials and Windows" program

	XDP(A)	YDP(A)	INC(°)	TIP(°)	AXD(A)	AYD(A)	AIN(°)	ATP(°)
S1C6-S2G5	-0.58	0.35	-2.48	4.79				
S1A7-S2T4	−0.65	0.11	−1.52	4.56	−0.16	−0.21	−1.53	1.43
S1C8-S2G3	−0.48	0.44	−1.4	−0.91	0.16	0.41	0.76	−1.95
S1C9-S2G2	−0.58	0.35	−3.87	−1	−0.01	−0.08	−0.4	0.78
S1G10-S2C1	−0.56	0.22	−5.3	−0.56	0.07	−0.12	−0.38	0.7
S3G8-S1C5	0.17	0.46	7.32	4.28				
S3C9-S1G4	−0.28	0.66	2.52	6.03	−0.29	0.34	0.35	6.55
S3A10-S1T3	−0.59	−0.09	1.84	−7.3	0	−0.45	4.03	−9.39
S3G11-S1C2	−0.36	−0.14	−0.48	−10.21	0.25	0.06	2.46	−0.37
S3C12-S1C1	−0.42	−0.23	−2.83	−12.06	−0.02	−0.1	−0.48	−2.84
S2C6-S5G5	−0.55	−0.09	−7.02	4.28				
S2G7-S3C4	−0.54	−0.39	−1.88	6.69	−0.11	−0.25	2.26	1.97
S2T8-S3A3	−0.51	−0.09	−1.38	−1.24	−0.07	0.44	−1.17	−6.82
S2C9-S3G2	−0.59	−0.09	−1.88	−5.39	−0.01	0.14	−0.12	−1.67
S2C10-S3G1	−0.47	−0.3	−5.28	−4.62	0.2	−0.14	−1.66	1.84
B-form	−0.7	0	−5.9	0	0	0	0	0
A-form	−5.4	0	19.1	0	0	0	0	0

	SHR(A)	STR(A)	STG(A)	BKL(°)	PRP(°)	OPN(°)	SHF(A)	SLD(A)	RIS(A)	TLT(°)	ROL(°)	TWS(°)
S1C6-S2G5	−0.46	0.13	0.26	4.78	−14.98	−8.34						
S1A7-S2T4	−0.38	−0.19	0.1	−3.93	−28.19	−2.85	−0.23	−0.46	3.38	−0.56	1.19	42.97
S1C8-S2G3	0.32	0.15	0.02	8.07	−8.53	−2.41	0.33	0.74	3.25	0.87	−7.42	35.68
S1C9-S2G2	0.51	−0.11	0.26	6.22	−14.12	−7.15	−0.11	−0.17	3.36	−2.51	0.69	43.49
S1G10-S2C1	−0.61	−0.06	0.18	5.68	−24.48	−2.9	0.09	−0.25	3.63	−1.81	1.14	34.64
S3G8-S1C5	0.22	0.04	0.56	−8.81	14.2	21.75						
S3C9-S1G4	0.52	−0.2	1.17	−7.77	10.36	−7.87	−0.74	0.55	4.09	−4.45	8.3	35.38
S3A10-S1T3	−0.27	−0.24	0.83	16.34	−23.54	−21.61	−0.32	−1.21	2.65	3.35	−22.73	33.34
S3G11-S1C2	−0.45	−0.03	0.49	19.41	−8.04	−6.52	0.48	0	3.56	0.15	−3.28	28.63
S3C12-S1C1	0.24	0.07	−0.15	25.83	−10.52	−1.58	−0.08	−0.18	3.32	−2.83	−4.69	47.76
S2C6-S5G5	0.5	−0.01	0.73	−7.81	−8.26	−13.07						
S2G7-S3C4	0.11	0.04	−0.27	−10.45	−25.7	−8.55	−0.11	−0.55	3.45	7.39	4.38	41.96
S2T8-S3A3	−0.07	−0.06	−0.18	6.71	−17.3	−10.53	−0.04	0.74	2.88	−0.67	−14.74	34.19
S2C9-S3G2	0.09	−0.08	−0.13	8.69	−15.96	−9.43	−0.1	0.13	3.36	−0.62	−5.82	35.62
S2C10-S3G1	0.17	0	0.19	2.97	−1.14	−6.18	0.32	−0.34	3.64	−5.06	2.82	44
B-form	0	0	0.1	0.2	3.7	−4.1	0	0	3.4	0	0	36
A-form	0	−0.4	0.2	−0.1	13.7	−4.6	0	0	2.6	0	0	32.7

E. Ravishankar et al. (1989). Intra-base pair parameters: shear (SHR), Stretch (STR), Stagger (STG), Buckle (BKL), Propeller (PRP), and Open (OPN). Inter-base pair parameters: Shift (SHF), Slide (SLD), Rise (RIS), Tilt (TLT), Roll (ROL), and Twist (TWS). Axis-base pair parameters: X-Displacement (XDP), Y-displacement (YDP), Inclination (INC), and Tip (TIP). Axis parameters: axis-X displacement (AXD), axis-Y displacement (AYD), axis-inclination (AIN), and axis-tip (ATP). In the two last lines, the values of the specified parameters for idealized DNA duplexes in B-form and A-form are presented.

glycosidic angle is initially set within the syn range. In Fig. 5 we present for comparison the local view of the position of the S3-T6 residue in structures refined with the S3-T6 base initially in the syn and in the anti conformation. When the base of S3-T6 is in the syn conformation, the other NOEs of protons that belong to the residues in the bulged region are also better satisfied. As mentioned above, the hydrophobic methyl group is in direct contact with the hydrophobic surface of the minor groove of Helix 1 in our refined structure. This feature of our structure is similar to that reported for the second unpaired thymidine in the DNA 5′-TTTA-3′ hairpin loop (Blommers et al., 1991). This base also lies in the minor groove of the helical stem.

The S3-G8 glycosidic angle is approximately 0° in our refined model. Comparison can be made to guanosines in Z-DNA, which typically exhibit glycosidic dihedral angles in the range 40° to 90° (Dickerson, 1992). A guanosine mismatched with an adenosine in a B-form oligonucleotide duplex was observed by x-ray crystallography to have a glycosidic angle of 86° (Brown et al., 1989). Recently an alternative structure to Z-DNA was proposed on the basis of modeling studies for left-handed structures generated in vivo by the torque available in naturally occurring DNA supercoils (Ansevin and Wang, 1990). This structure has been termed Z(WC)-DNA or W-DNA on account of its Watson-Crick-type backbone directions. The glycosidic angle of the guanosine is about 12° in this structure, close to the value which we observe for S3-G8.

Glycosidic angles in the range 40° to 90° bring the H8 and H1′ atoms quite close together, typically 2.3 to 2.5 Å, whereas for a glycosidic angle of about 0° a value of 2.9 to 3.0 Å is expected for this distance (Wüthrich, 1986). However, in our model a value of 3.3 Å is observed, consistent with the NOE crosspeak intensity. One explanation for this unusually long H8 to H1′ distance and the unusual dihedral angle is that the stacking forces that position the base generate some strain in the glycosidic bond.

Molecular dynamics simulation

To check the stability of our structure, we carried out molecular dynamics simulations of TWJ-TC in water. The final refined structure was solvated by surrounding it with 855

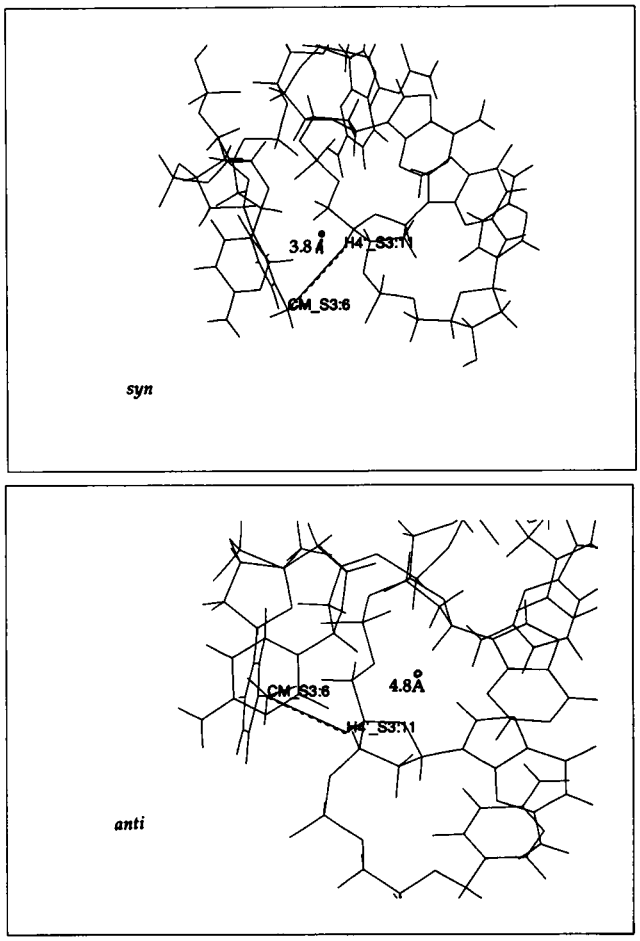

FIGURE 5 The position of the S3-T6 base in syn and anti conformations in the TWJ-TC complex. The distance between the methyl carbon of S3-T6 and the H4′ proton of S3-G11 in each conformation is shown. A strong NOE crosspeak is observed between the methyl protons of S3-T6 and H4′ of S3-G11.

molecules of water. At least two diameters of water molecules were used to construct the water shell surrounding the TWJ complex. The solvation procedure and the calculation were carried out using the corresponding feature of QUANTA/CHARMm 3.3 (Molecular Simulation). The solvated structure was subjected to 1000 steps of EM to remove close contacts among water molecules and between water molecules and the TWJ. Then the minimized structure was gradually heated from 0 K up to 300 K during 3 ps of MD simulation. The heating period was followed by equilibration for 6 ps. After equilibration, 30 ps of molecular dynamics simulation were conducted. It turned out that the equilibration period of the simulation was too short for our solvated molecule, and during the MD simulation the temperature of the system increased to 330 K. Fortunately, this did not affect the result of the simulation: the structure of the TWJ presented in this work was stable during the entire course of the simulation. During the course of the simulation, the angles between Helix 1 and Helix 3 and Helix 1 and Helix 2 remained in the region 70–85°. The glycosidic angles of S3-T6 and S3-G8 also remained unchanged in the syn region during the simulation. Only random fluctuations were observed in the glycosidic dihedral angles of S3-T6 and S3-G8.

In conclusion, the refined structure presented in Fig. 3 is consistent with 2D NOESY data acquired at 60 and 200 ms. The global structure is defined by the stacking of Helices 1 and 2 and the tertiary contact between S3-T6 and S3-G11 that determines how the two helical domains are joined together. The structure is stable during molecular dynamics simulation. Further refinement of the structure will require experimental constraints on backbone dihedral angles.

The authors thank Professor P. L. Rinaldi and Dr. D. G. Ray of Akron University for help in acquiring NMR spectra and our collaborators on the companion paper for useful discussions.

This work was supported by U.S. Public Health Service grant GM41454, by Ohio Board of Regents Research Challenge grants, by Petroleum Research Fund grant 20871-GB4, and by Research Corporation grant C-2314.

REFERENCES

Ansevin, A. T., and A. H. Wang. 1990. Evidence for a new Z-type left-handed DNA helix: properties of Z(WC)-DNA. *Nucleic Acids Res.* 18: 6119–6126.

Berendsen, H. J. C., J. P. M. Postma, N. F. van Gunsteren, A. DiNola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Berman, H. M. 1981. Conformational principles of nucleic acids. *In* Topics in Nucleic Acid Structure. S. Neidle, editor. John Wiley & Sons, New York.

Blommers, M. J. J., F. J. M. Van De Ven, G. A. Van Der Marel, J. H. Van Boom, and C. W. Hilbers. 1991. The three-dimensional structure of a DNA hairpin in solution. Two-dimensional NMR studies and structural analysis of d(ATCCTATTTATAGGAT). *Eur. J. Biochem.* 201:33–51.

Brown, T., G. A. Leonard, E. D. Booth, and J. Chambers. 1989. Crystal structure and stability of a DNA duplex containing A(anti)-G(syn) base-pairs. *J. Mol. Biol.* 207:455–457.

Brünger, A. T. 1993. X-PLOR, version 3.1. A System for X-Ray Crystallography and NMR. Yale University Press, London.

Dickerson, R. E. 1992. DNA structure from A to Z. *Methods. Enzymol.* 211:67–111.

Duckett, D. R., and D. M. Lilley. 1990. The three-way DNA junction is a Y-shaped molecule in which there is no helix-helix stacking. *EMBO J.* 9:1659–1664.

Eis, P. S., and D. P. Millar. 1993. Conformational distributions of a four-way DNA junction revealed by time-resolved fluorescence resonance energy transfer. *Biochemistry.* 32:13852–13860.

Guo, Q., L. Min, M. E. A. Churchill, T. D. Tullius, and N. R. Kallenbach. 1990. Asymmetric structure of a three-arm DNA junction. *Biochemistry.* 29:10927–10934.

von Kitzing, E., D. M. J. Lilley, and S. Diekmann. 1990. The stereochemistry of a four-way DNA junction: a theoretical study. *Nucleic Acids Res.* 18:2671–2683.

Kuszewski, J., M. Nilges, and A. Brünger. 1992. Sampling and efficiency of metric matrix distance geometry: a novel "partial" metrization algorithm. *J. Biomolecular NMR.* 2:33–56.

Lavery, R., and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

Leontis, N. B., W. Kwok, and J. S. Newman. 1991. Stability and structure of the three-way DNA junctions containing unpaired nucleotides. *Nucleic Acids Res.* 19:759–766.

Leontis, N. B., M. T. Hills, M. Piotto, A. Malhotra, J. Nussbaum, and D. G. Gorenstein. 1993. A model for the solution structure of a branched, three-strand DNA complex. *J. Biomol. Struct. Dyn.* 11: 215–223.

Ma, R.-I., N. R. Kallenbach, R. D. Sheardy, M. L. Petrillo, and N. C. Seeman. 1986. Three-arm nucleic acid junctions are flexible. *Nucleic Acids Res.* 14:9745–9753.

Macaya, R., E. Wang, P. Schultze, V. Sklenar, and J. Feigon. 1992. Proton nuclear magnetic resonance assignments and structural characterization of an intramolecular DNA triplex. *J. Mol. Biol.* 225:755–773.

Mujeeb, A., S. M. Kervin, G. L. Kenyon, and T. L. James 1993. Solution structure of a conserved DNA sequence from the HIV-1 genome: restrained molecular dynamics simulation with distance and torsion angle restraints derived from two-dimensional NMR spectra. *Biochemistry.* 32:13419–13431.

Ravishanker, G., S. Swaminathan, D. L. Beveridge, R. Lavery, and H. Sklenar. 1989. Conformational and helicoidal analysis of 30 ps of molecular dynamics on the d(CGCGAATTCGCG) double helix: "curves, dials and windows." *J. Biomol. Struct. Dyn.* 6:669–700.

Rosen, M. R., and D. J. Patel. 1993. Structural features of a three-stranded DNA junction containing a C-C junctional bulge. *Biochemistry.* 32: 6563–6576.

Wüthrich, K. 1986. NMR of Proteins and Nucleic Acids. John Wiley & Sons, New York.