Helical Stacking in DNA Three-Way Junctions Containing Two Unpaired Pyrimidines: Proton NMR Studies

Neocles B. Leontis,* Michael T. Hills,‡ Martial Piotto,§ Igor V. Ouporov,* Arun Malhotra,¶ and David G. Gorenstein‖

*Department of Chemistry, Bowling Green State University, Bowling Green, Ohio 43403-0213 USA; ‡Department of Chemistry, Purdue University, West Lafayette, Indiana 47907 USA; §Bruker Spectrospin, 67160 Wissembourg, France; ¶The Rockefeller University, New York, New York 10021 USA; and ‖Department of Human Biological Chemistry and Genetics, University of Texas, Medical Branch at Galveston, Galveston, Texas 77555 USA

ABSTRACT The proton NMR spectra of DNA three-way junction complexes (TWJ) having unpaired pyrimidines, 5′-TT- and 5′-TC- on one strand at the junction site were assigned from 2D NOESY spectra acquired in H_2O and D_2O solvents and homonuclear 3D NOESY-TOCSY and 3D NOESY-NOESY in D_2O solvent. TWJ are the simplest branched structures found in biologically active nucleic acids. Unpaired nucleotides are common features of such structures and have been shown to stabilize junction formation. The NMR data confirm that the component oligonucleotides assemble to form conformationally homogeneous TWJ complexes having three double-helical, B-form arms. Two of the helical arms stack upon each other. The unpaired pyrimidine bases lie in the minor groove of one of the helices and are partly exposed to solvent. The coaxial stacking arrangement deduced is different from that determined by Rosen and Patel (Rosen, M. A., and D. J. Patel. 1993. *Biochemistry.* 32:6576–6587) for a DNA three-way junction having two unpaired cytosines, but identical to that suggested by Welch et al. (Welch, J. B., D. R. Duckett, D. M. J. Lilley. 1993. *Nucleic Acids Res.* 21:4548–4555) on the basis of gel electrophoretic studies of DNA three-way junctions containing unpaired adenosines and thymidines.

INTRODUCTION

Multi-branch structures occur in DNA and RNA. Most currently accepted models for genetic recombination involve the four-way DNA junction (FWJ), or Holliday junction as an intermediate (Holliday, 1964). The Holliday junction is a transiently occurring species with a mobile branch point. Oligonucleotide sequences have been designed and synthesized to assemble in immobile junction structures as models of Holliday junctions (Seeman and Kallenbach, 1983). Numerous structural and thermodynamic studies of immobile DNA FWJ have appeared and have been the subject of several recent reviews (Lu et al., 1992; Lilley and Clegg, 1993a, b; Seeman and Kallenbach, 1994).

Three-dimensional models for DNA FWJ have been presented that feature approximately coaxial pairwise stacking of the four helical arms of the junction without disruption of basepairing (Sigal and Alberts, 1972; von Kitzing et al., 1990). Experimental evidence supporting this model has been obtained by solution NMR studies (Chen and Chazin, 1994). These studies provide further evidence that the strong tendency of basepairs to undergo stacking interactions is a major force for the stabilization of tertiary structure in nucleic acids. These observations suggest that coaxial stacking should also occur in junctions having an odd number of helices and raise the question, what determines which pairs of helices interact by stacking in such cases as three-way and five-way junctions where one or more helices will be left without a stacking partner?

Three-way DNA junctions, therefore, serve as useful structural benchmarks for comparison to DNA four-way junctions. However, they may be biologically significant in their own right. For example, a number of proteins that have been shown to interact specifically with DNA four-way junctions also bind to three-way junctions (Duckett et al., 1992). The best studied is endonuclease VII of T4 Bacteriophage (gene product 49), which appears to be responsible for the resolution of branched structures arising during replication and genetic recombination of phage DNA. Endonuclease VII has been shown to interact with and resolve synthetic DNA three-way junctions in vitro, as well as FWJ (Jensch and Kemper, 1986). Furthermore, studies using electron microscopy have shown that branched viral DNA structures with three-arms ("Y-structures") as well as four-arms accumulate in *Escherichia coli* cells infected with phage T4 lacking functional endonuclease VII (Minagawa et al., 1983; Flemming et al., 1993). Many small molecules ligands bind tightly to junction structures (Lu et al., 1992). For example, the cationic porphyrin *meso*-tetra(*para*-trimethylaniliniumyl)porphine (TMAP), which is too bulky to intercalate in duplex DNA, binds by an insertional mode in the junction region of DNA TWJ complexes (Nussbaum et al., 1994).

Three-way junction structures occur frequently in RNA molecules including the ribosomal RNAs (Noller, 1984, 1991), the hammerhead ribozymes (Forster and Symons, 1987; Uhlenbeck, 1987), and snU4/U6 RNA (Guthrie and Patterson, 1988). The hammerhead ribozyme is an example of a catalytic RNA that is a TWJ (Wilson and Szostak, 1992).

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

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

Abbreviations used: NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; FWJ, four-way junction; TWJ, three-way junction.

0006-3495/94/01/251/15 $2.00

5S ribosomal RNA is an integral component of all ribosomes, eukaryotic, eubacterial, and archaebacterial. All 5S RNAs contain a central junction loop that is a TWJ. It was proposed in the early 1980s that either helices I and V or helices II and V are coaxial (Luehrsen and Fox, 1981; Stahl et al., 1981).

The role of unpaired bases in biologically active three-way junctions has been examined in the hammerhead ribozyme and in 5S RNA. All but four of the nucleotides of the hammerhead ribozyme, all of which are unpaired, may be replaced with DNA analogs with retention of enzymatic activity (Yang et al., 1992). This provides evidence that the unpaired junction residues in this case are directly involved in catalysis. The unpaired bases in the junction region of 5S RNA are phylogenetically conserved. It has been shown using site-specific mutagenesis that the junction region is crucial for the recognition of protein TFIIIA by the 5S rRNA of *Xenopus laevis* (Baudin et al., 1991) and of ribosomal protein L18 by the 5S rRNA of *E. coli* (Egebjerg et al., 1989). Based on their modeling and chemical probing studies of 5S RNAs from eubacterial, chloroplast, eukaryotic, and archaebacterial sources, as well as consideration of the crystallographic structures of tRNA, the groups of Westhof, Ehresmann, and Romaniuk propose as a general rule of RNA folding that the strand with the least number of unpaired connecting nucleotides between helical sections should become part of the continuous double helix when there is a competition among helices for colinear stacking, as exists at multi-branched junctions (Westhof et al., 1989; Baudin et al., 1991; Brunel et al., 1991). Application of this concept to 5S RNA leads these workers to conclude that helices II and V are coaxially stacked in all 5S RNAs. Shen and Hagerman (1994) have shown using a combination of gel electrophoresis and transient electric birefringence measurements that two of the three helices in the 5S rRNA from the thermophilic bacterium *Sulfolobus acidocaldarius* are coaxially stacked. According to this principle, there should be two favorable stacking arrangements accessible to our simple TWJ that only have two unpaired bases on one strand. We will show that in fact the preferred stacking arrangement of our TWJ complexes is one of these two. We will compare our structure with that determined by Rosen and Patel (1993b), also using NMR methods, for a TWJ complex having two unpaired cytosines and a different junction base sequence, and with the helical stacking proposed for a third set of TWJs by Welch et al. (1993) based on gel electrophoretic measurements.

Using gel electrophoresis and UV melting techniques, we showed that two or more unpaired bases stabilize TWJ complexes (Leontis et al., 1991). The TWJ that have been successfully studied by NMR have included two unpaired nucleotides in the junction region (Leontis et al., 1993; Rosen and Patel, 1993a, b). The sequences and base-pairing of two related TWJ complexes, TWJ-TT and TWJ-TC, that we have studied are shown in the left and right panels Fig. 1, respectively. Rosen and Patel studied a TWJ called J3CC that is based on an unrelated sequence in the helical arms but is similar to our complexes in the sequence around the junction (with the only difference being that one flanking GC basepair in TWJ-TC is substituted with an AT basepair in J3CC) and in the inclusion of two unpaired pyrimidines (-CC-) at the junction. J3CC shares several common structural features with the TWJ complexes we have studied: a unique stacking arrangement of two of the three helices is found in each junction. Moreover, the unpaired pyrimidines are found to be extra-helical and exposed to solvent. However, as we will show, the stacking arrangement in TWJ-TT and TWJ-TC, relative to the unpaired bases, is different than in J3CC. Consequently, the tertiary interactions in these two sets of junctions are also different. The purpose of this paper is to present the NMR data demonstrating unequivocally the unique stacking arrangement adopted by TWJ-TT and TWJ-TC and to prove with relative confidence the identity of the tertiary contact mentioned above. These observations form the basis for construction of three-dimensional models that were subjected to quantitative refinement, the subject of the companion paper.

The significant structural questions regarding TWJ complexes that can be addressed using NMR spectroscopy include the following:

1. Do the helical arms of the junction retain the characteristic structure of B-type DNA helices?
2. Do all the expected Watson-Crick pairs form between the bases immediately flanking the junction region?
3. Does a unique pairwise stacking arrangement prevail between two of the three helical arms? If so, which of the helices interact with each other?
4. Do the unpaired bases in the junction region stack within the junction, or are they partially or largely exposed to solvent? Do the unpaired bases interact with the surface of one of the helical arms?

The NMR data will be presented to answer these questions. The approach taken to construct a qualititative three-dimensional model consistent with these data will be described. In the companion paper, the quantitative refinement of this model against 2D NOESY data for the TWJ-TC sample will be presented.

MATERIALS AND METHODS

DNA oligonucleotides

All oligonucleotides used in this study were obtained from the Midland Reagent Company (Midland, TX). All oligonucleotides were synthesized on a 10 μmole scale using phosphoramidite chemistry and purified using ion-exchange (for S1 and S2) or reverse-phase (for the G-rich S3TT sequence) HPLC chromatography.

Sample preparation

The purified oligonucleotides constituting each junction sample were dissolved individually in D_2O and quantitated using UV absorption. The molar extinction coefficients (at 260 nm) were calculated according to published nearest neighbor parameters (Puglisi and Tinoco, Jr., 1989). They are: 85,480 Absorbance Units (260 nm)/M for S1; 89,180 AU(260 nm)/M for S2; and 122,120 AU(260)/M for S3TT. Stoichiometric samples of the three strands were assembled by NMR titration using the quantities calculated from UV absorption as a guideline. Titrations were performed at 80°C in

D_2O solvent by simultaneously monitoring the well resolved thymidine methyl and aromatic regions of the spectrum. Alternatively, samples were titrated by gel electrophoresis. Samples were lyophilized and redissolved in buffers consisting of 10 mM sodium phosphate (pH 6.5–7.0), 0.5 mM EDTA, 10 mM $MgCl_2$, and 100 mM NaCl. Finally, samples were concentrated by ultrafiltration using Centricon-3 centrifugal concentrator devices (Amicon). Samples intended for D_2O solvent were lyophilized 3 times from 99.996% D_2O and dissolved to a concentration of about 2 mM in 0.6 ml of D_2O. Samples intended for study of exchangeable proton resonances were prepared by an identical procedure, except for dissolution in 95% H_2O/5% D_2O.

One-dimensional ^{1}H NMR temperature series

One-dimensional spectra were acquired on a VXR-500S 500 MHz NMR spectrometer from 25° to 80°C in 5° increments, using a standard one-pulse experiment, to determine the optimum temperature at which to run the 2D and 3D experiments.

Two-dimensional nuclear Overhauser spectroscopy

The 100 ms 2D-NOESY spectrum was acquired on a Varian VXR-500S 500 MHz spectrometer at 30°C in D_2O. The standard 90°-t1–90°-tm-90°-acquire pulse sequence was used with presaturation of the water resonance during the 4 s relaxation delay and the mixing period. Phase-sensitive data were collected using the States-Haberkorn hypercomplex method (States et al., 1982). 4096×768 complex points were acquired in t2/t1, respectively, with 32 scans per increment. The spectral width was 4100 Hz (8.1 ppm) in both dimensions.

The 2D NOESY spectra were obtained in H_2O solvent using a Bruker AMX spectrometer operating at 500 MHz and equipped with an inverse gradient probehead. The sequence used combines the 1-1 echo pulse scheme with pulsed field gradients. A 6 ms sine-shaped gradient pulse of intensity G_1 = 21 G/cm was applied during the NOE mixing time and two 300 μs sine-shaped gradient pulses of intensities $G_2 = G_3 = 31.5$ G/cm were applied during the 1-1 echo sequence (Sklenar et al., 1993). The delay was adjusted to give maximum excitation in the middle of the amino and imino proton region. Acquisition parameters were: 4K complex points in f2, 512 increments in f1 (States-TPPI), 32 scans per increment, recycle delay 1.5 s, 150 ms mixing time, spectral width 12820 Hz in both dimensions.

^{1}H homonuclear 3D-NOESY-TOCSY spectroscopy

The 3D-NOESY-TOCSY spectrum was acquired on a Varian VXR-500S 500 MHz NMR spectrometer as described by Piotto and Gorenstein (1991) at 30°C in D_2O. The NOESY and TOCSY mixing times were 200 and 35 ms, respectively. The spectral width was 4500 Hz (9.0 ppm) in all three dimensions. The residual solvent peak was presaturated with the decoupler during the 0.7 s relaxation delay. 233×128×1024 complex points were acquired in t1/t2/t3 dimensions, respectively, with 8 scans per t3 fid. Data processing was done on a Silicon Graphics 4D/25 workstation using FELIX NMR processing software (Biosym, Inc., San Diego, CA). Kaiser window apodization and first-order polynomial baseline correction were applied to the t3 dimension. Linear prediction was applied in t2 to the first data point and to extend the t2 fid from 64 to 90 real points. A 70° shifted sinebell, zero filling, and no baseline correction were also applied to the t2 dimension. Linear prediction of the 1st data point, 70° shifted sinebell, zero filling, and no baseline correction were applied to the t1 dimension. The final fully transformed spectrum contained 256×128×512 real points.

Further details on NMR methods used will appear elsewhere (Leontis et al., in press).

RESULTS

The sequence and secondary structure of the DNA junction consisting of the three strands S1, S2, and S3TT (TWJ-TT) is shown schematically in Fig. 1 (*left* panel). The GC-rich, thermally stable sequence of TWJ-TT was designed to include one AT base-pair in each helical arm and two thymidine residues as the stabilizing unpaired bases in the junction; the methyl groups of the thymidines serve as spectrally well resolved conformational probes. The helix formed by the 5′-half of S1 paired to the 3′-end of S3 is designated helix 1; that formed by the 5′-half of S2 paired to the 3′-half of S1, helix 2; and that formed by the 5′-end of S3 paired to the 3′-half of S2, helix 3. Assignment of the highly overlapped sugar and aromatic regions of the proton NMR spectrum required three-dimensional, homonuclear ^{1}H NOESY-TOCSY spectroscopy. The assignment of all the nonexchangeable ^{1}H resonances (with the exception of some H4′, H5′, and H5″ sugar protons), and of all observable imino and cytosine amino resonances is presented in Tables 1 (TWJ-TT) and 2 (TWJ-TC). The nonexchangeable proton resonances were assigned using strategies similar to those already reported (Mooren et al., 1991; Piotto and Gorenstein, 1991; Radhakrishnan and Patel, 1991; Radhakrishnan et al., 1992). The imino, cytosine amino, and adenosine H2 resonances were assigned from 2D NOESY spectra carried out in 90% H_2O using pulsed field gradient technology (Sklenar, et al., 1993). Short (<4.5 Å) ^{1}H-^{1}H distances were identified by qualitative analysis of the 3D spectra, recorded in 100% D_2O and of 2D ^{1}H NOESY spectra recorded in 100% D_2O and in 90% H_2O. Analysis of the spectrum of the parent junction, TWJ-TT, was assisted by NMR studies of the closely related junction, S1(G1):S2:S3TC(C12), TWJ-TC, also shown schematically in Fig. 1 (*right* panel). In this complex, one of the unpaired thymidines in Strand 3 is replaced with a cytosine (S3-C7), and one terminal basepair is re-

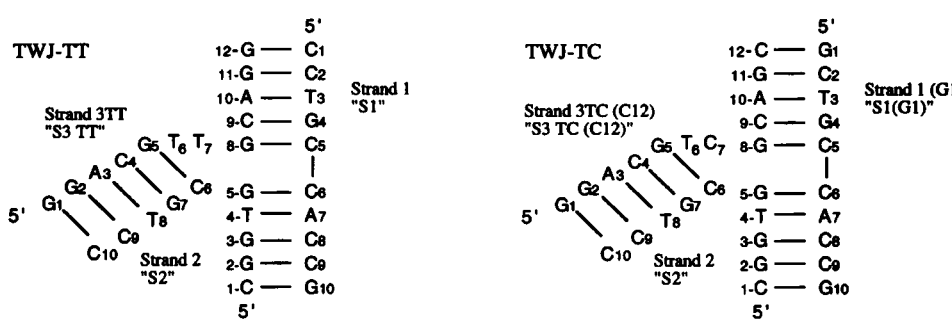

FIGURE 1 The sequences of the three-arm junctions, S1:S2:S3TT (TWJ-TT) and S1(G1):S2:S3TC(C12) (TWJ-TC) studied in this work. Note that the terminal base-pair of Helix 1 is reversed in TWJ-TC. This change, which had no effect on the structure of the junction, aided in assigning overlapping resonances.

TABLE 1 Proton chemical shifts for TWJ-TT at 30°C

	Arom	H2/5/ME	H1′	H2″	H2′	H3′	H4′	H5′/H5″	G/T imino	C amino
S1-C1	7.83	5.97	6.02	2.61	2.29	4.74	4.18	3.84/3.81		
S1-C2	7.71	5.72	6.07	2.55	2.24	4.87	4.20	4.26/4.15	13.02	8.44, 7.07
S1-T3	7.44	1.74	5.87	2.55	2.18	4.94	4.23	4.19	13.97	
S1-G4	7.97		6.03	2.81	2.72	5.04	4.37		12.96	8.46, 6.61
S1-C5	7.48	5.43	6.03	2.47	2.24	4.87	4.28	4.23/4.19	13.12	8.37, 6.61
S1-C6	7.46	5.45	5.55	2.33	1.92	4.85	4.18		12.72	8.23, 6.61
S1-A7	8.34	7.79	6.28	2.94	2.77	5.06	4.46	4.20/4.15	13.78	
S1-C8	7.37	5.38	5.90	2.43	2.04	4.85	4.32	4.22/4.11	12.88	8.15, 6.60
S1-C9	7.49	5.66	5.74	2.38	2.00	4.86	4.13		13.17	8.64, 7.00
S1-G10	8.00		6.22	2.44	2.68	4.74	4.24	4.14/4.09		
S2-C1	7.61	5.90	5.78	2.41	1.89	4.73	4.11	3.76		
S2-G2	7.98		5.69	2.81	2.74	5.01	4.37		13.17	8.64, 7.00
S2-G3	7.82		6.09	2.84	2.62	5.02	4.46	4.26	12.88	8.15, 6.60
S2-T4	7.20	1.42	5.94	2.49	2.08	4.98	4.30	4.24/4.16	13.78	
S2-G5	7.92		6.13	2.71	2.71	5.02	4.48	4.21/4.17	12.72	8.23, 6.61
S2-C6	7.50	5.48	5.97	2.62	2.30	4.86	4.15	4.13		
S2-G7	8.08		6.23	2.97	2.72	5.06	4.49	4.24	12.92	8.39, 6.62
S2-T8	7.40	1.52	6.25	2.64	2.20	4.97	4.21		13.85	
S2-C9	7.66	5.80	6.14	2.53	2.27	4.92	4.26	4.16	13.00	8.58, 7.03
S2-C10	7.66	5.75	6.30	2.34	2.34	4.61	4.10	4.13		
S3-G1	7.74		5.71	2.70	2.53	4.87	4.41	3.76/3.70		
S3-G2	7.94		5.61	2.84	2.75	5.08	4.42	4.32/4.22	13.00	8.58, 7.03
S3-A3	8.30	8.05	6.29	2.93	2.77	5.12	4.53	4.23/4.22	13.85	
S3-C4	7.32	5.37	5.67	2.39	2.01	4.84	4.24	4.15	12.92	8.39, 6.62
S3-G5	8.05		6.12	2.71	2.71	5.02				
S3-T6	7.88	2.01	5.92	2.55	2.34	4.94		4.28/4.13		
S3-T7	7.51	1.49	5.88	2.30	2.09	4.76	4.09			
S3-G8	7.80		6.02	2.73	2.68	5.04	4.38	4.21	13.12	8.37, 6.61
S3-C9	7.51	5.47	5.65	2.39	2.00	4.84	4.24	4.21	12.96	8.46, 6.61
S3-A10	8.20	7.76	5.98	2.84	2.69	5.05	4.40	4.20/4.10	13.97	
S3-G11	7.71		5.89	2.66	2.52	5.03	4.38	4.21/4.11	13.02	8.44, 7.07
S3-G12	8.07		6.15	2.40	2.44	4.65	4.26	4.15		

Samples were dissolved in D_2O or H_2O buffer consisting of 100 mM sodium chloride, 10 mM sodium phosphate, 5 mM magnesium chloride, and 0.5 mM EDTA.

versed (S1-G1/S3-C12). The NMR data of TWJ-TC demonstrate that these two complexes have very similar conformations. The chemical shifts of corresponding residues are very similar, and essentially identical NOESY connectivities were observed. We were therefore able to use the spectra of TWJ-TC to resolve ambiguities in the assignments of TWJ-TT and, most notably, to sort out the assignments of the unpaired thymidines, S3-T6 and S3-T7.

One-dimensional spectra as a function of temperature

One-dimensional proton spectra of the aromatic region of junction TWJ-TT, acquired in D_2O solvent, are displayed as a function of temperature in Fig. 2. The spectra indicate that the complex is stable at 308 K or below. Below 298 K, the resonances become quite broad because of the increased viscosity of the solvent and increased dipolar interactions. Above 308 K the resonances broaden because of chemical exchange, indicating that the complex begins to dissociate (helix melting). Above 323 K the lines sharpen again, indicating complete strand dissociation. These effects correlate with those observed in the imino (downfield) region. The downfield region (12–15 ppm) of proton spectra of nucleic acids acquired in H_2O solvent reveals resonances attributable to imino protons protected from fast chemical exchange with solvent protons by hydrogen bonding as, for example, in base-pairing interactions. The imino proton region of one-dimensional NMR spectra of TWJ-TT is shown as a function of temperature in Fig. 3. Three resolved resonances corresponding to three base-paired thymidines appear in the most downfield region. The guanosine imino region is more crowded. The imino resonances broaden with increasing temperature in a concerted fashion and largely disappear by 318 K. The temperature effects seen in the imino proton region correlate with those seen in the aromatic region and, taken together, indicate that the complex dissociates in a cooperative manner to the three component strands. On the basis of the temperature studies, we chose to acquire 2D and 3D spectra between 298 and 303 K.

Examination of the methyl region of the 1D spectrum of TWJ-TT at 303 K reveals the chemical shifts of the five thymidine methyl resonances (1.42, 1.49, 1.52, 1.74, and 2.01 ppm). Four methyl resonances are observed for the TC junction, resonating at 1.45, 1.52, 1.71, and 1.98 ppm. This comparison permits the assignment of the methyl group of S3-T7 in TWJ-TT to the resonance at 1.49 ppm. Three aromatic resonances are well resolved and isolated in the extreme downfield part of the aromatic region. These arise from

TABLE 2 Proton chemical shifts for TWJ-TC at 30°C

	Arom	H2/5/Me	H1′	H2″	H2′	H3′	H4′	H5′/H5″	G/T imino	C amino
S1-G1	8.00		6.02	2.80	2.70	4.88	4.31	3.78		8.16, 6.52
S1-C2	7.57	5.40	6.13	2.58	2.19	4.87	4.31	4.26/4.15	13.01	8.35, 6.65
S1-T3	7.42	1.71	5.87	2.53	2.18	4.94	4.21	4.19	14.07	
S1-G4	7.97		6.01	2.81	2.69	5.05	4.47		12.95	8.46, 6.58
S1-C5	7.52	5.46	6.10	2.47	2.24	4.87	4.28	4.23/4.19	13.14	8.39, 6.67
S1-C6	7.49	5.45	5.54	2.35	1.92	4.85	4.24		12.75	8.23, 6.54
S1-A7	8.35	7.75	6.28	2.94	2.77	5.06	4.46	4.20/4.15	13.75	
S1-C8	7.36	5.36	5.88	2.42	2.04	4.82	4.32	4.22/4.11	12.92	8.14, 6.58
S1-C9	7.48	5.65	5.74	2.38	2.00	4.86	4.15		13.17	8.63, 6.97
S1-G10	7.98		6.22	2.44	2.67	4.73	4.39	4.14/4.09		
S2-C1	7.65	5.93	5.81	2.44	1.92	4.73	4.10	3.76		
S2-G2	7.96		5.69	2.81	2.74	5.06	4.37		13.17	8.63, 6.97
S2-G3	7.80		6.07	2.84	2.62	5.01	4.49	4.26	12.92	8.14, 6.58
S2-T4	7.22	1.45	5.92	2.56	2.16	4.96	4.30		13.75	
S2-G5	7.89		6.08	2.71	2.71	5.01	4.48	4.21/4.17	12.75	8.23, 6.56
S2-C6	7.50	5.48	5.93	2.58	2.20	4.97	4.15	4.13		
S2-G7	8.03		6.17	2.95	2.71	5.05	4.47	4.24	12.90	8.27, 6.54
S2-T8	7.35	1.52	6.25	2.60	2.17	4.94	4.21		13.79	
S2-C9	7.64	5.79	6.14	2.53	2.27	4.90	4.26	4.16	13.00	8.54, 7.00
S2-C10	7.66	5.75	6.27	2.33	2.33	4.61	4.09	4.13		
S3-G1	7.88		5.68	2.71	2.52	4.86	4.40	3.76/3.72		
S3-G2	7.90		5.59	2.84	2.75	5.07	4.41	4.30/4.22	13.00	8.54, 7.00
S3-A3	8.27	8.00	6.28	2.93	2.77	5.12	4.52	4.23/4.22	13.79	
S3-C4	7.28	5.34	5.67	2.36	1.97	4.83	4.22	4.15	12.90	8.27, 6.54
S3-G5	8.02		6.14	2.71	2.71	5.02				
S3-T6	7.75	1.98	5.92	2.66		4.94		4.28/4.13		
S3-C7	7.57	5.69	5.91	2.38	2.00	4.71				
S3-G8	7.76		6.09	2.73	2.68	5.04	4.38	4.21	13.14	8.39, 6.67
S3-C9	7.49	5.47	5.62	2.42	2.08	4.87	4.24	4.21/4.08	12.95	8.46, 6.58
S3-A10	8.23	7.76	6.09	2.92	2.75	5.09	4.44	4.18/4.10	14.07	
S3-G11	7.71		5.88	2.68	2.51	5.00	4.41	4.21/4.11	13.01	8.35, 6.65
S3-C12	7.38	5.27	6.16	2.25	2.18	4.50	4.09			8.16, 6.52

Samples were dissolved in D_2O or H_2O buffer consisting of 100 mM sodium chloride, 10 mM sodium phosphate, 5 mM magnesium chloride, and 0.5 mM EDTA.

the H8 protons of the three adenosine residues. The H6 resonance of S2-T4 (7.20 ppm) is resolved in the upfield part of the aromatic region. All of these resolved resonances are relatively sharp singlets, indicating that both TWJ complexes exist in unique configurations.

Evidence for the existence of three helical arms

The sequence was designed to include one AT basepair in each helical arm. The sequence at each AT basepair is unique, allowing unequivocal assignments to be made from examination of the well resolved adenosine H8 and thymidine methyl proton resonances in two-dimensional NOESY spectra carried out in D_2O solution. The two-dimensional NOESY spectra reveal the existence of three helical arms. Each helix contains one AT basepair. The 2D NOESY spectra show that all three expected AT basepairs form. The aromatic H6 resonances of each thymidine base are unambiguously assigned to their respective methyl resonances from the intense NOESY crosspeaks observed in the aromatic-methyl region. The assignments are presented in Tables 1 (TWJ-TT) and 2 (TWJ-TC). An obvious correspondence in chemical shift is observed between the H6/Me resonances of the two samples. Three of the thymidine methyl/H6 resonance pairs in each sample (those resonating at 1.43, 1.52, and 1.74 ppm in the TT junction and at 1.45, 1.52, and 1.71 ppm in the TC junction) exhibit the intra- and inter-nucleotide NOEs to aromatic, H1′, H2′, H2″, and H3′ protons characteristic of right-handed, B-form DNA. The relevant NOESY regions involving the methyl resonances for the two junctions are compared in Fig. 4. Stronger NOEs are observed between each of the three intra-helical thymidine methyl resonances and the sugar protons of the preceding (5′) nucleotides compared with their own sugar resonances, as expected for B-DNA (Wüthrich, 1986).

The sequence-specific assignments of the three intra-helical thymidines proved straightforward. They were made from the 2D NOESY spectra in Fig. 4. Only one thymidine H6 resonance (7.40 ppm) displays NOE to a cytosine H5. The methyl (1.52 ppm) to H5 (5.80 ppm) NOE is also observed. The thymidine with H6/Me resonances at 7.40 ppm/1.52 ppm can therefore be unambiguously assigned to S2-T8, the only thymidine that is 5′ to a cytosine. The only thymidine neighboring a cytosine on the 3′-side is the one having H6 resonating at 7.44 ppm and Me at 1.74 ppm, identifying it as S1-T3. This is proved by the observation of NOE between H5 of S1-C2 and Me of S1-T3 but not between CH5 and TH6. The upfield change in the chemical shift position of H5 of S1-C2 from 5.72 ppm in the TT-junction to 5.40 ppm

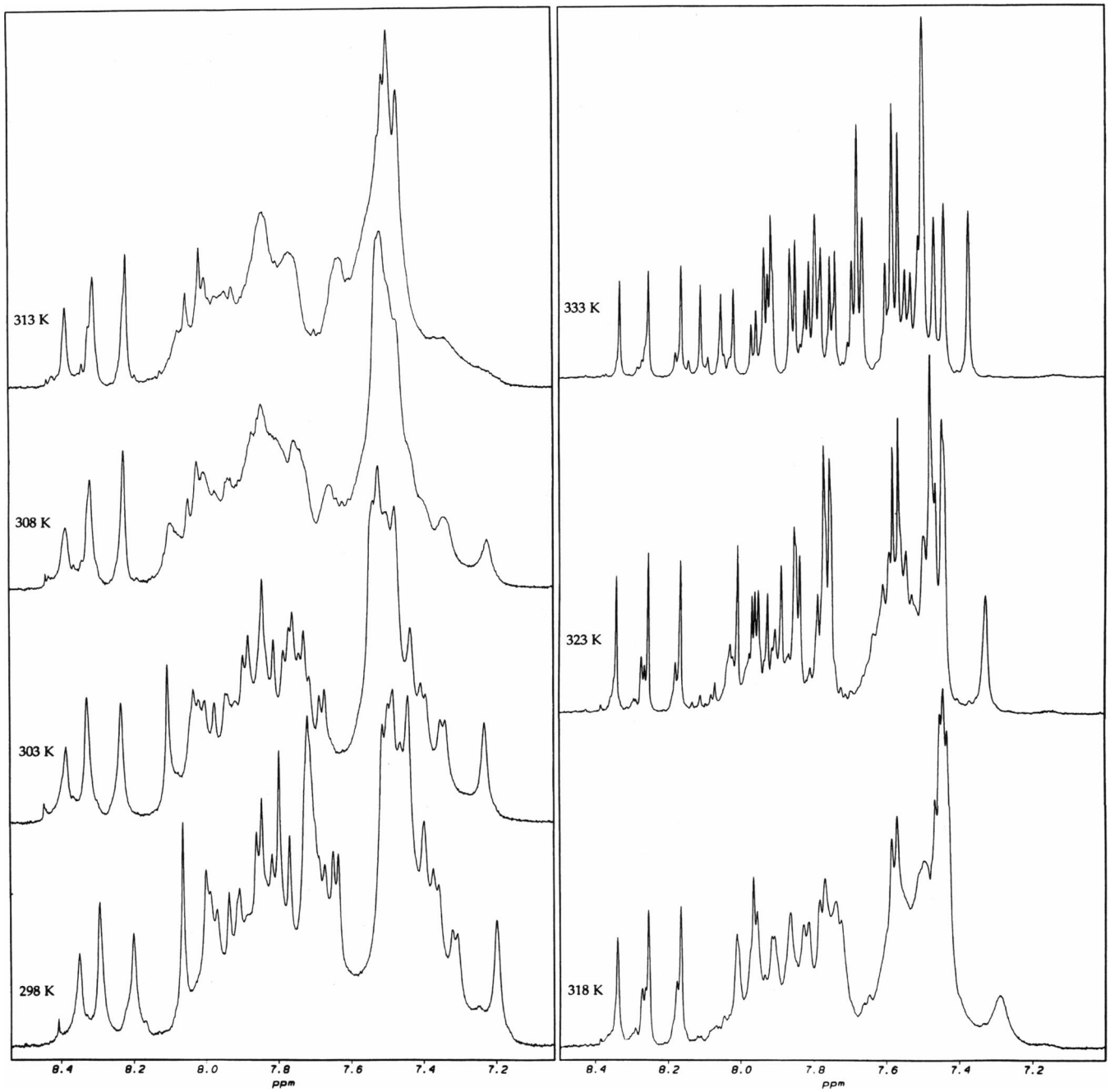

FIGURE 2 Aromatic region of one-dimensional proton NMR spectra of TWJ-TT acquired in D_2O solvent as a function of temperature. Spectra were processed with exponential multiplication (2 Hz).

in the TC-junction, which is due to the substitution of a guanosine at position 1 of S1, is revealed by the Me to H5 NOESY crosspeak. By a process of elimination, one deduces that the thymidine having H6/Me resonating at 7.20ppm/1.42 ppm is S2-T4.

The imino protons of the three intra-helical thymidines were assigned from 2D NOESY spectra acquired in H_2O using pulsed field gradients from NOEs to the thymidine methyl resonances (Leontis et al., 1993). The adenosine H2 protons of the corresponding base-paired adenosines could then be assigned by the intense NOEs observed to the corresponding imino resonances.

Because all three adenosine H8 resonances are resolved in the downfield aromatic region, one can identify and assign the H1′, H2′, H2″, and H3′ resonances of all three adenosines and their 5′ neighboring nucleotides from the 2D NOESY spectra. Sequence-specific assignments were made from the NOESY spectra. The relevant aromatic to sugar regions of a NOESY spectrum of the TWJ-TC are shown in Fig. 5. Only adenosine S1-A7 is flanked on both sides by cytosines. It is

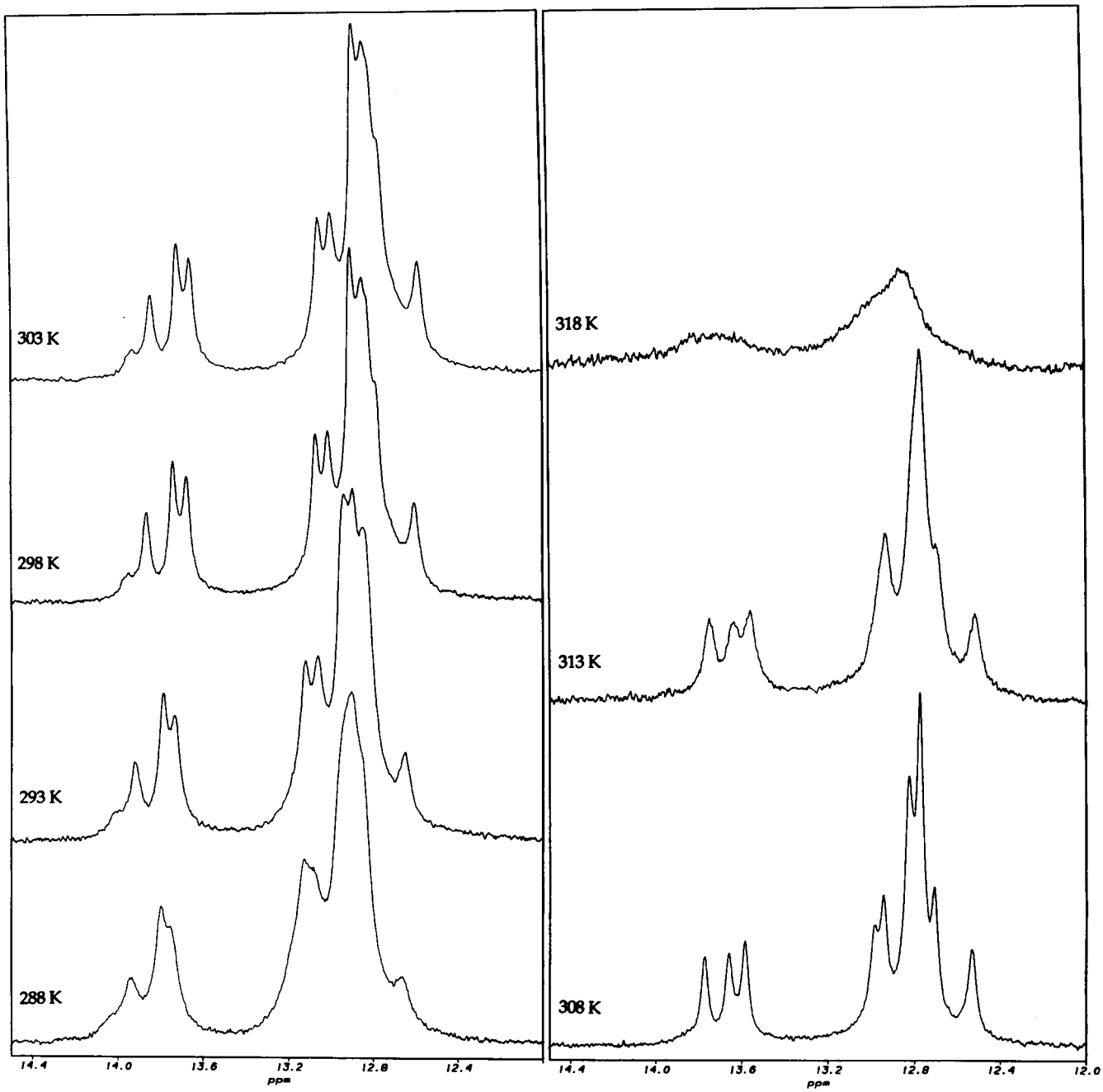

FIGURE 3 Imino region of one-dimensional proton NMR spectra of TWJ-TT acquired in H_2O solvent as a function of temperature. Spectra were processed with exponential multiplication (2 Hz).

identified by the NOE between its H8 (8.34 ppm) and H1′ of S1-C6 (5.55 ppm) and between its H8 and the H5 of S1-C8 (5.38 ppm). S3-A3 is identified by the NOE between its H8 (8.30 ppm) and the H5 proton of S3-C4 (5.37 ppm). The third adenosine, S3-A10, is identified by NOEs to the sugar protons of S3-C9.

The assignment of the intra-helical adenosine and thymidine resonances established that the complex did indeed form three helical arms and provided a starting point within each arm for sequentially assigning each strand in both the 5′ and 3′ directions. This was possible, once all of the cytosine H5/H6 and the sugar ring spin systems had been identified. The resolution and identification of all the spin systems were achieved using the 3D NOESY-TOCSY spectrum.

Assignment of cytosine H5-H6 NOESY crosspeaks

The cytosine H5-H6 NOESY crosspeaks for the two junctions are compared in Fig. 6. The data are plotted at a high contour level so that only the H5-H6 crosspeaks are displayed. These data show that the chemical shifts of the cor-

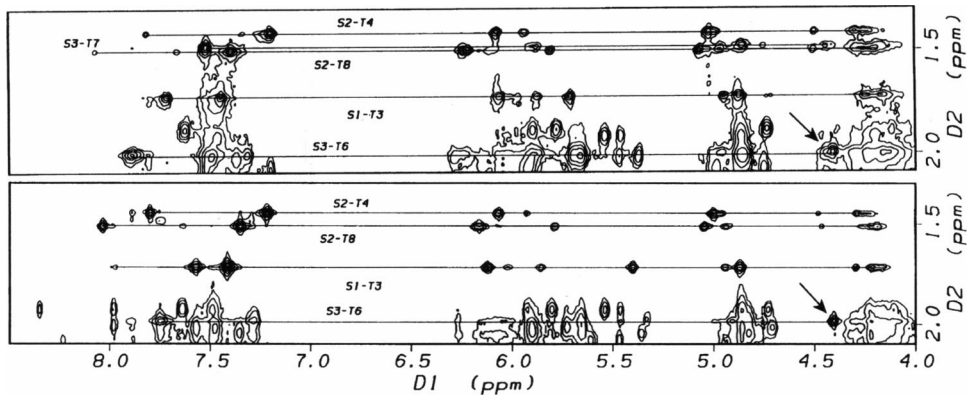

FIGURE 4 Selected regions involving methyl resonances from NOESY spectra acquired in D_2O at 30°C, 300 ms mixing time. (*top*) TWJ-TT. (*bottom*) TWJ-TC. The corresponding NOEs between the methyl group of S3-T6 and H4′ of S3-G11 are shown with arrows.

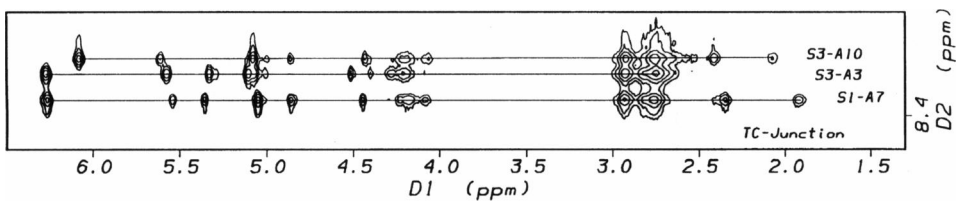

FIGURE 5 Selected regions involving adenosine aromatic resonances from NOESY spectra of TWJ-TC acquired in D_2O at 30°C, 300 ms mixing time.

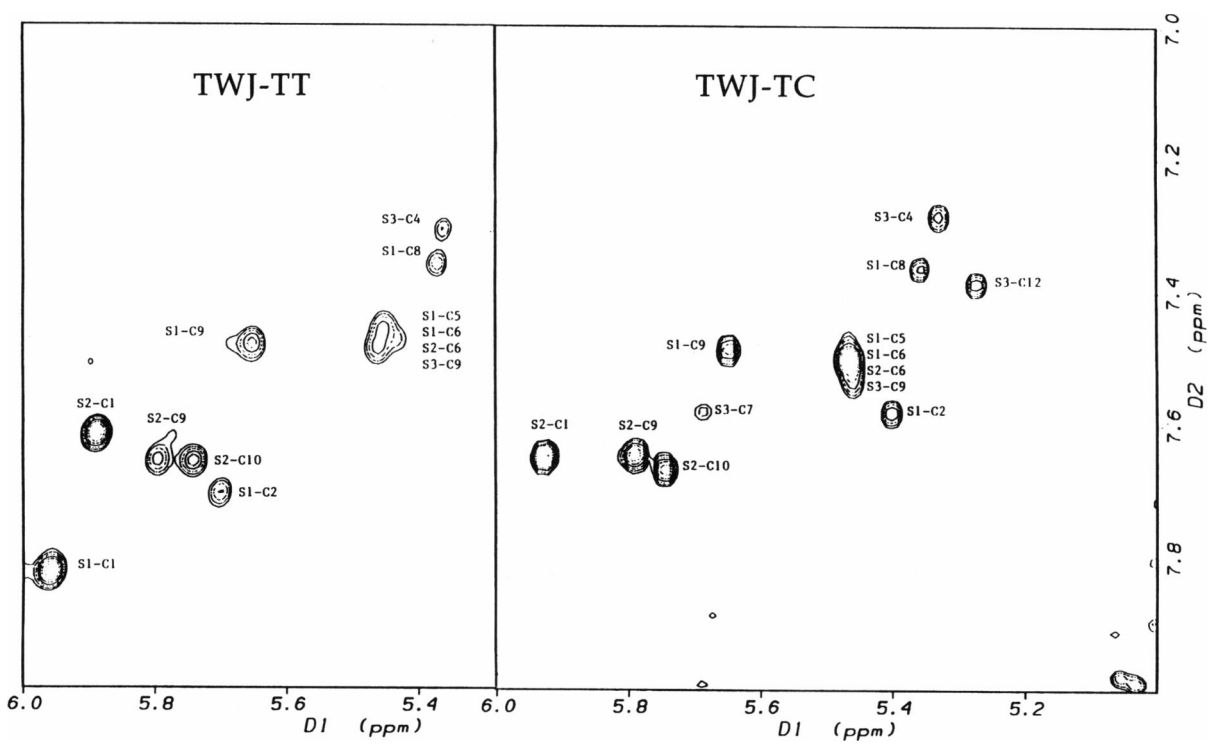

FIGURE 6 Comparison of CH6/CH5 NOESY crosspeaks in TWJ-TT (*left* panel) and TWJ-TC (*right* panel). The spectra were acquired in D_2O at 30°C with a 300 ms mixing time.

responding cytosine residues in the two samples are very similar. Significant differences are only observed for S1-C2, because of the change of basepair S3-G12/S1-C1 in TWJ-TT to S3-C12/S1-G1 in TWJ-TC. The unpaired cytosine S3-C7 can be assigned unambiguously from Fig. 6. Eight cytosines are well resolved in the spectrum of the TT junction and nine in that of the TC junction. The four remaining cytosines in each sample, S1-C5, S1-C6, S2-C6, and S3-C9, appear to share degenerate chemical shifts for both their aromatic H6 (7.46–7.51 ppm) and H5 (5.43–5.48 ppm) proton resonances. Three of these cytosines, S1-C5, S1-C6, and S3-C9, could be resolved in NOESY spectra collected in H_2O by virtue of differences in the chemical shifts of the imino resonances of the guanosines to which they are base-paired. Eight crosspeaks were

observed in this region, corresponding to all of the GC basepairs that exhibit observable imino resonances. Crosspeaks for S1-C1, S2-C1, S2-C6, and S2-C10 are not observed.

Stacking of S3-G8 and S2-G5

This interaction is shown by the observation of the inter-basepair imino-imino crosspeak in NOESY spectra acquired in H_2O, as has already been described (Leontis et al., 1993). In the NOESY spectrum of both TWJ-TT and TWJ-TC acquired in 90% H_2O, sequential NOEs were observed between hydrogen-bonded imino protons in all three helical arms. The imino proton resonances of the three terminal guanosines (S1-G10, S3-G1, and S3-G12) and of the junction nucleotides (S3-G5, S3-T6, and S3-T7) were not observable, presumably because these protons are accessible to and rapidly exchange with H_2O solvent. We did observe, however, the imino protons of S2-G5 and S3-G8, which also flank the junction. Moreover, we observed an inter-basepair imino-imino NOESY crosspeak linking these two bases. We also observed NOEs from the imino protons of S3-G8 and S2-G5 to the amino and H5 protons of S1-C5 and S1-C6, respectively, which proves that S3-G8 hydrogen-bonds S1-C5 and that S2-G5 hydrogen-bonds to S1-C6, as shown in Fig. 1. These data conclusively show that these base-pairs are stacked on each other.

Assignment of resonances belonging to junction residues

The assignment of resonances of individual residues flanking the junction region will be described in greater detail.

S1-C5

Because of the chemical shift degeneracy of H5 of S1-C5 and H5 of S1-C6, the NOEs between H5 of S1-C5 and H2′/H2″ of S1-G4 could not be resolved in 2D spectra but could in the 3D NOESY-TOCSY. In the 3D spectrum we can also see NOEs from H5 of S1-C6 to H2′ and H2″ of S1-C5 showing stacking across the junction. The H1′ of S1-C5 is obtained from the 2D spectrum. The NOE between H6 of S1-C6 and H1′ of S1-C5 and between H6 and H1′ of S1-C5 can be distinguished in the TWJ-TC spectra, although they partly overlap. The NOE between H6 of S1-C5 and H1′ of S1-G4 is well resolved, indicating stacking between these bases. Sequential aromatic to aromatic NOEs are observed between S1-C5 and S1-G4 and between S1-C6 and S1-A7, as are sequential imino to imino NOEs all along Strand 1.

S1-C6

This residue is 5′ to S1-A7. Well resolved sequential NOEs between H8 of S1-A7 and all of the sugar protons of S1-C6 as well as between the aromatic protons of the two residues are observed. The dinucleotide step between S1-C6 and S1-A7 appears to conform well to B-form DNA. In NOESY spectra recorded in H_2O with long mixing times, an NOE attributable to spin-diffusion is observed between the imino proton of S2-T4 (paired to S1-A7) and the H5 of S1-C6, further demonstrating the stacking interaction between S1-C6 and S1-A7.

S2-G5

The aromatic H8 proton of S2-G5 is assigned by its NOE to H6 of S2-T4. The sequential NOEs between H8 of S2-G5 and its own H1′ and the H1′ of S2-T4 are well resolved in 2D spectra. There is no evidence for the sequential NOE between S2-G5 H1′ and the aromatic H6 proton of S2-C6. The observation of NOEs between the imino protons of S2-G5 and of S3-G8 in H_2O spectra of both TWJ-TT and TWJ-TC provides crucial confirmation of the stacking of these two bases across the junction. The expected NOEs between the imino proton of S2-G5 and the amino and H5 protons of S1-C6 provide further confirmation of base pairing between S2-G5 and S1-C6.

S2-C6

Clear evidence for the stacking of S2-C6 on S2-G7 is shown by sequential NOEs observed between H8 of S2-G7 and H1′, H2′, H2″, H3′, and H4′ of S2-C6. Intra-residue H6 to sugar proton crosspeaks for this residue are hard to resolve because of overlap with the aromatic H6 protons of three other cytosines. The aromatic to aromatic crosspeak between S2-C6 and S2-G7 is not observed, although the corresponding crosspeak between S2-G7 and S2-T8 is. The imino proton of S3-G5, which is presumably base-paired with S2-C6 is also not observed. This does not mean that this basepair does not form. The imino protons of terminal basepairs usually exchange too fast with solvent protons to give distinct resonances in the downfield (12 to 15 ppm) region of the proton spectrum.

S3-G5

The H8 of the S3-G5 residue is identified by NOE to the H1′ of S3-C4 (5.67 ppm). This NOE is fairly weak, but well resolved. This may be due to motional flexibility of this residue. The H5/H6 resonances of S3-C4 are easily assigned from well resolved NOEs to the sugar protons of S3-A3, as discussed above. The sugar protons of S3-C4 were assigned from NOEs to H6. The H1′ of S3-G5 is found at 6.12 ppm by NOE to its own H8. Inter- and intra-residue NOEs involving S3-G5 H8 and sugar H1′ protons are also observed in the spectrum of TWJ-TT. NOEs between S3-G5 H8 and its own H2′/H2″ cannot be resolved because of possible overlap with H8 of S2-G7. It is likely that S3-G5 H2′ and H2″ resonate at 2.71, at the chemical shift of the S2-G7 H2′ resonance. NOEs between S3-G5 H8 and the H3′, H2′, and H2″ of S3-C4 are not observed even at 200 ms. No aromatic-aromatic NOE is observed for S3-C4 and S3-G5. The imino proton of S3-G5 is not observed in NOESY spectra acquired

in H_2O. This is consistent with the structural model because S3-G5 is the terminal base in helix 3.

S3-T6 and S3-T7/C7

The next two residues exhibit a richer set of NOEs. The aromatic and methyl resonances were assigned as described above. Intra-residue crosspeaks between S3-T6 H6 and its own H1′, H2′, H2″, and H3′ protons are observed in TWJ-TT and TWJ-TC spectra. Likewise, the H6, H5, H1′, H2′, H2″, and H3′ resonances of S3-C7 in TWJ-TC were assigned from intra-residue crosspeaks involving H6. A crucial inter-residue crosspeak is observed between S3-T6 H6 and S3-C7 H5. For TWJ-TT junction we identified the chemical shift of S3-T7 H6 using the crosspeak between the methyl group of this base and H6. Analysis revealed only NOE peaks between H6 and H1′ and H3′ for this nucleotide. We have checked our assignment concerning this nucleotide in a 3D-NOESY-NOESY spectrum where we found a Me > H6 > H1′ 3D crosspeak.

The 2D NOESY data reveal that for two of the thymidines in the TT junction and one of the thymidines in the TC junction, the inter-nucleotide NOEs from the methyl groups to the aromatic, H1′, H2′, and H2″ atoms of the preceding base are not observed. From this we conclude that these bases are looped out rather than stacked in the junction.

S3-G8

The H8 resonance of S3-G8 was assigned by comparing the TWJ-TT and TWJ-TC NOESY spectra in D_2O to find the equivalent NOE crosspeak between S3-G8/H8 and S3-C9/H5. The observation of this crosspeak provides further proof of the stacking of S3-G8 and S3-C9. The 2resonance of H8 of S3-G8 overlaps that of H6 of S3-T6 in TWJ-TC spectra. But the NOE observed between the aromatic resonance at the chemical shift shared by H8 of S3-G8 and H6 of S3-T6 and the H5 of S3-C9 cannot be due to S3-T6 H6 because only S3-G8 is close enough to S3-C9.

We see NOEs from the imino of S3-G8 (13.12) to the S1-G4 imino (12.95) and to the S2-G5 imino (12.75) in both TWJ-TT and TWJ-TC NOESY spectra acquired in H_2O. We also see strong NOEs from the imino of S3-G8 to the amino protons of S1-C5 (8.39/8.37 and 6.67/6.61) in both samples. NOEs from the S3-G8 imino to the aromatic (H6) and H5 resonances of S1-C5 are also observed. These are important observations because the H5 and H6 resonances of S1-C5 are degenerate with those of at least two other cytosines. However, the imino protons of the Gs to which they are base-paired are all well resolved, allowing one to identify these contacts unambiguously.

In our models, in which the base of S3-G8 lies in the syn conformational range, we observe the following intra-residue distances for S3-G8: H8 to H1′, 3.3 Å; H8 to H2′, 3.0 Å; H8 to H2″, 3.0 Å. The H8-H1′ distance in S3-G8 is shorter than the average distance of 3.9 Å observed for the other purines in the complex, which all have glycosidic angles in the *trans* conformation. The distance is consistent

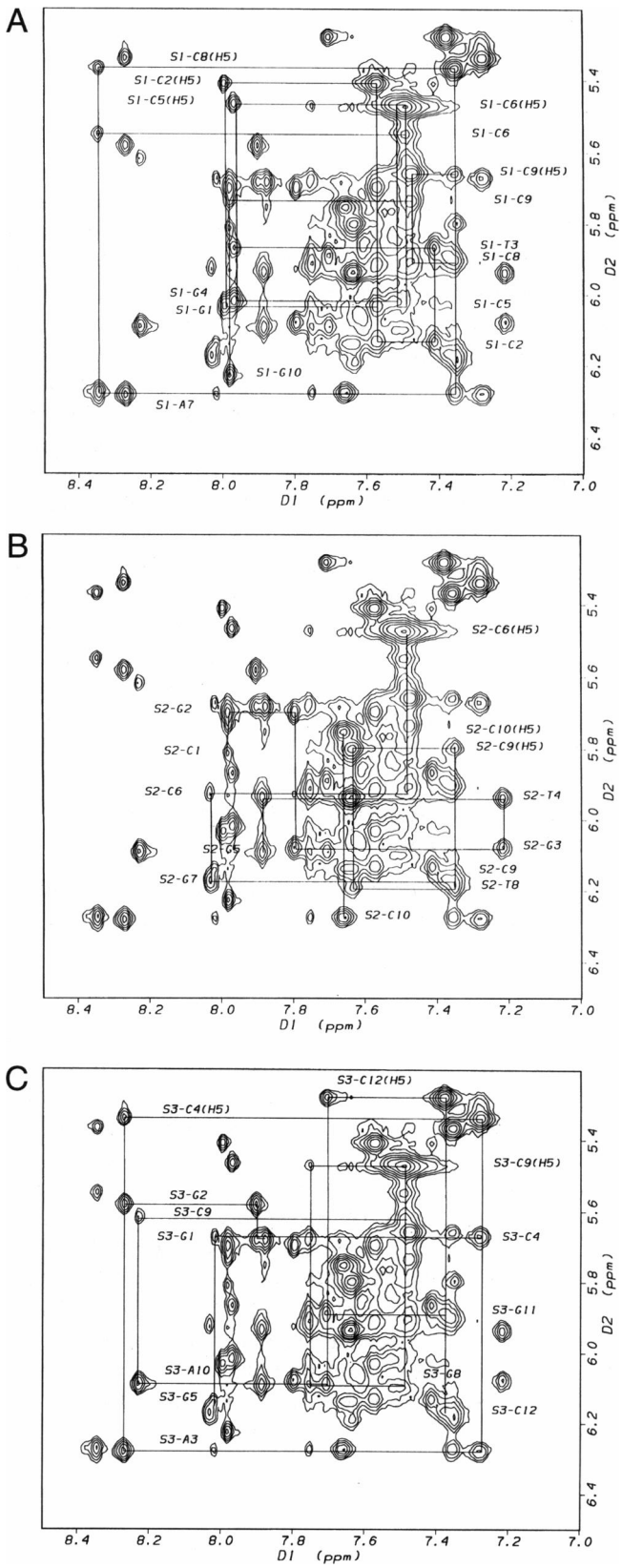

FIGURE 7 Sequential NOEs ("NOESY walks") in aromatic-H1′/H5 region of TWJ-TC acquired in D_2O at 30°C with a 300 ms mixing time. (*A*) Strand 1 connectivities. (*B*) Strand 2 connectivities. Panel C: Strand 3TC connectivities.

with the intensity of the NOE assigned to the H8 to H1′ interaction in S3-G8. However, it is longer than the distances typically observed for the guanosines in the syn conformation. This is discussed in more detail in the companion paper. The H8-H2′ distance is longer than in typical B-DNA with the base in the anti configuration, consistent with our failure to observe a strong H8 to H2′ NOE for S3-G8.

Comparison of chemical shifts of TWJ-TT and TWJ-TC

The chemical shifts of proton resonances are collected in Tables 1 and 2. It is apparent that the chemical shifts of corresponding resonances are very similar between the two samples. This provides strong evidence that the conformations of the two junctions are very similar.

Sequential connectivities in all three strands

By combining all the spectral information it is possible to trace sequential connectivities in the aromatic to sugar regions of the NOESY spectra, to generate NOESY walks for each of the three strands. NOESY walks in the aromatic to H1′ region of a 300 ms spectrum acquired in D_2O solvent are shown for the TC-junction in Fig. 7. Sequential connectivities can be followed along the entire length of Strand 1. The NOESY walks in Strands 2 and 3 are interrupted by the junction. As discussed in our first report on the NMR spectroscopy of these junctions, seven independent, sequential networks of NOEs can be traced through the spectra (Leontis et al., 1993). One network consists of the imino, amino, and adenosine H2 protons of the bases of S1 and of the bases to which they are paired on the 5′-half of S2 and the 3′-half of S3TT. A second network consists of the aromatic, CH5, methyl, and sugar protons of S1. Inter-residue NOESY crosspeaks are observed between H5 of S1-C6 and the sugar protons of S1-C5, indicating continuous stacking along S1 across the junction region. The aromatic, CH5, methyl, and sugar protons of the 5′-half of S2 and of the 3′-half of S3TT constitute two more networks. These four networks are interconnected by 1) intra-residue NOEs between the cytosine H5 and amino ^{1}H resonances and between the thymidine imino and methyl ^{1}H resonances, and by 2) intra-basepair NOEs between the guanosine imino and cytosine amino and H5 resonances. Three independent networks of NOE connectivities are observed for the resonances attributable to the 3′-end of S2 and the 5′-end of S3TT. These consist of the aromatic, CH5, methyl, and sugar resonances of 1) the 3′-half of S2, 2) the 5′-half of S3TT, and 3) the imino, amino, and AH2 resonances of the base-pairs S3-G2/S2-C9, S3-A3/S2-T8, and S3-C4/S2-G7. A cross-strand NOE is observed between H2 of S3-A3 and H1′ of S2-C9.

Identification of tertiary contact involving the methyl group of unpaired S3-T6

The structural model depends critically on the correct assignment of the sugar resonance at 4.40 ppm, which is in NOE contact with the methyl group of unpaired S3-T6. This gives rise to a distinctive and well resolved crosspeak at 4.41/1.98 ppm in the TWJ-TC spectrum (4.38/2.01 ppm in TWJ-TT spectra). As mentioned earlier, the methyl resonance of S3-T6 was assigned by comparison between the spectra of TWJ-TT and TWJ-TC. The identification of the contact point of this methyl group as the H4′ of S3-G11 rests on careful comparison of corresponding regions in the 2D NOESY spectra of TWJ-TT and TWJ-TC.

Because the aromatic resonance of S3-G11 is fortuitously well resolved in the spectrum of TWJ-TC, we will focus on NOESY spectra of this sample. In spectra of TWJ-TC with mixing times ranging between 60 and 300 ms, three strong crosspeaks are observed between the aromatic region and resonances at 4.40 ppm: these have aromatic chemical shifts at 7.71, 7.75, and 7.90 ppm. Two of these are due to intra-residue H8/H4′ crosspeaks, S3-G2 (7.90 ppm) and S3-G11 (7.71). We assigned the third to inter-residue contact between H6 of unpaired S3-T6 and H4′ of S3-G11. As mentioned, there is only one aromatic resonance at 7.71 ppm, and it was unambiguously assigned to the H8 of S3-G11 by sequential crosspeaks to the H5 of S3-C12 (5.27 ppm), to H1′ of its own sugar (5.88 ppm), and to H1′ of S3-A10. The H8 resonance of S3-G2 could be assigned from sequential NOEs involving the sugar protons of S3-G2 and S3-G1. The chemical shift of the H1′ resonance of S3-G2 was readily identified as 5.59 ppm. The assignments of the 4.40 ppm resonances to the H4′ protons of S3-G2 and S3-G11 were verified by weak spin-diffusion crosspeaks, observed in the 300 ms NOESY spectrum, to H6 of S3-C12, presumably via H1′ and H2″ of S3-G11 and to H8 of S3-A3. The (4.40/7.76 ppm) crosspeak was assigned to the H6 of S3-T6.

Cross-correlating the H4′/aromatic and the H4′/methyl regions with the H4′/H1′ region, we observed the expected strong intra-sugar H1′/H4′ crosspeaks for S3-G11 and S3-G2 and only those resonances at 4.40 ppm. A small difference of less than 0.01 ppm in the chemical shift of the H4′ coordinate was observed in some spectra. On both sides of the diagonal of these spectra, the H4′ coordinate of the crucial methyl/H4′ crosspeak (1.98/4.41 ppm) aligns exactly with the intra-sugar H1′/H4′ crosspeak of S3-G11 and not with that of S3-G2.

If the S3-T6 base is in fact located near H4′ of S3-G11, the H6 and methyl protons would be expected to lie close to the H5′/H5″ protons as well. The H5′/H5″ resonances of S3-G2 and S3-G11 were identified with reasonable certainty by cross-correlating the aromatic-sugar and the sugar-sugar regions of NOESY spectra. Examining the H2′/H2″ to H4′/H5′/H5″ region of the 60 ms spectrum of TWJ-TC, we observed the expected intra-residue crosspeaks, H2′/H2″ to H4′ and H2′/H2″ to H5′/H5″ NOEs are observed for both residues. This permits complete assignment of the spin systems of both S3-G2 and S3-G11 as shown in Tables 1 and 2. The resonances are partly distinct: 4.30 and 4.22 ppm for S3-G2, and 4.23 and 4.11 ppm for S3-G11. At long mixing times, one observes crosspeaks from the methyl and H6 resonance of S3-T6 to the chemical shift positions of the H5′/H5″ of S3-G11 rather than those of S3-G2. Concerning NOEs from

the S3-T6 methyl resonance to other sugar protons of S3-G11, at long mixing times we observe a weak interaction with H3′ of S3-G11 (5.00 ppm). No corresponding peak is seen at the chemical shift of the H3′ of S3-G2 (5.07 ppm), providing further support for the assignment. The expected NOE crosspeak between the S3-T6 methyl and H1′ of S3-G11 cannot be observed because of overlap due to the H1′/H2′ crosspeaks of S3-C7 and S1-C8.

In TWJ-TT spectra we can resolve and identify the H1′/H4′ crosspeaks of S3-G2, S3-G1, S3-G11, and S3-A10. These crosspeaks all share the same H4′ coordinate (between 4.39 and 4.42 ppm), but can be distinguished by their H1′ chemical shifts. Because the linewidths are wider in spectra of the TT junction, these peaks are harder to resolve. The S3-T6 H6/Me resonance occurs at 7.88/2.01 ppm. The crucial H4′/Me crosspeak occurs at 4.38/2.01 ppm.

Construction of a qualitative structural model

Atomic structural models for TWJ-TT and TWJ-TC were constructed using the QUANTA molecular modeling software (Molecular Simulations Inc., Waltham, MA). Because the NMR data suggested that the helical stems were B-type, right-handed duplexes and that Helices 1 and 2 were coaxially stacked, a 10 basepair ideal B-form DNA helix corresponding to the sequences of Helix 1 and 2 was constructed. Another ideal B-form DNA helix corresponding to Helix 3 was constructed with a two-base overhang consisting of the unpaired pyrimidine bases of Strand 3.

The longer helix was nicked at the position corresponding to the dinucleotide step between S2-G5 and S3-G8. The two helices were manually positioned with respect to each other. Small adjustments were made to the dihedral angles at the 5′ end of S2-C6 to facilitate attachment to S2-G5 at the midpoint of the longer helix. The attachment on the other side of the shorter helix, of S3-T7 to S3-G8, was guided by the NOE contact between the methyl group of S3-T6 and the sugar ring of S3-G11 in the minor groove at the H4′ position. This required that the phosphodiester backbone of the unpaired bases S3-T6 and S3-T7 pass back along the minor groove of Helix 1 into the junction region to connect with S3-G8. This creates a fairly tight turn at the S3-T7 and S3-G8 dinucleotide step, requiring a syn conformation for S3-G8. The necessary contact with S3-G11 was established by placing S3-T6 in a syn conformation. Although no distortion of Helix 3 was required to establish this contact, dihedral angles in the two unpaired Ts were adjusted to minimize steric clashes at the junction and to maintain the phosphodiester backbone. Because strong tertiary NOE contacts were not observed for S3-T7, it was left exposed to the solvent.

Energy minimization with CHARMm (Brooks et al., 1983) was used to clean up the manually built structure. The full CHARMm 21.3 force field was used. Rather than explicitly including hydrogen bonds, electrostatic interactions with full partial charges were used with a distance-dependent dielectric constant. 42 NOE constraints involving G and T imino, C amino, adenosine H2, and the S3-T6 methyl protons were included in the force field using a flat well-type potential function with harmonic walls. Because energy refinement was applied only to relieve structural distortions, only gentle energy minimization was used (two cycles of steepest descent with 500 steps each while increasing the force constants for the NOE constraints after the first cycle). This preliminary model was published recently (Leontis et al., 1993) and is the basis (the starting structure) for the refinement procedures described in the next paper. For the purposes of this paper, it provides a qualitative description of the global structure of TWJ-TT and TWJ-TC.

DISCUSSION

We have demonstrated unequivocally the stable coaxial stacking of helices 1 and 2 in TWJ-TT and TWJ-TC. We have further shown that all expected basepairs flanking the junction region form. The S3-G5/S2-C6 basepair may be less stable or more flexible than the others, but this should not be surprising because it is a terminal basepair, not involved in a stacking interaction across the junction. We have shown that all three helical arms TWJ-TT and TWJ-TC form stable, B-form helices. Finally, we have demonstrated with a reasonable degree of certainty that the unpaired base S3-T6 contacts helix 1 in the minor groove at S3-G11. The assignment we have made allows one to build easily a reasonable model without greatly distorting any of the component helices of the junction complex and maintaining the important base-stacking interactions across the junction.

Comparison with the structure of J3CC

All of our data indicate that TWJ-TT and TWJ-TC exhibit very similar structures. Therefore, in what follows we will only refer to TWJ-TC, the structure of which we have further refined as described in the companion paper. But it should be understood that these statements also apply to TWJ-TT. Our structure for TWJ-TC bears superficial resemblance to those determined by Rosen and Patel (1993b) for the three-way junction complex, J3CC: coaxial stacking is observed between two of the three helices; the unpaired pyrimidine bases are extrahelical and exposed to solvent; one unpaired pyrimidine contacts the minor groove of one of the helical arms. The stacking arrangements in TWJ-TC and J3CC are compared schematically in Fig. 8. J3CC is redrawn from Rosen and Patel to facilitate comparison with TWJ-TC. The figure makes clear the difference in the stacking arrangement between the two junctions, J3CC and TWJ-TC. In J3CC, it is the strand equivalent to Strand 2 of TWJ-TC, with respect to the position of the unpaired bases, which is continuously stacked. Thus, Stems I and III of J3CC (using the nomenclature of Rosen and Patel) are coaxially stacked in J3CC. With regard to the position of the unpaired bases, Stem I of J3CC is equivalent to Helix 3 of TWJ-TC, whereas Stem II of J3CC is equivalent to Helix 1 of TWJ-TC. The excluded (nonstacked) helix in J3CC is Stem II. However, in TWJ-TC, it is Helices 1 and 2 that are coaxially stacked and Helix 3

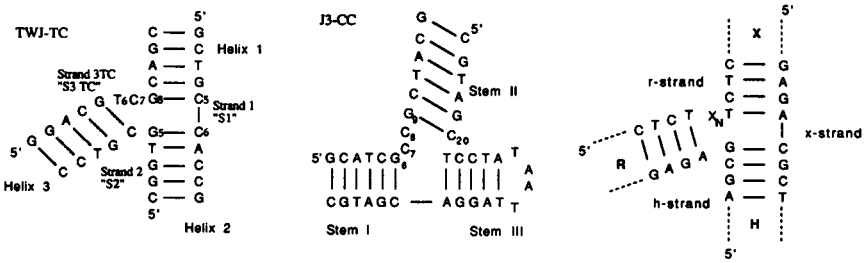

FIGURE 8 Comparison of qualitative structural features of three different TWJ systems showing the preferred coaxial stacking elucidated for each. (*left*) TWJ-TC from our work showing stacking of Helix 1 and Helix 2. (*middle*) J3CC from the NMR work of Rosen and Patel (1993b) showing coaxial stacking of Stem I and Stem III. (*right*) Stacking of Helix X and Helix H, as deduced from gel electrophoresis experiments as shown in Welch et al. (1993).

that is the excluded helix. In the models of J3CC presented by Rosen and Patel, Stem II forms a fairly acute angle with Stem III. The unpaired base, C7, of J3CC, which is equivalent to S3-T6 in TWJ-TC, is in contact with the minor groove of Stem I, equivalent to Helix 3. Our data indicate that S3-T6 is most likely in contact with the minor groove of Helix 1, not Helix 3. The phosphodiester backbone at the unpaired bases of J3CC follows a path across the top of the excluded helix, such that unpaired base C8, equivalent to S3-T7/C7, stacks on top of basepair G9:C20. This basepair is equivalent to S3-G8:S1-C5 of TWJ-TC. The reconnection of the excluded helix is made without a tight hairpin-like turn as in our model for TWJ-TC.

The base sequence of J3CC is very similar to that of TWJ-TC in the region flanking the junction. The only differences in the residues immediately surrounding the junction are the substitution of C or methylated C in J3CC at the position equivalent to S3-T6 in TWJ-TC and the substitution of a TA basepair (T21-A34) for the CG base-pair S1-C6/S2-G5 in TWJ-TC. Thus, the same pattern of purines and pyrimidines exists in J3CC and TWJ-TC. Nonetheless, it is clear from the above discussion that a different stacking arrangement exists in the two complexes. Before further discussing the factors that may play decisive roles in determining the preferred stacking conformations of three-way junctions, we review the results of gel electrophoretic studies of another series of three-way junctions containing unpaired bases.

Comparison with TWJ complexes studied by gel electrophoresis

Welch et al. (1993) analyzed the global structures of DNA three-way junctions containing unpaired bases using the gel electrophoretic method first applied by Cooper and Hagerman (1987) to four-way DNA junctions. Three-way junctions were constructed having extended helical arms and unique restriction sites near the junction region on each arm. Each arm was shortened in turn using the appropriate restriction enzyme, and the electrophoretic mobilities of the resulting asymmetric complexes were compared. Conclusions were drawn regarding the relative angles between the junction arms under the assumption that complexes in which the two extended arms subtend acute angles exhibit lower electrophoretic mobilities than those that subtend obtuse angles. Using this method, Welch et al. (1993) found evidence for unique stacking of three-way junctions having the core sequence shown in Fig. 8 (*right* panel), drawn in such a way as to facilitate comparison with TWJ-TC/TT. They concluded that the most probable coaxial stacking arrangement is that shown, in which Helix H stacks on Helix X, using their nomenclature. Furthermore, they concluded that as the number of unpaired bases increased from 1 to 5 (X_N), the angle between helices H and R became more acute, whereas that between helices H and X became more obtuse. Similar results were obtained when thymidines or adenosines were used as the unpaired nucleotides, X_N. Fig. 8 shows that the "x-strand," which is equivalent to our Strand 1 in relation to the position of the unpaired bases, is continuously stacked. The junctions of Welch et al. (1993), therefore, exhibit the same stacking arrangement as we have deduced for TWJ-TT and TWJ-TC, with respect to the position of the unpaired bases. This is despite the fact that the sequences at the junction site are very different in the two junctions.

Comparison of these three junction structures, two of which were examined with high resolution NMR methods, shows that all three conform to the general rule of nucleic acid folding enunciated by Westhof and co-workers that states that in multi-stem junctions the strand with the least number of unpaired connecting nucleotides between helical sections should become part of the continuous double-helix when there is a competition among helices for collinear stacking (Baudin et al., 1991). In the case of "simple" junctions in which only one strand has unpaired (connecting) bases, this still leaves two possible stacking arrangements, along each of the other two strands.

Clearly, the work of three separate groups on three different TWJ systems shows that coaxial stacking is an important driving force in organizing nucleic acid structure. And clearly, other factors besides the relative positioning of the unpaired bases are involved in determining the preferred stacking arrangement in TWJs. These factors may include:

1. The nature of the unpaired bases. Rosen and Patel (1993a) presented preliminary evidence indicating that junctions with unpaired purines (adenosines or inosines) have very different structures from J3CC. Their data seem to indicate that the unpaired purines are inserted within the junction itself.
2. The number of unpaired bases may be a factor. Welch et al. (1993) did not see a change in stacking preference using gel electrophoresis as they varied the number of unpaired bases between 1 and 5. However, this

is a low resolution technique. Definitive answers to this question will require more NMR work.

3. The base sequence of the basepairs immediately flanking the junction region. This may be the determinative factor that causes J3CC and TWJ-TC to assume different conformations. If so, then it appears that a seemingly minor change in sequence, in this case the substitution of a CG basepair by a TA basepair, is sufficient to cause a major change in conformation.
4. Longer range effects caused by sequence variation in the helical arms. These may be communicated to the junction region by changes in the helical parameters or by tertiary contacts involving unpaired bases, as observed in J3CC and TWJ-TC.
5. Solvent conditions, in particular salt concentration effects. It should be noted that J3CC was studied in 200 mM NaCl in the absence of divalent cations, whereas TWJ-TC and TWJ-TT were studied in 100 mM NaCl containing 5 mM Mg^{2+}. The possibility that the difference in ionic conditions employed by the two laboratories played some role in favoring one junction conformation over another cannot be excluded. Conversely, it is possible that by changing ionic conditions a particular junction (J3CC or TWJ-TC) can be induced to switch to the alternative conformation involving stacking along the other available strand. Without exception, all laboratories working on TWJ structures have noted that at low salt concentrations and in the absence of divalent cations, TWJ complexes assume flexible, unstructured conformations in which coaxial stacking at the junction appears to be completely disrupted, presumably by electrostatic repulsions between the three converging polyelectrolyte phosphodiester chains.

The answer to the general question, what determines the preferred coaxial stacking in the simplest multi-stem loops found in nucleic acids, three-way junctions, will require further work. In any case the factors listed above should be explored. Comparison of the structures that have already been determined, at least as regards coaxial stacking, provides a starting point for planning further investigations.

As a final point of discussion, it should be mentioned that the energetic difference between alternative stacking arrangements may not be very large and, therefore, the possibility exists that three-way junctions occurring in biologically active molecules are conformationally flexible and that dynamic equilibria exist between alternative stacking arrangements. In fact, dynamic flexibility may be required for the function of these structures. The work summarized above on DNA TWJs shows that in the case of simple junctions having only two unpaired bases, one coaxial stacking conformation is favored under a given set of conditions. But it may be that in naturally occurring three-way junctions, such as in 5S rRNA, which usually have more than two unpaired bases distributed among all three strands, more than one conformation may be energetically accessible under physiological conditions or as a result of interactions with other molecules.

The authors thank Dr. Florence Lebreton for the 3D NOE-NOE pulse sequence and Dr. Dean Carlson for general technical assistance with NMR spectroscopy, and M. Rosen for electronically communicating a computer model of J3CC.

This work was supported by U.S. Public Health Service grants GM41454 (to N. B. Leontis) and AI27744 (to D. G. Gorenstein), by Ohio Board of Regents Research Challenge grants, by Petroleum Research Fund grant 20871-GB4, and by Research Corporation grant C-2314 (to N. B. Leontis). The Purdue University Biochemical Magnetic Resonance Center, where most of the spectroscopic work was carried out, is supported by the National Institutes of Health designated AIDS Research Center at Purdue (grant AI727713 to D. G. Gorenstein) and by the National Science Foundation Biological Facilities Center on Biomolecular NMR, Structure, and Design at Purdue (grants BBS 8614177 and 8714258 to D. G. Gorenstein from the Division of Biological Instrumentation).

REFERENCES

Baudin, F., P. J. Romaniuk, P. Romby, C. Brunel, E. Westhof, B. Ehresmann, and C. Ehresmann. 1991. Involvement of "hinge" nucleotides of *Xenopus laevis* 5S rRNA in the RNA structural organization and in the binding of transcription factor TFIIIA. *J. Mol. Biol.* 218:69–81.

Baudin, F., P. J. Romaniuk, P. Romby, C. Brunel, E. Westhof, C. Ehresmann, and B. Ehresmann. 1991. Involvement of "hinge" nucleotides of *Xenopus laevis* 5S rRNA in the RNA structural organization and in the binding of transcription factor TFIIIA. *J. Mol. Biol.* 218: 69–81.

Brooks, C. L., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187.

Brunel, C., P. Romby, E. Westhof, C. Ehresmann, and B. Ehresmann. 1991. Three-dimensional model of *Escherichia coli* ribosmal 5S RNA as deduced from structure probing in solution and computer modeling. *J. Mol. Biol.* 221:293–308.

Chen, S. M., and W. J. Chazin. 1994. Two-dimensional ^{1}H NMR studies of immobile Holliday junctions: nonlabile proton assignments and identification of crossover isomers. *Biochemistry.* 33:11453–11459.

Cooper, J. P., and P. J. Hagerman. 1987. Gel electrophoretic analysis of the geometry of a DNA four-way junction. *J. Mol. Biol.* 198:711–719.

Duckett, D. R., A. I. H. Murchie, A. Bhattacharyya, R. M. Clegg, S. Diekmann, E. von Kitzing, and D. M. J. Lilley. 1992. The structure of DNA junctions and their interactions with enzymes. *Eur. J. Biochem.* 207: 285–295.

Egebjerg, J., J. Christiansen, R. S. Brown, N. Larsen, and R. A. Garrett. 1989. Protein L18 binds primarily at the junctions of helix II and internal loops A and B in *Escherichia coli* 5S RNA: implications for 5S RNA structure. *J. Mol. Biol.* 206:651–668.

Flemming, M., B. Deumling, and B. Kemper. 1993. Function of gene 49 of bacteriophage T4 III. Isolation of Holliday structures from very fast-sedimenting DNA. *Virology.* 196:910–913.

Forster, A. C., and R. H. Symons. 1987. Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell.* 49:211–220.

Guthrie, C., and B. Patterson. 1988. Spliceosomal snRNAs. *Annu. Rev. Genet.* 22:387–419.

Holliday, R. 1964. A mechanism for gene conversion in fungi. *Genet. Res.* 5:282–304.

Jensch, F., and B. Kemper. 1986. Endonuclease VII resolves Y-junctions in Branched DNA in vitro. *EMBO J.* 5:181–189.

Leontis, N. B., M. T. Hills, M. E. Piotto, A. Malhotra, J. M. 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.

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

Leontis, N. B., M. E. Piotto, M. T. Hills, A. Malhotra, I. V. Ouporov, J. M. Nussbaum, and D. G. Gorenstein. Structural studies of DNA three-way junctions. *Methods Enzymol.* In press.

Lilley, D. M. J., and R. M. Clegg. 1993a. The structure of branched DNA species. *Q. Rev. Biophys.* 26:131–175.

Lilley, D. M. J., and R. M. Clegg. 1993b. The Structure of the four-way junction in DNA. *Annu. Rev. Biophys. Biomol. Struct.* 22:299–328.

Lu, M., Q. Guo, and N. R. Kallenbach. 1992. Interaction of drugs with branched DNA structures. *Crit. Rev. Biochem. Mol. Biol.* 27: 157–190.

Luehrsen, K. R., and G. E. Fox. 1981. Secondary structure of eukaryotic cytoplasmic 5S ribosomal RNA. *Proc. Natl. Acad. Sci. USA.* 78: 2150–2154.

Minagawa, T., A. Murakami, Y. Ryo, and H. Yamagishi. 1983. Structural features of very fast sedimenting DNA formed by gene 49 defective T4. *Virology.* 126:183–93.

Mooren, M. M. W., C. W. Hilbers, G. A. van der Marel, J. H. van Boom, and S. S. Wijmenga. 1991. Three-dimensional homonuclear TOCSY-NOESY of nucleic acids. Possibilities for improved assignments. *J. Magn. Res.* 94:101–111.

Noller, H. F. 1984. Structure of ribosomal RNA. *Annu. Rev. Biochem.* 53: 119–162.

Noller, H. F. 1991. Ribosomal RNA and translation. *Annu. Rev. Biochem.* 60:191–227.

Nussbaum, J. M., M. E. A. Newport, M. Mackie, and N. B. Leontis. 1994. Structure-specific binding and photo-sensitized cleavage of branched, DNA three-way junction complexes by cationic porphyrins. *Photochem. Photobiol.* 59:515–528.

Piotto, M. E., and D. G. Gorenstein. 1991. ^{1}H 3D NOESY-TOCSY nuclear magnetic resonance spectrum of an oligonucleotide dodecamer duplex containing a GG mismatch. *J. Am. Chem. Soc.* 113:1438–40.

Puglisi, J. D., and I. Tinoco Jr. 1989. Absorbance melting curves of RNA. *Methods. Enzymol.* 180:304–325.

Radhakrishnan, I., and D. Patel. 1991. NMR assignment strategy for DNA protons through three-dimensional proton-proton connectivities. Application to an intramolecular DNA triplex. *J. Am. Chem. Soc.* 113: 8542–8544.

Radhakrishnan, I., D. J. Patel, and X. Gao. 1992. Three-dimensional homonuclear NOESY-TOCSY of an intramolecular pyrimidine-purine-pyrimidine DNA triplex containing a central G-TA triple: nonexchangeable proton assignments and structural implications. *Biochemistry.* 31: 2514–2523.

Rosen, M. A., and D. J. Patel. 1993a. Conformational differences between bulged pyrimidines (C-C) and purines (A-A, I-I) at the branch point of three-stranded DNA junctions. *Biochemistry.* 32:6563–6575.

Rosen, M. A., and D. J. Patel. 1993b. Structural features of a three-stranded DNA junction containing a C-C junctional bulge. *Biochemistry.* 32: 6576–6587.

Seeman, N. C., and N. R. Kallenbach. 1983. Design of immobile nucleic acid junctions. *Biophys. J.* 44:201–209.

Seeman, N. C., and N. R. Kallenbach. 1994. DNA branched junctions. *Annu. Rev. Biophys. Biomol. Struct.* 23:53–86.

Shen, Z., and P. J. Hagerman. 1994. Conformation of the central, three-helix junction of the 5S ribosomal RNA or *Sulfolobus acidocaldarius*. *J. Mol. Biol.* 241:415–430.

Sigal, N., and B. Alberts. 1972. Genetic recombination: the nature of a crossed strand-exchange between two homologous DNA molecules. *J. Mol. Biol.* 71:789–793.

Sklenar, V., M. Piotto, R. Leppik, and V. Saudek. 1993. Gradient-tailored water suppression for ^{1}H-^{15}N HSQC experiments optimized to retain full sensitivity. *J. Magn. Res.* 102:241–245.

Stahl, D. A., K. R. Luehrsen, C. R. Woese, and N. R. Pace. 1981. An unusual 5S rRNA, from *Sulfolobus acidocaldarius*, and its implications for a general 5S rRNA structure. *Nucleic Acids Res.* 9:6129–6137.

States, D. J., T. A. Haberkorn, and D. J. Rueben. 1982. A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Res.* 48:286–292.

Uhlenbeck, O. C. 1987. A small catalytic oligoribonucleotide. *Nature.* 328: 596–600.

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

Welch, J. B., D. R. Duckett, and D. M. J. Lilley. 1993. Structures of bulged three-way DNA junctions. *Nucleic Acids Res.* 21:4548–4555.

Westhof, E., P. Romby, P. J. Romaniuk, J. Ebel, C. Ehresmann, and B. Ehresmann. 1989. Computer modeling from solution data of spinach chloroplast and of *Xenopus laevis* somatic and Oocyte 5S rRNAs. *J. Mol. Biol.* 207:417–431.

Wilson, C., and J. W. Szostak. 1992. Ribozyme catalysis. *Curr. Opin. Struct. Biol.* 2:749–756.

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

Yang, J. H., N. Usman, P. Chartrand, and R. Cedergren. 1992. Minimum ribonucleotide requirement for catalysis by the RNA hammerhead domain. *Biochemistry.* 31:5005–5009.