Two-Dimensional ^{1}H NMR Studies of Immobile Holliday Junctions: Nonlabile Proton Assignments and Identification of Crossover Isomers†

Shiow Meei Chen and Walter J. Chazin*

Department of Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037

Received April 13, 1994; Revised Manuscript Received July 8, 1994®

ABSTRACT: The nonlabile protons of two 32-base-pair models of the Holliday junction intermediate in genetic recombination have been studied by two-dimensional ^{1}H nuclear magnetic resonance (NMR) spectroscopy. The sequence of these models is designed to fully inhibit branch migration of the junction and to probe the possible sequence dependence of these four-arm DNA structures. Overlap of resonances in homonuclear two-dimensional nuclear Overhauser enhancement (NOE) spectra necessitates the use of a multipathway approach for obtaining sequence-specific assignments, wherein all possible NOE connectivities are analyzed in parallel. Using this strategy, ^{1}H resonance assignments were obtained for virtually all nonlabile base protons and C1′, C2′, and C3′ sugar protons. Several unambiguous cross-arm NOE connectivities were identified, directly establishing the stacking arrangements of each contiguous (two-arm) helical domain. The distribution of the two possible stacking isomers is distinctly different for the two junctions studied, thereby indicating that the relative stability of the isomers is dependent on the sequence at the junction.

The Holliday junction (HJ)[1] is a four-arm branched DNA structure (Holliday, 1964) that is a common intermediate to most currently accepted models for genetic recombination. Although it is the cleavage of the HJ by enzymes that generates the parental or recombinant product, a significant body of evidence has accumulated indicating that the structure at the junction has a central role in determining the outcome of the recombination event. HJs or HJ-like structures may also play a role in other cellular processes such as replication (Mosig, 1983) and telomere resolution (Szostak & Blackburn, 1982). An understanding of the molecular basis of these cellular processes, in particular genetic recombination, will require a detailed knowledge of the HJ structure and of the nature of interactions with HJ-resolving enzymes.

In 1972, Sigal and Alberts published CPK and Kendrew models of the Holliday junction which showed that strand crossover can be readily achieved without unpairing any bases and that HJs can exist in two energetically favorable symmetrical conformations (Sigal & Alberts, 1972). Their models suggested that the junction regions are compact and that the crossover strands would be susceptible to cleavage, yielding two intact molecules with either parental or recombined configuration. In fact, cleavage of the crossover strand does proceed as predicted by their models (Kemper et al., 1984; deMassey et al., 1984; West & Korner, 1985; Evans & Kolodner, 1987). Experimental characterization of the physical properties of cellular HJs is made difficult by their inherent property of branch migration (Thompson et al., 1976), in which the junction point is translated along homologous stretches of duplex DNA. Seeman, Kallenbach, and coworkers proposed to circumvent this problem by studying synthetic models of Holliday junctions in which the sequence is designed so that the branch point is immobile (Seeman, 1982; Seeman & Kallenbach, 1983). Their original 32-base-pair immobile HJ system (J1) has been characterized extensively (Seeman & Kallenbach, 1983; Kallenbach et al., 1983; Seeman et al., 1985; Wemmer et al., 1985; Marky et al., 1987; Churchill et al., 1988; Chen et al., 1991, 1993). Working with immobile HJs, a consensus view consisting of two pairs of smoothly stacked, approximately coaxial duplex domains with the two contiguous strands running antiparallel to each other and the two crossover strands running in opposite directions in the two stacking domains (Figure 1) has been generated from a wide array of physical, chemical, and biological studies (Lilley & Clegg, 1993a,b; Seeman & Kallenbach, 1994). Molecular mechanics calculations suggest that this stacked-X structure has distorted B-DNA geometry at the junction with widening of the minor groove (von Kitzing et al., 1990). Complementary experiments indicate that the pairing preference of the four arms to form the stacking domains is dependent upon the DNA sequence at the branch point (Chen et al., 1988; Duckett et al., 1988; Murchie et al., 1989; Guo et al., 1991).

The next step required to interpret this wealth of biophysical information is to determine at the molecular level how the sequence at the junction influences HJ structure. Clearly, there is an urgent need for high-resolution structural data [e.g., Cooper and Hagerman (1991) and Seeman and Kallenbach (1994)]. To this end, we have set out to characterize the three-dimensional solution structure and dynamics of 32-base-pair immobile HJs using NMR spectroscopy. Sequence-specific NMR assignments for the labile protons of Seeman and Kallenbach's 32-base-pair synthetic immobile HJ (J1) and an analog that differs only in the relative position of two base pairs at the junction (J2) have been

† Supported by the National Science Foundation (DMB 9019250) and in part by a Junior Faculty Research Award from the American Cancer Society (JFRA-294).

® Abstract published in *Advance ACS Abstracts,* September 1, 1994.

[1] Abbreviations: HJ, Holliday junction; CD, circular dichroism; UV, ultraviolet; NMR, nuclear magnetic resonance; FID, free induction decay; 1D, one dimensional, 2D, two dimensional; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; JR-NOESY, NOESY spectrum acquired with the observe pulse replaced by a jump–return composite sequence; EDTA, ethylenediaminetetraacetic acid; PAGE, polyacrylamide gel electrophoresis; Tris, tris(hydroxymethyl)aminomethane; d_i(A;B), intranucleotide distance between protons A and B; d_s(A;B), sequential distance between protons A and B, where A is in the 5′ direction relative to B.

0006-2960/94/0433-11453$04.50/0

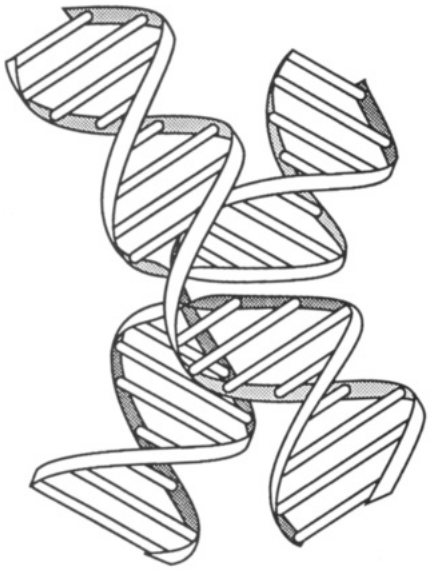

FIGURE 1: Schematic diagram of the stacked-X structure of the Holliday junction.

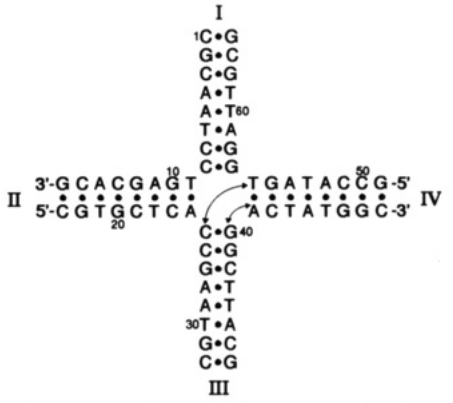

FIGURE 2: Sequence and numbering systems of J1 and J2. The sequence of J1 is shown, with the changes required to produce J2 indicated by the arrows: the G_{40}-C_{25} and T_{56}-A_{41} base pairs in J1 exchange places, becoming A_{40}-T_{25} and C_{56}-G_{41} base pairs in J2. The oligonucleotide sequence runs consecutively from the 5′ to the 3′ end of strand 1 (residues 1–16) and then on in the same manner to strands 2 (17–32), 3 (33–48), and 4 (49–64). The four arms of the junction are labeled with roman numerals.

reported previously (Chen et al., 1993). These studies provided conclusive evidence that the two stacking domains of synthetic immobile HJs are fully base-paired up to and at the site of strand crossover. A preliminary report describing the strategy to assign the nonlabile proton NMR resonances has also been published (Chen et al., 1991). We present here the sequence-specific ^{1}H NMR assignments for the nonlabile protons of J1 and J2 and further analysis of the data to define the duplex stacking arrangement and local geometry at the junction.

MATERIALS AND METHODS

The preparation of the immobile Holliday junctions J1 and J2 has been described in detail (Chen et al., 1991). Briefly, the four oligodeoxynucleotide hexadecamers were synthesized, deblocked, and purified by standard methods, and the strands were carefully titrated to a 1:1:1:1 molar ratio as monitored by NMR and gel electrophoresis. The mixture of strands was lyophilized and then redissolved in a buffer containing 20 mM Tris-HCl-d_{11} at pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 0.2 mM EDTA, and 0.1% (w/v) NaN_3. The D_2O sample was prepared by repeated lyophilization and dissolution in 99.996% D_2O (MSD Isotopes, Pointe Claire, Québec, Canada).

All NMR experiments were performed using a Bruker AM-600 spectrometer. 2Q spectra were acquired at 305 and 310 K using the standard pulse sequence (Braunschweiler et al., 1983). TOCSY spectra (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) were acquired at 310 K as described by Rance (1987) with 40–60 ms of DIPSI-2 spin-locking (Shaka et al., 1988). For 2Q and TOCSY, the transmitter was placed to the low-field side of the 1′H region, and base proton correlations were folded in the ω_1 dimension. NOESY spectra were acquired in H_2O with 200-ms mixing times at 293, 296 (J2 only), 300, 305, and 310 K and in D_2O with 100-ms mixing times at 300 and 310 K using the standard pulse sequence (Macura & Ernst, 1980). For spectra in H_2O, the last pulse was replaced by a composite jump–return sequence (Plateau & Guéron, 1982). For spectra in D_2O, a short Hahn-echo sequence was inserted just before acquisition to improve the quality of the baseline (Rance & Byrd, 1983; Davis, 1989). An additional NOESY spectrum with ω_1-decoupling (Rance et al., 1984) was acquired for each HJ in D_2O at 310 K. All NMR data were processed using a CONVEX C240 computer with FTNMR software (Hare Research, Inc., Woodinville, WA). Other relevant acquisition and processing parameters have been reported previously (Chen et al., 1991, 1993).

The generation of HJ molecular models has been described (Macke et al., 1992). Figure 3 was made within the Biopolymer module of InsightII (Biosym, Inc.) by creating the two 16′mer duplex domains and then forming the U-form crossover isomers that have 60° scissor angles using the previous models as templates for the interhelical angles.

RESULTS AND DISCUSSION

Figure 2 shows the 32-base-pair J1 Holliday junction along with the numbering systems used to designate each nucleotide and the four arms. J2 is a permutation of J1 with one base pair exchanged between arms III and IV, as indicated by the arrows. The four base pairs at the center of these four-arm DNA structures comprise the "junction", and the base pair at the open end of each duplex is termed the arm terminus.

Although the utilization of 2D NMR spectroscopy for sequential ^{1}H resonance assignments of the nonlabile protons of these models of the Holliday junction proved to be laborious and challenging, after careful examination of all possible spectral regions, virtually complete base proton, 1′H, 2′H, 2″H, and 3′H resonance assignments are obtained. The few missing resonance assignments were all from residues at the junction which tend to have broad line widths and very weak signals. The analysis of these 32-base-pair DNA structures thus represents a substantial increase over the largest oligonucleotide for which complete ^{1}H NMR assignments have been obtained, a non-self-complementary 23-base-pair duplex (Otting et al., 1987; Grütter et al., 1988). A shorthand notation similar to that in Wüthrich (1986) will be utilized to specify NOEs and the corresponding interproton distances: intraresidue and sequential (5′ to 3′ direction) NOEs are indicated by d_i(A;B) and d_s(A;B), respectively.

Assignment Protocol. The strategy adopted for sequential assignment follows the standard two-step process developed for smaller DNA systems [e.g., Wüthrich (1986)]. The initial stage involved identifying scalar (through-bond) correlations in 2Q and TOCSY spectra for the protons within each sugar ring and 5H/6H of the cytosine and $5CH_3$/6H of the thymine bases. The corresponding base and sugar protons for each residue were then paired, and the sequentially adjacent residue was identified, via dipolar (through-space) correlations in NOESY spectra. Problems arise for the HJ system due to severe cross-peak overlap in the standard regions of the

NOESY spectrum that can usually be analyzed to make complete sequential assignments. The strategy adopted to overcome these problems involved analysis of *all regions* of the spectrum in parallel, circumventing problems in one region by finding connectivities in another, and making use of sequential pathways not ordinarily followed to make assignments (e.g., $d_s(2',2'';5)$; Chen et al., 1991). To achieve complete assignments of J1 and J2, it was also necessary to acquire a series of spectra in the temperature range from 293 to 310 K and allow a limited degree of NOE spin diffusion in the spectra and then search in parallel for *all possible assignment pathways in all spectra.*

In general, the degree of difficulty of assigning residues followed the order guanosine > adenosine > cytidine > thymidine. Thymidine is distinguished by the unique methyl group resonances, cytidine by the 5H resonances, and adenosine by reasonable dispersion of its 8H resonances. The guanosine 8H resonances tended to be much less dispersed; most were assigned via connectivities to an adjacent residue that was already assigned. As a consequence of these different degrees of difficulty in assigning residue types, the sequential assignment was made by first branching out from the various thymidines and then systematically selecting cytidines for further analysis. This proved sufficient to build up a substantial set of partial assignments for each strand that were then connected together. Such an approach was possible because the base-pair patterns are different for each arm and the sequence along the strands is already fairly unique at the trinucleotide level.

In the initial stages, the J2 spectra were analyzed, and assignments were made independently of J1. After noting the extreme similarity in the chemical shifts for the preponderance of residues, it was deemed advantageous to proceed further by analyzing J1 and J2 in parallel. It is important to stress that every effort was made to assign *all* cross peaks observed in each NOESY spectrum, thereby overcoming the severe complexity of the assignment problem and increasing the reliability of the final result.

Sequence-Specific Assignments of J1. The first step involved a thorough analysis of the 2Q spectrum, which yielded 18 of the 19 cytosine 5H/6H correlations, 11 of the 13 thymine $5CH_3$/6H correlations, and 62 1'H/2'H/2''H sugar ring spin subsystems. The broadened 5H/6H peaks for C_{25} were identified in the TOCSY spectrum, whereas the remaining $5CH_3$/6H and 1'H/2'H/2''H resonances were identified during analysis of the NOESY spectra. 3'H assignments were made from a combined analysis of the appropriate regions of the 2Q, TOCSY, and NOESY spectra. The resonances from residues at the strand termini were identified directly because they have unique spectral characteristics that arise from the combination of end-fraying (narrow line widths) and the absence of phosphate linkages at the 3' or 5' position (distinctive 2'H, 2''H, and 3'H chemical shifts for G_{16}, C_{32}, C_{48}, and C_{64}; distinctive 5'H/5''H chemical shifts for C_1, C_{17}, G_{33}, and G_{49}).

Once all scalar correlations were identified, the NOESY spectra were analyzed. Since J1 is comprised of four oligonucleotide strands with 16 residues on each, 60 sequential connectivities are ultimately sought. However, it is possible to make complete sequence-specific assignments with less than this number, as long as there are no cases where sequential connectivities to two consecutive residues are missing. In fact, 56 of the 60 connectivities for J1 could be identified on the basis of at least one sequential NOE, and most were made by many more. A summary of the sequential assignment of J1, along with a detailed description for strand 1, is provided as supplementary material.

Upon completion of the exhaustive analysis of NOEs, there was no unambiguous experimental evidence for NOE connectivity between four pairs of residues (C_{23}-A_{24}, A_{24}-C_{25}, T_{56}-G_{57}, G_{57}-G_{58}); hence, only A_{24} and G_{57} have no direct sequential NOE-based assignment in either the 5' or 3' direction.[2] However, assignments for these two residues could be made from the pool of two remaining sets of sugar and base proton resonances, on the basis of A_{24}-8H having a lower field chemical shift. The complete sequence-specific resonance assignments for J1 are provided in Table 1.

Sequence-Specific Assignments of J2. A thorough analysis of the 2Q spectrum of J2 yielded all 19 cytosine 5H/6H correlations, 11 of the 13 thymine $5CH_3$/6H correlations, and all 64 1'H/2'H/2''H sugar ring spin subsystems. The remaining $5CH_3$/6H correlations were subsequently identified in the NOESY spectra, and the 3'H assignments were made from a combined analysis of the appropriate regions of the 2Q, TOCSY, and NOESY spectra. From the total of 60 potential sequential connectivities for J2, 57 could be identified on the basis of at least one sequential NOE, and most were made by many more. A summary of the sequential assignment of J2 is provided as supplementary material.

Upon completion of the exhaustive analysis of NOEs, there remained no experimental evidence for NOE connectivity between three pairs of residues (T_{25}-C_{26}, C_{56}-G_{57}, G_{57}-G_{58}), all around the junction area. Since only G_{57} lacked sequential NOEs in either the 5' and 3' directions, its resonances could be assigned by default, and the sequential assignment was complete. The chemical shifts of J2 resonances that differ significantly ($>\pm 0.03$ ppm) from the corresponding values for J1 are listed in parentheses in Table 1.

The Arm Stacking Geometry of J1 Is I/II. The NMR determination of the arm stacking geometries of the J1 and J2 model HJs is dependent on the observation of the series of NOE connectivities involving the eight residues at the junction. All other NOEs follow the expected patterns for standard B-form DNA duplexes and are essentially independent of the arm stacking geometry. Thus, the four arms are comprised of normal B-DNA duplexes right up to the junction base pairs. These conclusions are also reflected in the extreme similarity of the chemical shifts of J1 and J2 for the first seven base pairs within each arm (Table 1).

At the junction, there are two types of strands, contiguous and crossover, that are expected to exhibit very different patterns of NOE cross peaks. A standard set of sequential NOE connectivities is expected for the residues on the contiguous strands. On the crossover strand, NOEs between residues that are not adjacent in sequence but are spatially proximate in the manner of sequential NOEs (i.e., between the residues at the junction on the crossover strands that are stacked upon each other within the duplex domains) report directly on the stacking geometry. In an ideal situation, all of the critical sequential-type NOEs between residues that are not adjacent in the sequence would be observed, and the analysis of stacking geometry would be straightforward. However, the detection of these NOEs is greatly complicated by (1) the large line widths of most of the resonances of the residues at the junction, (2) the general problem of resonance degeneracy in this 64-residue structure, and (3) the strong likelihood of conformational distortion away from standard B-form geometry of the residues at the point of strand crossover.

[2] Note that as a result of the crossing over of strands from one duplex stacking domain to the other, the sequential NOE connectivities at the two points of crossover will be absent [I/II stacking, no $d_s(24;25)$ or $d_s(56;57)$; I/IV stacking, no $d_s(8;9)$ or $d_s(40;41)$].

Table 1: ^{1}H NMR Resonance Assignments of Nonlabile Protons of J1 and J2 at pH 7.5, 310 K[a]

residue	base	M/5/2	1′	2′	2″	3′	others
strand 1							
C1	7.62	5.91	5.79	1.92	2.39	4.69	4.09 (4′), 3.73 (5′)
G_2	7.94	–[b]	5.88	2.65	2.72	5.01	
C_3	7.34	5.44	5.49	1.90	2.27	4.82	4.08 (4′)
A_4	8.24	7.34	5.95	2.79	2.93	5.08	4.33 (4′)
A_5	8.16	7.72	6.19	2.59	2.91	5.00	4.43 (4′), 4.29 (5′), 4.17 (5″)
T_6	7.11	1.41	5.84	2.08	2.52	4.82	
C_7	7.56	5.58	6.06	2.24	2.53	4.80	4.28 (4′)
C_8	7.45	5.42 (5.47)	5.94	2.09 (2.20)	2.53	4.82	
T_9	7.14	1.47	5.68	1.63	2.04 (2.11)	4.81	
G_{10}	7.89	–	5.33	2.67	2.76	4.98	4.29 (4′)
A_{11}	8.09	7.64	6.06	2.66	2.88	5.04	
G_{12}	7.59	–	5.69	2.45	2.61	4.94	
C_{13}	7.29	5.27	5.62	1.92	2.35	4.81	
A_{14}	8.26	7.83	6.19	2.68	2.85	5.01	4.40 (4′), 4.15 (5′), 4.03 (5″)
C_{15}	7.27	5.37	5.72	1.86	2.29	4.79	
G_{16}	7.89	–	6.15	2.38	2.59	4.66	
strand 2							
C_{17}	7.66	5.92	5.82	2.02	2.44	4.72	4.08 (4′), 3.75 (5′)
G_{18}	7.99	–	6.02	2.69	2.82	4.99	
T_{19}	7.21	1.49	5.89	2.10	2.50	4.91	
G_{20}	7.89	–	5.91	2.66	2.70	4.99	
C_{21}	7.40	5.32	5.89	2.05	2.49	4.67 (4.72)	
T_{22}	7.41	1.57	6.02	2.08	2.43 (2.50)	4.85	
C_{23}	7.43	5.68	5.48	1.83 (1.91)	2.07 (2.17)	4.81	
A_{24}	8.33 (8.27)	7.69[c] (7.62)[c]	6.27	2.74	3.03 (2.90)		
C_{25}	7.64	5.63	5.80	1.93	2.45		
(C_{56})	(7.44)	(5.20)	(5.89)		(2.40)		
C_{26}	7.47	5.60 (5.70)	5.39	1.97 (2.03)	2.28	4.78	4.08 (4′)
G_{27}	7.84	–	5.33	2.65	2.73	4.97	
A_{28}	8.17	7.32	5.98	2.78	2.93	5.07	
A_{29}	8.15	7.67	6.16	2.56	2.88	5.01	
T_{30}	7.01	1.38	5.74	1.87	2.31	4.84	
G_{31}	7.85	–	5.94	2.60	2.69	4.97	
C_{32}	7.47	5.46	6.21	2.16	2.25	4.49	
strand 3							
G_{33}	7.94	–	5.98	2.59	2.78	4.85	
C_{34}	7.49	5.47	5.67	2.15	2.47	4.90	4.24 (4′), 4.16 (5′), 3.72 (5″)
A_{35}	8.37	7.73	6.34	2.78	2.99	5.07	4.47 (4′), 4.22 (5′), 4.13 (5″)
T_{36}	7.22	1.45	5.97	2.01	2.54	4.84	
T_{37}	7.38	1.60	6.05	2.12	2.46	4.87	
C_{38}	7.43	5.65	5.49	2.00	2.30	4.87	4.16 (4′)
G_{39}	7.81	–	5.63 (5.49)	2.62	2.75	4.97 (4.85)	
G_{40}	7.64	–	5.69	2.54	2.66		
(G_{41})			(5.80)				
A_{41}	8.08	7.64	6.19	2.64	2.85	5.00	
(A_{40})	(8.16)	(7.73)	(6.10)	(2.75)	(2.80)		
C_{42}	7.25 (7.39)	5.22	5.77 (5.82)	1.98 (2.07)	2.45	4.69	
T_{43}	7.36 (7.42)	1.63	5.75	2.17	2.53	4.86	
A_{44}	8.24 (8.29)	7.29 (7.36)	6.21	2.61	2.88	5.00	
T_{45}	7.09	1.37	5.69	1.92	2.33	4.86	
G_{46}	7.79	–	5.66	2.64	2.73	4.98	
G_{47}	7.73	–	5.97	2.52	2.71	4.96	4.37 (4′), 4.20 (5′), 4.06 (5″)
C_{48}	7.44	5.40	6.19	2.17	2.20	4.50	
strand 4							
G_{49}	7.96	–	6.00	2.64	2.77	4.85	
C_{50}	7.52	5.41	6.09	2.18	2.52	4.87	4.22 (4′)
C_{51}	7.53	5.66	5.57	2.15	2.45	4.87	4.15 (4′)
A_{52}	8.33	7.64	6.27	2.71	2.95	5.03	
T_{53}	7.14	1.50	5.66 (5.60)	2.01	2.41 (2.34)	4.86	
A_{54}	8.15	6.99[c] (7.19)[c]	6.03	2.70	2.84	5.03	
G_{55}	7.60 (7.66)	–	5.87 (5.79)	2.45	2.69 (2.64)	4.63 (4.98)	
T_{56}	7.33	1.23	5.85	1.89	2.35		
(T_{25})	(7.40)	(1.49)	(6.02)	(1.97)	(2.42)		
G_{57}	7.70 (7.62)	–	5.57	2.40	2.56		
G_{58}	7.40 (7.53)	–	5.77	2.47 (2.52)	2.71	4.85	
A_{59}	8.05	7.84	6.26	2.61	2.94	4.96	
T_{60}	7.24	1.23	6.07	2.06	2.58	4.88	
T_{61}	7.32	1.63	5.86	2.09	2.44	4.92	
G_{62}	7.94	–	5.90	2.40	2.70		
C_{63}	7.36	5.47	5.84	1.90	2.34		
G_{64}	7.95	–	6.17	2.62	2.38	4.68	

[a] Chemical shifts referenced to the residual HOD signal, precalibrated using methanol. The majority of the chemical shifts of J1 and J2 are identical within experimental error (±0.02 ppm); hence, a separate listing for J2 is included in parentheses only when a chemical shift is >0.03 ppm different. [b] A dash indicates no proton present in the C2 or C5 position. [c] Chemical shift identified only from connectivities to labile protons; taken from Chen et al. (1993) wherein values are reported at 300 K.

The unambiguously identified sequential NOEs for the eight residues at the junction of J1 are listed in Table 2, including those reported previously involving exchangeable protons (Chen et al., 1993). At the junction of arm I and arm II in

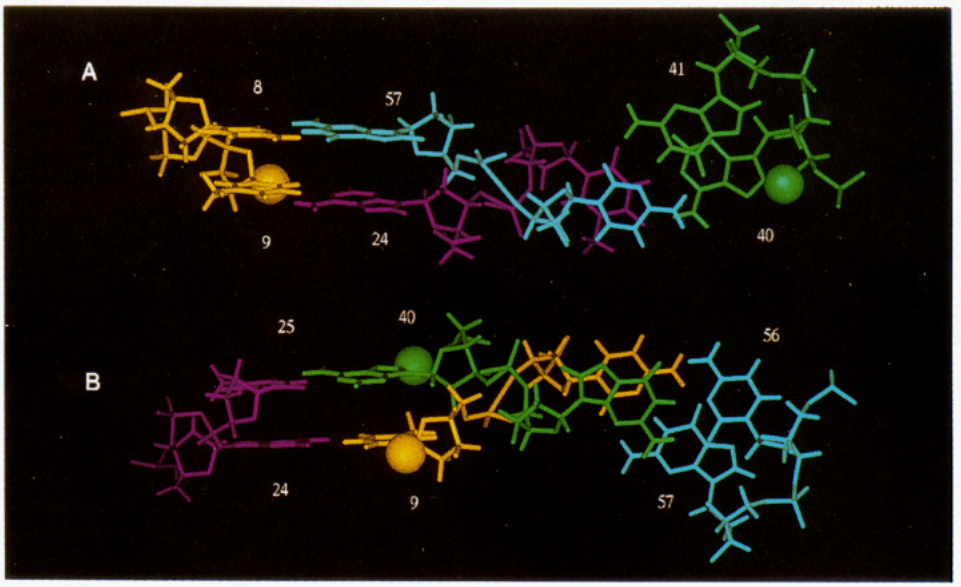

FIGURE 3: Molecular graphics tube diagram of molecular models of the I,II/III,IV (A) and I,IV/II,III (B) isomers of J2, indicating the relative distances between T_9 and A_{40} in the two conformations. The models are constructed from two duplexes with regular B-form geometries that have been nicked and religated to form the crossover structure in the stacked-X conformation with a 60° scissor angle. The four strands are color coded as follows: 1, yellow; 2, magenta; 3, green; 4, blue. In both models spheres are drawn to indicate the location of T_9-6H and A_{40}-1′H.

strand 1, virtually all of the $d_s(C_8;T_9)$ NOEs possible between nonexchangable protons were present and at least partially resolved, as well as several additional NOEs involving exchangable protons. These data indicate that strand 1 is one of the contiguous strands. The identification of NOEs for the other strands proved to be more complicated. In strand 2, although C_{23} was readily assigned, the A_{24}-8H resonance was very broad, and only one tentatively assigned NOE could be identified from a C_{23} resonance. No $d_s(A_{24};C_{25})$ connectivities were observed. Although the C_{25} resonances were broad, connectivities to the 5H and 6H resonances of C_{26} could be identified. In strand 3, there was difficulty with G_{39} due to a severe case of multiple resonance degeneracy, but $d_s(C_{38};G_{39})$ and $d_s(G_{39};G_{40})$ connectivities could eventually be identified. Several NOEs were identified between G_{40} and A_{41} resonances, although these tended to be weak in contour plots as the signals of both residues exhibited exchange broadening. Numerous sequential NOEs provided connectivity on to C_{42}. In strand 4, three of the four residues at or adjacent to the junction are intrinsically difficult guanosine residues; hence, working from T_{56} proved to be critical to assigning this region. The $5CH_3$ resonance of T_{56} is sufficiently resolved from T_{60}-$5CH_3$ at temperatures below 300 K to allow NOEs to T_{56}-6H and G_{55}-8H to be identified. The resonances of both T_{56} and G_{57} are exchange broadened as for the G_{40}-A_{41} pair, but in this case no $d_s(T_{56};G_{57})$ NOE connectivities could be identified. G_{58}-8H is uniquely high field shifted among the guanosine residue but is severely overlapped with C_{21}-6H, which in turn has several resonance degeneracies with T_{56}. Some G_{58}-A_{59} connectivities could however be distinguished, providing unambiguous sequential assignment for G_{58}, but no NOEs to G_{57} were identified. The cumulative analysis of the results for strands 2, 3, and 4 strongly suggested that strands 2 and 4 were crossover and that strand 3 was the second contiguous strand.

In addition to the sequential NOEs, five nonsequential NOEs were unambiguously assigned between residues in the junction region involving T_9-G_{57}, C_{25}-T_{56}, and G_{40}-T_{56} (Table 2). These NOEs specifically identify close spatial proximity between base pair C_8-G_{57} in arm I and T_9-A_{24} in arm II and between base pair C_{25}-G_{40} in arm III and A_{41}-T_{56} in arm IV. Furthermore, these contacts are fully consistent with the observation of standard sequential NOEs observed between the junction residues of strands 1 and 3 and with the absence of NOEs between A_{24} and C_{25} in strand 2 and between C_{56} and G_{57} in strand 4, precisely at the points of strand crossover. Thus, the stacking domains of J1 are comprised of arm I with arm II and arm III with arm IV (the I/II, III/IV isomer). Our results are fully consistent with, and provide conclusive evidence to confirm, the conclusions drawn by Seeman, Kallenbach, and co-workers about J1 stacking geometry from a range of more indirect methods [reviewed in Seeman and Kallenbach (1994)].

J2 Exists as a Mixture of I/II and I/IV Arm Stacking Geometries. The determination of stacking geometry was more difficult for J2 than for J1 because the detection of NOEs for the residues at and near the junction was more difficult. A list of all of the unambiguously assigned NOEs between the junction residues of J2 is provided in Table 2. There are a number of typical $d_s(C_8;T_9)$ and $d_s(A_{40};G_{41})$ connectivities that identify strands 1 and 3 as contiguous and clearly resolved NOEs indicating a close proximity between T_{25} and C_{56} that identify strands 2 and 4 as crossover. The stacking geometry corresponding to these data is arm I with arm II and arm III with arm IV, as determined for J1. The absence of the complementary NOEs between A_{24} and G_{57} can be attributed to the substantial exchange broadening of the resonances from these residues.

Close scrutiny of Table 2 reveals, however, that the ensemble of all NOEs cannot be explained solely by the I/II, III/IV isomer; a set of unambiguous but uniformly weaker NOEs [$d(T_9;T_{25})$, $d(T_9;A_{40})$, $d(A_{24};T_{25})$] were identified that are clearly inconsistent with this stacking geometry. This apparent anomaly can be explained by reference to molecular models created for the two J2 crossover isomers (Macke et al., 1992), as shown in Figure 3. For example, in the I/II, III/IV isomer it is impossible to explain the observation of an NOE between T_9 and A_{40} (Figure 4) because these residues are located on the opposite contiguous strands at the outer extrema of the opposing stacking domains. In the molecular model built with stranded B-form geometry shown in Figure 3, these residues are rather distant with no proton pairs closer than 15 Å from each other. Even with substantial distortion from standard B-form geometry, it is not possible to bring T_9 and A_{40} close

Table 2: Unambiguously Assigned NOE Cross Peaks between Junction Residues of J1 and J2

J1	J2
	major
C_8-1′H; T_9-6H	C_8-2′H; T_9-5M
C_8-1′H; T_9-5M	C_8-2″H; T_9-5M
C_8-2′H; T_9-5M	C_8-1′H; T_9-6H
C_8-2″H; T_9-5M	C_8-8H; T_9-6H
C_8-2″H; T_9-6H	C_8-5H; T_9-5M
C_8-5H; T_9-5M	C_8-1′H; T_9-5M
C_8-6H; T_9-6H	C_8-4NH1; T_9-5M
C_8-6H; T_9-5M	T_{25}-5M; C_{56}-5H
C_8-4NH1; T_9-5M	T_{25}-5M; C_{56}-6H
T_9-5M; G_{57}-N1H	A_{40}-1′H; G_{41}-8H
C_{25}-5H; T_{56}-5M	A_{40}-8H; G_{41}-8H
C_{25}-6H; T_{56}-5M	minor
G_{40}-1′H; A_{41}-8H	T_9-6H; A_{40}-1′H
G_{40}-8H; A_{41}-8H	T_9-5M; A_{40}-8H
G_{40}-N1H; A_{41}-2H	T_9-5M; A_{40}-1′H
G_{40}-N1H; T_{56}-N3H	T_9-N3H; T_{25}-N3H
G_{40}-N1H; T_{56}-5M	A_{24}-2H; T_{25}-5M
	A_{24}-8H; T_{25}-6H

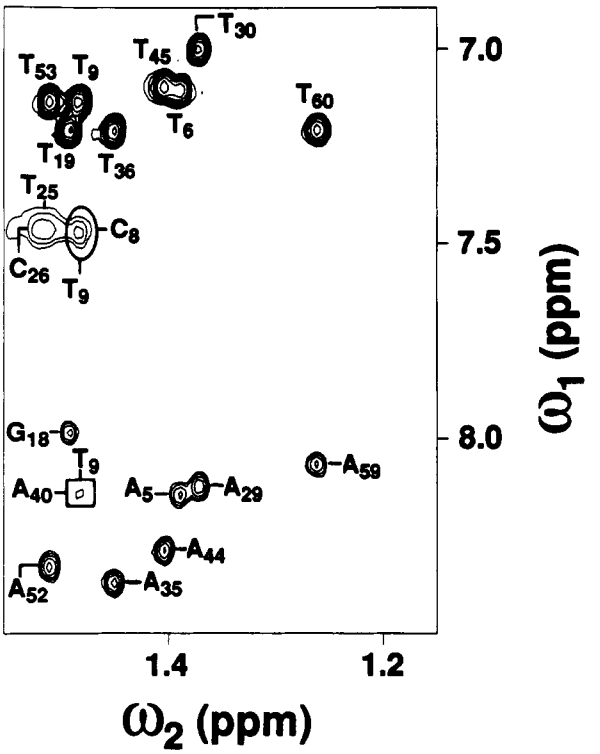

FIGURE 4: Region of the 600-MHz NOESY spectrum of 1 mM J2 containing cross peaks in the base proton to methyl proton region. The spectrum was acquired at 310 K with a mixing time of 100 ms from a D_2O solution containing 20 mM Tris-d_{11}, 50 mM NaCl, 5 mM $MgCl_2$, 0.2 mM EDTA, and 0.1% (w/v) NaN_3, at pH 7.5. Intraresidue NOEs are identified by labeling above or below the cross peak, whereas sequential NOEs are identified by labeling at the left or right of the cross peak. This plot is made with a relatively high minimum contour level for clarity of presentation and to clearly reveal relative intensities; hence, certain peaks appear to be very weak. All of the assigned NOE cross peaks have been unambiguously identified, working in plots made with much greater expansion and with lower minimum contour levels. The critical $d_s(8;9)$ and $d_s(40;9)$ NOEs are highlighted with a circle and a square, respectively.

enough to explain the observed NOEs. However, in the J2 molecular model with the alternate stacking arrangement (I/IV, II/III), T_9 and A_{40} are stacked directly adjacent to each other (Figure 3), well within the range to give rise to NOEs between them. The only explanation that is consistent with all of the available experimental evidence is that, under the conditions of the NMR experiments J2 exists in solution as an equilibrium distribution of the two possible stacking geometries in fast exchange on the NMR chemical shift time scale, with a preference for the I/II, III/IV versus the I/IV, II/III isomer of the order of 5:1.

CONCLUDING REMARKS

The experiments described in this report demonstrate the validity of using ^{1}H NMR to obtain important insights into the structure of Holliday junctions. The spatial proximities of stacked bases along each strand and within each duplex domain can be analyzed in a site-specific manner to determine the local geometry at the junction. The unambiguous evidence for full base pairing up to and including the residues at the junction has been reported previously (Chen et al., 1993). The ^{1}H chemical shifts and patterns of NOEs reported here provide direct and conclusive evidence for base stacking up to and including the junction core. Furthermore, the relative magnitudes of sequential NOEs indicate that all four arms have normal B-DNA conformations all the way up to the junction base pairs.

The arm stacking geometry of the J1 model HJ has been determined to be arm I with arm II and arm III with arm IV, confirming conclusions drawn from a range of gel electrophoresis, chemical, and enzymatic protection experiments (Seeman & Kallenbach, 1994). However, although all of the available data are fully consistent with I/II, III/IV geometry, many of the resonance lines from protons of the four base pairs at the junction are broader than the others, with the effect conspicuously stronger on the crossover strands. These data are indicative of dynamic and structural perturbation relative to standard duplex DNA, which we attribute to the conformational strain associated with strand crossover. The extreme similarity of the chemical shifts for all seven base pairs preceding the junction in each of the four duplex arms strongly suggests that these perturbations are highly localized.

The corresponding analysis of the stacking geometry for J2 revealed an equilibrium distribution of the two possible stacking isomers in fast exchange on the NMR chemical shift time scale, with a decided preference for I/II, III/IV. Results from time-resolved fluorescence resonance energy transfer experiments on J2 appear to be consistent with this analysis (M. Yang and D. P. Millar, in preparation). A distribution of the two crossover isomers has also been invoked to explain the observation of minor sites in restriction enzyme cleavage experiments [e.g., Duckett et al. (1988) and Mueller et al. (1988)]. The large stabilization energies associated with base stacking presumably drive the distribution between the two stacking isomers in the Holliday junction. But a critical question remains: what is the molecular mechanism by which this isomerization occurs? Current experimental evidence offers very little insight.

The J1 sequence appears to have a particularly strong preference for the I/II, III/IV stacking geometry; J2 has a decidedly weaker preference under identical conditions. This substantial difference observed in the relative ratios of isomers for J1 and J2, which differ only by the exchange of two base pairs at the junction, provides additional evidence in support of the concept that the relative stability of the two stacking isomers is determined solely by the sequence at the junction [reviewed in Lilley and Clegg (1993a,b) and Seeman and Kallenbach (1994)]. The examination of additional analogs of J1 is currently in progress to further test the validity of this hypothesis and determine the molecular basis for relative junction stabilities. This information will be used to establish a set of guidelines for predicting HJ crossover isomer distribution and to search for a correlation between HJ

sequence and the propensity to generate parental or recombinant products.

ACKNOWLEDGMENT

We thank Dr. Siobhan Miick for the preparation of Figure 3, Dr. Ned Seeman for several helpful discussions, Dr. Mark Rance for continued support with experimental methodology, and Beth Larson for help with typing the manuscript.

SUPPLEMENTARY MATERIAL AVAILABLE

A summary of the sequential resonance assignment procedure for J1 and J2 and two figures with four panels each containing the $d_{i,s}(1';6,8)$ connectivities of J1 and $d_{i,s}(2',2'';6,8)$ connectivities of J2 from 600-MHz NOESY spectra, with the cross peaks labeled with sequence-specific assignments for each strand in separate panels (11 pages). Ordering information is given on any current masthead page.

REFERENCES

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.
Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) *Mol. Phys. 48*, 535–560.
Chen, J.-H., Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R., & Seeman, N. C. (1988) *Biochemistry 27*, 6032–6038.
Chen, S.-m., Heffron, F., Leupin, W., & Chazin, W. J. (1991) *Biochemistry 30*, 766–771.
Chen, S.-m., Heffron, F., & Chazin, W. J. (1993) *Biochemistry 32*, 319–326.
Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R., & Seeman, N. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4653–4656.
Cooper, J. P., & Hagerman, P. J. (1991) *Curr. Opin. Struct. Biol. 1*, 464–468.
Davis, D. G. (1989) *J. Magn. Reson. 81*, 603–607.
deMassey, B., Studire, F. W., Dorgai, L., Appelbaum, E., & Weisberg, R. A. (1984) *Cold Spring Harbor Symp. Quant. Biol. 49*, 715–726.
Duckett, D. R., Murchie, A. I. H., Diekmann, S., Kitzing, E., Kemper, B., & Lilley, D. M. J. (1988) *Cell 55*, 79–89.
Evans, D., & Kolodner, R. (1987) *J. Biol. Chem. 262*, 9160–9165.
Grütter, R., Otting, G., Wüthrich, K., & Leupin, W. (1988) *Eur. Biophys. J. 16*, 279–286.
Guo, Q., Lu, M., & Kallenbach, N. R. (1991) *Biopolymers 31*, 359–372.
Holliday, R. (1964) *Genet. Res. 5*, 282–304.
Kallenbach, N. R., Ma, R. I., & Seeman, N. C. (1983) *Nature 305*, 829–831.
Kemper, B., Jensch, F., Depka-Prodzynski, M. V., Fritz, H.-J., Borgmeyer, U., & Mizuuchi, K. (1984) *Cold Spring Harbor Symp. Quant. Biol. 49*, 815–825.
Lilley, D. M. J., & Clegg, R. M. (1993a) *Annu. Rev. Biomol. Struct. 22*, 299–328.
Lilley, D. M. J., & Clegg, R. M. (1993b) *Q. Rev. Biophys. 26*, 131–175.
Macke, T., Chen, S.-m., & Chazin, W. J. (1992) in *Structure & Function, Volume 1: Nucleic Acids* (Sarma, R. H., & Sarma, M. H., Eds.) pp 213–227, Adenine Press, Schenectady, NY.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.
Marky, L. A., Kallenbach, N. R., McDonough, K. A., & Seeman, N. C. (1987) *Biopolymers 26*, 1621–1634.
Mosig, G. (1983) in *Relationship of T4 DNA replication and recombination*, pp 120–130, American Society for Microbiology, Washington, DC.
Murchie, A. I. H., Clegg, R. M., Kitzing, E. V., Duckett, D. R., Diekmann, S., & Lilley, D. M. J. (1989) *Nature 341*, 763–766.
Otting, G., Grütter, R., Leupin, W., Minganti, C., Ganesh, K. N., Sproat, B. S., Gait, M. J., & Wüthrich, K. (1987) *Eur. J. Biochem. 166*, 215–220.
Plateau, P., & Guéron, M. J. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.
Rance, M. (1987) *J. Magn. Reson. 74*, 557–564.
Rance, M., & Byrd, R. A. (1983) *J. Magn. Reson. 54*, 221–240.
Rance, M., Wagner, G., Sørensen, O. W., Wüthrich, K., & Ernst, R. R. (1984) *J. Magn. Reson. 59*, 250–261.
Seeman, N. C. (1982) *J. Theor. Biol. 99*, 237–247.
Seeman, N. C., & Kallenbach, N. R. (1983) *Biophys. J. 44*, 201–209.
Seeman, N. C., & Kallenbach, N. R. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 53–86.
Seeman, N. C., Maestre, M. F., Ma, R.-I., & Kallenbach, N. R. (1985) in *The Molecular Basis of Cancer* (Rein, R., Ed.) pp 99–108, Liss, New York.
Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274–293.
Sigal, N., & Alberts, B. (1972) *J. Mol. Biol. 71*, 789–793.
Szostak, J. W., & Blackburn, E. H. (1982) *Cell 29*, 245–255.
Thompson, B. J., Camien, M. N., & Warner, R. C. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2299–2303.
von Kitzing, E., Lilley, D. M. J., & Diekmann, S. (1990) *Nucleic Acids Res. 18*, 2671–2683.
Wemmer, D. E., Wand, A. J., Seeman, N. C., & Kallenbach, N. R. (1985) *Biochemistry 24*, 5745–5749.
West, S. C., & Korner, A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6445–6449.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.