Two-Dimensional ^{1}H NMR Studies of 32-Base-Pair Synthetic Immobile Holliday Junctions: Complete Assignments of the Labile Protons and Identification of the Base-Pairing Scheme†

Shiow-meei Chen, Fred Heffron, and Walter J. Chazin*

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

Received August 25, 1992; Revised Manuscript Received October 26, 1992

ABSTRACT: The labile protons of two 32-base-pair, four-arm models of immobile Holliday junctions have been studied by two-dimensional ^{1}H nuclear magnetic resonance (NMR) spectroscopy. Overlap of resonances in the imino–imino region of two-dimensional nuclear Overhauser enhancement (NOE) spectra necessitates the use of a multi-pathway approach for obtaining sequence-specific assignments wherein all possible NOE connectivities to the labile protons are utilized, including those from the 2H of adenine, 5CH$_3$ of thymine, and 5H of cytosine. Resonance assignments are obtained for all slowly exchanging imino and cytosine amino protons. Base-pairing up to and including the junction point is found in all four arms of both Holliday junctions. Several cross-arm NOE connectivities are identified and can be used to infer the geometry of the helical stacking domains. The two Holliday junctions studied, which differ only by the exchange of two base pairs at the branch point, appear to have opposite arm stacking geometries. These assignments form an important part of the critical background for detailed NMR analysis of Holliday junction structure and dynamics.

Genetic recombination is one of the fundamental processes of biology, leading to genetic variation and diversity. Most mechanisms proposed for recombination proceed via a four-arm branched DNA intermediate termed the Holliday junction (Holliday, 1964), which is cleaved by enzymes (resolvases) to generate either parental or recombinant products. A significant body of evidence has accumulated indicating that the structure of the Holliday junction (HJ)[1] has a central role in determining the outcome of the recombination event. Thus, an understanding of the molecular basis of genetic recombination requires a detailed knowledge of the structure of the HJ and the nature of the interactions with resolvase enzymes.

Characterization of the physical properties of cellular HJs is made difficult by the inherent property of branch migration (Thompson et al., 1976), in which the junction point is translated along homologous stretches of duplex DNA. To circumvent this problem, Seeman, Kallenbach, and co-workers proposed the study of synthetic analogs of HJs, in which the sequence is designed so that the branch point is immobile (Seeman, 1982; Seeman & Kallenbach, 1983). These synthetic immobile HJs have been studied by an array of techniques, leading to a consensus view of the HJ as two pairs of smoothly stacked coaxial duplex domains with the two contiguous strands running antiparallel to each other and the two cross-over strands running in opposite directions in the two stacking domains (Sigal & Alberts, 1972; Gough & Lilley, 1985; Cooper & Hagerman, 1987, 1989; Duckett et al., 1988; Churchill et al., 1988; Lu et al., 1989; Seeman et al., 1989; Murchie et al., 1989, 1991; Clegg et al., 1992). 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). There is strong evidence that the arrangement of the stacking domains is dependent upon the DNA sequence at the branch point [e.g., Chen et al. (1988), Duckett et al. (1988), Murchie et al. (1989), and Guo et al. (1991)].

The experimental approaches utilized to date to study immobile HJs all involve indirect methods to *infer* the structure of the HJ; hence, it is desirable to obtain a detailed structural characterization of HJs. Since it has not been possible to study HJs by X-ray crystallography, we have set out to characterize the three-dimensional structure and dynamics of 32-base-pair immobile Holliday junctions by NMR spectroscopy. These studies will provide critical information to verify and refine the structural aspects implicit in the models presently available for the HJ. A preliminary report on our efforts described the strategy to assign the nonlabile proton NMR resonances (Chen et al., 1991). We present here the sequence-specific NMR assignments for the labile imino and cytosine amino protons of two 32-base-pair synthetic immobile HJs that differ only in the relative position of two base pairs at the junction. One of these corresponds to a model HJ studied previously by 1D NMR techniques (Wemmer et al., 1985), and the results of these two studies are compared. In addition to the complete assignments, our data provide sufficient information to define the base-pairing scheme and to infer the duplex stacking arrangements of these two HJs.

MATERIALS AND METHODS

The preparation of the immobile Holliday junctions J1 and J2 (Figure 1) has been described in detail elsewhere (Chen et al., 1991). Briefly, the four hexadecamer strands are synthesized, deblocked, and purified by standard methods, and then the oligodeoxynucleotide strands are carefully titrated

† Supported in part by a grant from the National Science Foundation (DMB 9019250 to W.J.C. and F.H.) and a fellowship to W.J.C. from the American Cancer Society (JFRA-294).

[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; d_c(A;B), cross-strand distance between protons A and B.

0006-2960/93/0432-319$04.00/0

to a 1:1:1:1 molar ratio before annealing. The buffer for the NMR samples contains 20 mM Tris-HCl at pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 5% D_2O, and 0.1% (w/v) NaN_3.

NMR experiments were performed using Bruker AM-500 and AM-600 spectrometers equipped with Aspect 3000 computers and digital phase shifting hardware. NOESY spectra were acquired with 200-ms mixing times at 293 K, 296 K (J2 only), 300 K, 305 K, and 310 K using the standard pulse sequence (Macura & Ernst, 1980), except that the last pulse was replaced by a jump-and-return composite sequence (Plateau & Guéron, 1982). No other means of solvent suppression was necessary. The excitation sequence was adjusted to give either one maximum just to the high-field side of the imino protons or two maxima, one each in the middle of the imino and the aromatic proton spectral regions. The pulse width and phase of the second 90° pulse in the jump-and-return sequence was adjusted, along with the receiver reference phase and preacquisition delay, in order to minimize baseline distortions and any need for postacquisition phase correction. For each t_1 value, 96 scans were acquired over a spectral width of 12 500 Hz into 8192 data points. The transmitter was placed on the solvent resonance, and the TPPI method was used to achieve quadrature phase detection in the ω_1 dimension (Marion & Wüthrich, 1983). In some of the spectra, the imino proton resonances (12–14 ppm) were folded in ω_1 into the low-field region of the spectrum that is devoid of signals (9–12 ppm). In all spectra, the number of t_1 values was set to give t_{1max} = 22 ms.

All data were processed using a CONVEX C240 computer with FTNMR software (Hare Research, Inc., Woodinville, WA). A convolution difference technique (Marion et al., 1989) was used to treat the time-domain data for removing the solvent signal prior to Fourier transformation (as implemented in FTNMR by Dr. M. Rance). Each data set was processed many times, with the parameters initially adjusted to achieve a balance between sensitivity and resolution and then to specifically enhance one extreme or the other. For example, the imino proton resonances have inherently large line widths (>50 Hz) and relatively weak NOEs, so when processing to examine the corresponding region of the NOESY spectrum, only the first ~40 ms of the FIDs was utilized and window functions were selected to enhance the sensitivity (typically squared sine-bell shifted by 90°). Stronger window functions (typically squared sine-bell shifted by 40°) and zero-filling were required for processing the ω_1 dimension to reduce artifacts related to the much poorer digital resolution in this dimension. For regions of the NOESY spectra containing resonances with narrower line widths, most or all of the FID was utilized and the window functions were adjusted to the desired compromise between sensitivity and resolution. The typical digital resolution of these spectra was 11.6 Hz/pt (8.0 Hz/pt for folded spectra) in the ω_1 dimension and 3.05 Hz/pt in ω_2.

RESULTS AND DISCUSSION

The critical first step in any detailed analysis by NMR spectroscopy is the assignment of resonances to specific atoms in the molecule. Sequence-specific assignment of the imino proton resonances in small polynucleotides was first obtained by the temperature dependence of ^{1}H NMR resonances and 1D NOE difference spectroscopy [reviewed in Wemmer and Reid (1985); Wüthrich, 1986]. For larger duplexes (M_r > 10 000), these methods cannot provide unambiguous resonance assignments, even with the use of high-field spectrometers.

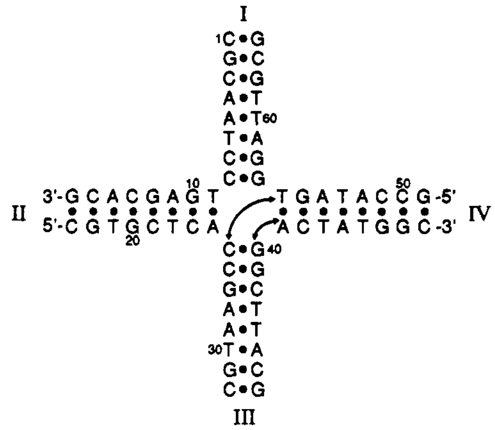

FIGURE 1: Sequence and numbering systems of J1 and J2. The sequence of J1 is shown, with the changes required to produce J2 indicated by the arrows: the G_{40}-C_{25} and T_{56}-A_{41} base-pairs in J1 exchange places, becoming A_{40}-T_{25} and C_{56}-G_{41} base-pairs in J2. The oligonucleotide sequence 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.

Two-dimensional NOESY spectra without solvent saturation (Boelens et al., 1985) allow the study of the labile protons of larger oligonucleotide structures, and the application of the method is now well-established [e.g., Rajagopal et al. (1988)]. However, the assignments reported here for the 32-base-pair model HJs represent a substantial increase over the largest oligonucleotide to be fully assigned by this method, a non-self-complementary 23-base-pair duplex (Grütter et al., 1988). The large size and complexity of the assignment problem for HJs necessitates some documentation of how the assignments have been obtained, and this information is provided in the text and Table I. Figures are included to demonstrate the overall quality of the spectra and to demonstrate the various types of NOEs that are used to make the assignments. A shorthand notation similar to that in Wüthrich (1986) will be utilized to specify NOEs and the corresponding interproton distances: intraresidue and sequential (5′ and 3′ direction) NOEs are indicated by d_i(A;B) and d_s(A;B), respectively; cross-strand NOEs are distinguished by an additional subscript c (e.g., d_{ci}(A;B), d_{cs}(A;B)).

The two 32-base-pair Holliday junctions selected for detailed analysis are shown in Figure 1 along with the numbering systems used to designate each nucleotide and the four arms. J1 was designed and has been extensively studied by Seeman, Kallenbach, and co-workers [e.g., Seeman and Kallenbach (1983), Kallenbach et al. (1983), Seeman et al. (1985), Marky et al. (1987), and Churchill et al. (1988)]. J2 is a simple permutation of J1, with an exchange of the base pairs at the junction between branches III and IV as indicated by the arrows in Figure 1. 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. The imino proton resonances of the four terminal base pairs of J1 and J2 will not be observed under the experimental conditions of pH and temperature used in these studies; thus, the assignment goal is for the 7 nonterminal base pairs in each arm, 28 in total.

The fundamental design of the sequence of J1 and J2 greatly facilitated sequence-specific assignment of the labile protons because each of the four arms of the HJ has a different arrangement of bases. In fact, for the assignment of the labile protons of J1 and J2, it is not even necessary to distinguish the internal order of base pairs (i.e., differentiate a G-C base

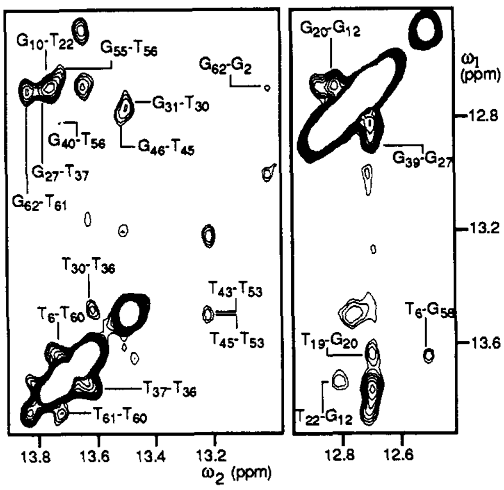

FIGURE 2: Region of the 600-MHz NOESY spectrum of 1 mM J1 containing cross-peaks between imino proton resonances. The spectrum was acquired at 300 K with a jump-and-return read pulse and a mixing time of 200 ms. The buffer is 20 mM Tris-d_{11}, 50 mM NaCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 0.1% (w/v) NaN_3, and 5% D_2O, with the pH adjusted to 7.5. A selected number of cross-peaks are identified by labeling with their sequence-specific assignment. This plot is made with a relatively high minimum contour level for clarity of presentation; 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.

pair from a C-G base pair, or an A-T base pair from a T-A base pair). For clarity, nondifferentiated sequences are represented by generic GC or AT identifiers, and the order of GC versus AT for each arm is termed the generic base-pair pattern. The generic base-pair patterns are sufficient to distinguish the four arms of J1 and all but arms I and IV of J2 (Figure 1). In the latter case, the unique assignments for J1 can be used to distinguish the corresponding resonances in J2. If it were necessary, the specific internal order within the base pairs could be established directly on the basis of the relative intensities of certain intrastrand and cross-strand sequential NOEs and, of course, by correlations to sequentially assigned nonlabile protons.

The uniquely low-field-shifted resonances of the imino protons in the region 12–15 ppm are the focus of the labile proton resonance assignment task. The most direct sequential assignment pathway is based on NOEs arising from the relatively short distances between imino protons of adjacent base pairs, 3.4 Å in canonical B-DNA. The corresponding region from a 600-MHz JR-NOESY spectrum of J1 is shown in Figure 2, demonstrating the overall high quality of the spectra, the clean separation of thymine N3H resonances from guanine N1H resonances, and the large number of cross-peaks observed. However, there are certain portions of the imino proton spectral range that are crowded with resonances precluding direct assignment of the corresponding cross-peaks. The acquisition of spectra at five different temperatures in the range 293–310 K proved useful for overcoming some of the problems due to resonance degeneracy, as the changes in chemical shift were sufficient to resolve some of the overlapped cross-peaks yet small enough to not necessitate reassigning the entire spectrum.

An alternative to direct imino–imino assignment involves two-step pathways, first establishing intra-base-pair correlations from the imino proton to the cytosine $4NH_2$ resonances within GC base pairs or to the adenine 2H resonance within AT base pairs, then following sequential NOEs from the $4NH_2$ or 2H resonances to adjacent imino protons. Wüthrich (1986) briefly discusses the strategy (Chapter 13.3) and tabulates the corresponding proton–proton distances (Tables 11.3 and 11.4). For GC base pairs, the two cytosine $4NH_2$ resonances can be readily identified by their characteristic chemical shifts and the very strong NOE between them, as shown in Figure 3. These are aligned with the corresponding imino proton by two d_{ci} (1;$4NH_2$) NOEs, one arising from the short distance between the guanine imino proton and the base-paired cytosine amino proton (2.5 Å) and the second due to spin diffusion from the base-paired proton to the non-base-paired amino proton. Examples of such cross-peak pairs are identified in the lower panel of Figure 4. The sequential assignments for the amino and imino protons can then be obtained from connectivities between the cytosine amino protons and the imino protons of adjacent base pairs [intrastrand or interstrand d(1,3;$4NH_2$) and d($4NH_2$;1,3)]. A number of such cross-peaks are present in the lower panel of Figure 4, labeled with lowercase letters. For AT base pairs, sequence-specific assignments can be made from the NOE between the thymine imino proton and its base-paired adenine 2H (d_i(3;2), ~2.8 Å), and then on to the adjacent imino protons via [d(1,3;2) and d(2;1,3)] sequential NOEs. Examples of these connectivities can be found in the upper panel of Figure 4. Although not used for making the assignments of J1 and J2, we note that the two-step assignment pathways have an additional useful property in that they are directional [e.g., $d_s(1,3;2) \neq d_s(2;1,3)$] and can be used to distinguish a G-C from a C-G or an A-T from a T-A base pair.

In contrast to studies on smaller DNA fragments [e.g., Chazin et al. (1986)], long-range connectivities and a limited degree of spin diffusion of the NOE were purposely sought after to assist in obtaining the assignments for the labile protons of J1 and J2; thus, ~200 ms mixing times were used for the NOESY spectra in this study. This provided a significant number of additional cross-peaks that proved useful for working through overlapped spectral regions and for confirming assignments.[2] For GC base pairs, if the cross-peaks corresponding to the direct NOEs between the imino and cytosine amino protons are not well-resolved due to resonance overlap of the imino proton, the d_{ci}(5;1) spin diffusion cross-peak can be useful to distinguish the exact imino proton resonance frequency. Figure 5 shows a spectral region containing a number of such cross-peaks. These peaks, in combination with the d_i(5;$4NH_2$) cross-peaks, can serve to indirectly correlate the guanine N1H and cytosine $4NH_2$ resonances of the base pair. Sequential assignment can then be made from the sequential NOEs to the resolved amino protons. For AT base pairs, the thymine $5CH_3$ resonances can be used, combining the d_i(Me;N3) NOE with d_s(Me;1,3), d_s(Me;$4NH_2$), and d_s(Me;2) NOEs.

The dominant feature of the approach taken for obtaining the sequential resonance assignments of J1 and J2 is the identification of as many of the observed NOE cross-peaks as possible and, correspondingly, to assign each base pair via as many sequential pathways as possible. This was deemed necessary for overcoming the complexity of the assignment problem and served to increase the reliability of the final result. It is important to note that the assignments were made from simultaneous analysis of all of the NOESY spectra for the two HJs, not from analysis of any one region of a single spectrum. The sequence-specific assignments obtained for the imino and resolved amino protons of J1 and J2 are listed

[2] For greater insight into the origins of these cross-peaks, the cross-relaxation pathways can be readily worked out from the proton–proton distances tabulated in Wüthrich (1986).

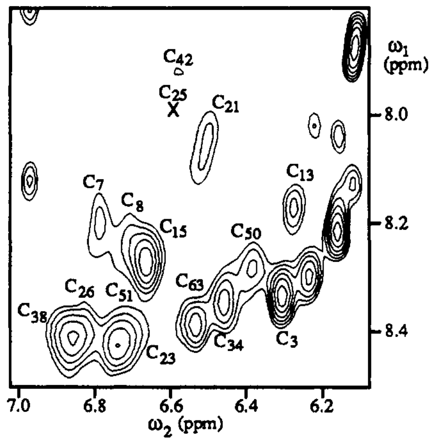

FIGURE 3: Spectral region of the same 600 MHz NOESY spectrum of J1 as shown in Figure 2, containing connectivities between the pairs of cytosine $4NH_2$ resonances. The assignments for all 15 cytosine residues that have observable cross-peaks at this pH and temperature are labeled with their sequence-specific assignments. The four cross-peaks corresponding to the $4NH_2$ resonances of the four terminal cytosine residues (C_1, C_{17}, C_{32}, C_{48}) are not observable at this temperature due to end-fraying. The cross-peak of amino protons of C_{25} is too weak to be seen at this contour level, and its position is marked with an X.

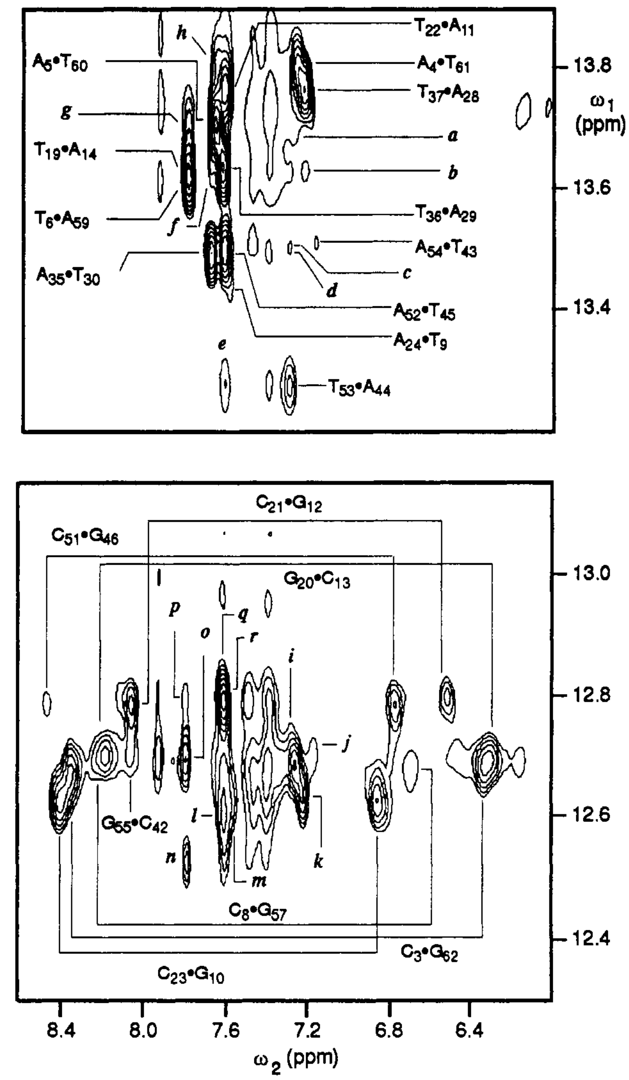

FIGURE 4: Regions of the 600-MHz NOESY spectrum of 1 mM J2 containing cross-peaks between imino and adenine 2H and cytosine $4NH_2$ resonances. The spectrum was acquired with a 200-ms mixing time at 296 K. The buffer was the same as that described for the solution of J1 in Figure 2. The imino protons of each base pair are abbreviated as G and T, and the adenine 2H and cytosine $4NH_2$ are abbreviated as A and C, respectively. The intra-base-pair thymine N3H imino to adenine 2H and guanine N1H to cytosine $4NH_2$ cross-peaks are directly labeled in the figure. Eighteen inter-base-pair NOE cross-peaks are identified by letters that correspond to (a) A_4-T_{60}, (b) T_{36}-A_{28}, (c) A_{44}-T_{43}, (d) A_{44}-T_{45}, (e) A_{52}-T_{53}, (f) A_{35}-T_{36}, (g) A_{59}-T_{60}, (h) A_5-T_{61} (i) A_4-G_{62}, (j) A_{54}-C_{55}, (k) A_{28}-G_{27}, (l) A_{11}-G_{10}, (m) A_{24}-G_{10}, (n) A_{59}-G_{58}, (o) A_{14}-G_{20}, (p) A_{14}-G_{18}, (q) A_{52}-G_{46}, (r) A_{11}-G_{12}. All contours seen in this figure correspond to cross-peaks that have been clearly distinguished in working plots made with much greater expansion and finer contouring. The unlabeled cross-peaks arise from amino proton resonances of adenine and guanine that have not been assigned.

in Table I. A summary of the NOEs identified is presented in Figure 6 and discussed in the following sections and the supplementary material. These comprise a completely de novo set of resonance assignments, independent of the sequence-specific assignments made for the nonlabile protons (S. Chen, F. Heffron, and W. J. Chazin, in preparation).

Imino Proton Resonance Assignments for J1. A good starting point for the ^1H NMR assignment of arm I is the segment of three consecutive AT base pairs (A_4-T_{61}, A_5-T_{60}, T_6-A_{59}). The T_{60} imino proton was readily identified at 13.71 ppm on the basis of its chemical shift, the intra-base-pair NOEs to the unique resonance frequencies of T_{60} $5CH_3$ (1.17 ppm) and A_5 2H (7.67 ppm), and fully resolved direct NOEs to the imino protons of T_6 and T_{60} (Figure 2). The assignment sequence is extended to the C_3-G_{62} imino proton by $d_s(1;3)$ and $d_{cs}(1;2)$ NOEs to A_4-T_{61} and in turn to G_2-C_{63} via an NOE between the imino protons of G_{62} and G_2 (Figure 2). The extension of assignments toward the junction from T_6-A_{59} was possible from direct $d_s(1;3)$ NOEs between T_6 and G_{58} (Figure 2) and a $d_s(1;1)$ between G_{58} and G_{57} that could be unambiguously identified only at 293 K. This set of connectivities gives a generic base-pair pattern of $[(GC)_2(AT)_3(GC)_2]$, which corresponds exclusively to arm I. The identification of the G_{57} imino proton resonance at a chemical shift of 12.82 ppm indicates full base-pairing in arm I, including the base pair at the junction.

The imino protons of arm II resonate at crowded spectral frequencies, which greatly reduces the possibility for obtaining assignments directly from NOEs between adjacent imino protons. However, the combination of direct and indirect assignment pathways involving adenine 2H, cytosine $4NH_2$, and cytosine 5H proved sufficient for making the assignments. Two blocks of three residues [(G_{18}-C_{15}, T_{19}-A_{14}, G_{20}-C_{13}) and (C_{21}-G_{12}, T_{22}-A_{11}, C_{23}-G_{10})] are readily identified by the observation of NOEs between a single adenine 2H resonance and three different imino proton resonances (Figure 6): A_{14} 2H exhibits cross-peaks to its base-pairing partner T_{19} N3H and to the two neighboring imino protons of G_{18} and G_{20} (Figure 4, peaks o and p); A_{11} 2H exhibits connectivities to the imino protons of its base-pairing partner T_{22} and the two adjacent residues, G_{10} and G_{12} (Figure 4, peaks l and r). These two blocks of three residues are aligned adjacent with each other by a critical direct NOE between the imino protons of G_{20} and G_{12} (Figure 2). In the overall sequence of J1, the generic [(GC)(AT)(GC)(GC)(AT)(GC)] pattern fits exclusively to arm II. The extension of the assignments to the T_9-A_{24} base pair at the junction is made on the basis of the expected intraresidue NOEs and a sequential NOE between A_{24} 2H and the neighboring imino proton of G_{10}. The identification of the T_9 imino proton resonance at a chemical shift of 13.61 ppm establishes that arm II is also fully base paired up to and including the junction.

Since the description of the sequential assignment of arms I and II covers the two primary procedures utilized, the level

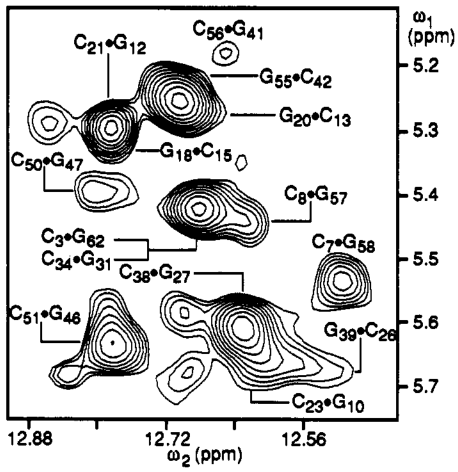

FIGURE 5: Region of the 600-MHz NOESY spectrum of 1 mM J2 containing intra-base-pair cross-peaks between guanine imino proton and cytosine 5H resonances. The spectrum was acquired at 305 K with a 200-ms mixing time from the same solution as for Figure 4. Unlabeled peaks arise from NOEs between guanine imino and amino proton resonances. The peak for G_2-C_{63} is observed outside the region shown at ω_1 = 5.44 ppm, ω_2 = 13.00 ppm.

of detail in the descriptions for arms III and IV will be reduced. For simplicity, we now skip to the assignment of the imino proton resonances of arm IV. The starting point is the imino proton resonance of G_{55}-C_{42} because it is fully resolved at 293 K and cross-peaks are observed to the imino proton resonances of both adjacent base pairs, T_{56}-A_{41} and A_{54}–T_{43}. Assignments were extended to base pairs T_{53}-A_{44}, A_{52}-T_{45}, and C_{51}-G_{46} from the system of connectivities involving the 2H resonances of the adenine residues (Figure 6), similar to the description given above for arm II. The series of sequential connectivities observed provide a [(GC)(AT)$_3$(GC)(AT)] generic base-pair pattern that is unique to arm IV. The chemical shift of the T_{56} imino proton (13.71 ppm) indicates base-pairing at the junction point in arm IV.

The assignment of the imino proton resonances of C_{34}-G_{31}, A_{35}-T_{30}, T_{36}-A_{29}, T_{37}-A_{28}, and C_{38}-G_{27} in arm III parallels that described above for arm II, i.e., extensive use of the interconnectivity of NOEs to adenine 2H [d(2;1,3), d(1,3;2), and d(2;2)]. For example, A_{35} 2H exhibits an NOE to the imino proton of its base-pairing partner T_{30}, NOEs to the two adjacent imino protons of G_{31}-C_{34} and A_{29}-T_{36}, and a fourth NOE to the 2H of A_{29}. Although the imino proton resonance of G_{27} is overlapped with other signals, an NOE from the imino proton of G_{27} to the resolved resonance of G_{39} N1H permits extension of the assignments to G_{39}-C_{26}. The series of connectivities for arm III (Figure 6) identifies a generic base-pair pattern of [(GC)(AT)$_3$(GC)$_2$] that fits the sequence of arms I, III, and IV; however, this set of resonances can be assigned to arm III by default, since the corresponding residues of arms I and IV have already been assigned.

To this point, the assignments for 26 of the 28 nonterminal base pairs of J1 have been accounted for. Two other imino proton resonances from GC base pairs can be distinguished in the NOESY spectra by their chemical shifts and intra-base-pair NOEs from the 5H and amino proton resonances to the imino proton resonances. Although sequential NOE connectivities for them cannot be unambiguously identified due to spectral overlap problems, these resonances can be assigned by default to the only two nonterminal base pairs for which no imino proton assignments had been made, C_{25}-G_{40} and G_{47}-C_{50}. The two resonances are differentiated on the basis of the line widths observed for the various protons within each base pair. G_{47}-C_{50} is located one base pair from the

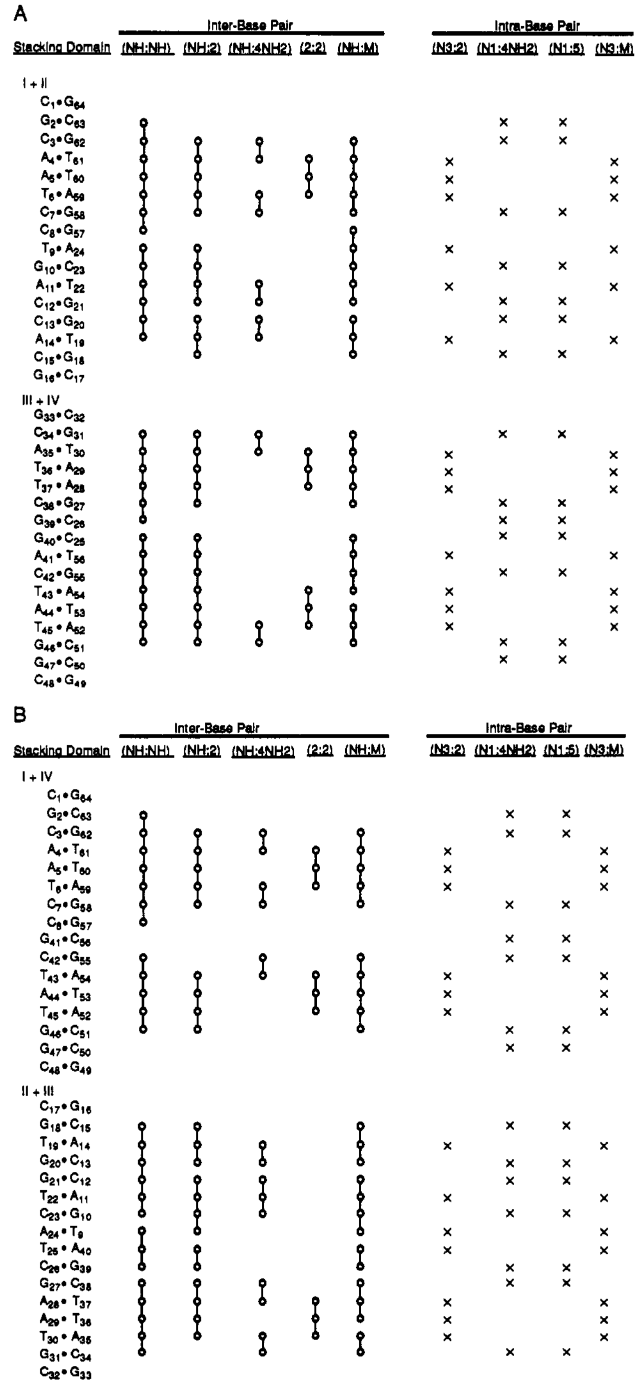

FIGURE 6: Summary of NOE cross-peaks involving the labile protons of J1 (A) and J2 (B) observed in 2D NOESY spectra obtained at 293 K, 296 K, 300 K, 305 K, and 310 K.

terminus of duplex arm IV and exhibits fairly sharp resonance lines, whereas C_{25}-G_{40} has rather broad resonance lines, even for the nonlabile cytosine 5H [see Figure 2 of Chen et al. (1991)], that are characteristic of the base pairs at the junction. The assignment of C_{25}-G_{40} was corroborated by NOEs between the two arms in stacking domain, vide infra. With these two additional assignments, the resonances for all 28 imino protons of J1 are assigned, and all eight residues at the junction are shown to be fully base paired.

An examination of the imino proton portion of the ¹H NMR spectrum of J1 was reported several years ago (Wemmer et al., 1985). Their strategy consisted of preparation of the four oligonucleotide duplexes corresponding to the four arms of the HJ, sequential assignment of the imino proton resonances

Table I: [1]H NMR Assignments of Adenine 2H and Exchangeable Protons of J1 and J2 at pH 7.5, 300 K[a]

arm	imino	$C4NH_2$ 2, 1	A2H
		Arm I	
C_1-G_{64}			–
G_2-C_{63}	13.00	8.39, 6.53	–
C_3-G_{62}	12.71	8.31, 6.31	–
A_4-T_{61}	13.82	–	7.30
A_5-T_{60}	13.71	–	7.67
T_6-A_{59}	13.64	–	7.79
C_7-G_{58}	12.49	8.19, 6.74	–
	12.53	8.19, *6.68*	–
C_8-G_{57}	12.82	8.24, 6.69	–
	12.63	8.27, 6.68	–
		Arm II	
C_{17}-G_{16}			–
G_{18}-C_{15}	12.83	8.27, 6.67	–
T_{19}-A_{14}	13.62	–	7.78
G_{20}-C_{13}	12.69	8.15, 6.28	–
C_{21}-G_{12}	12.81	8.03, 6.49	–
T_{22}-A_{11}	13.75	–	7.61
G_{23}-G_{10}	12.69	8.41, 6.72	–
	12.61	8.42, *6.84*	–
A_{24}-T_9	13.61	–	7.69
	13.50	–	*7.62*
		Arm III	
G_{33}-C_{32}		7.83[b]	–
C_{34}-G_{31}	12.67	8.34, 6.47	–
A_{35}-T_{30}	13.48	–	7.68
T_{36}-A_{29}	13.61	–	7.63
T_{37}-A_{28}	13.75	–	7.26
C_{38}-G_{27}	12.70	8.41, 6.87	–
	12.64	8.40, 6.84	–
G_{39}-C_{26}	12.87	8.41, 6.85	–
	12.57	8.40, *6.78*	–
G_{40}-C_{25}	12.82	8.02, 6.65[c]	–
A_{40}-T_{25}	*13.78*[d]	–	7.80
		Arm IV	
G_{49}-C_{48}			–
C_{50}-G_{47}	12.77	8.29, 6.40	–
C_{51}-G_{46}	12.75	8.43, 6.74	–
	12.80	8.47, 6.75	–
A_{52}-T_{45}	13.49	–	7.59
T_{53}-A_{44}	13.21	–	7.21
	13.27	–	*7.32*
A_{54}-T_{43}	13.50	–	6.99
	13.50		*7.19*
G_{55}-C_{42}	12.62	7.94, 6.59	–
	12.73	*8.04*, *6.43*	–
T_{56}-A_{41}	13.71	–	7.79
C_{56}-G_{41}	*12.69*	*8.10*, *6.36*	–

[a] Chemical shifts referenced to the ^1H_2O signal, precalibrated using methanol. The majority of chemical shifts of these two HJs are identical within the experimental error of each measurement (±0.02 ppm); hence, separate listings for J2 are included just below the corresponding value for J1 only when at least one of the resonances of the base pair are >±0.04 ppm different. Significant chemical shift differences are listed in italics. A dash (–) is given for entries that are not applicable for a particular base pair. [b] Assigned for J2 only. [c] Assignments obtained at 305 K. [d] Assignment obtained at 293 K.

for each of the duplexes by 1D NOE experiments (the requisite 2D techniques had not yet been developed), and then transfer of as many of the assignments as possible to the intact HJ by comparison of 1D ^1H NMR spectra. Although a limited number of the assignments could be corroborated by 1D NOE difference experiments on the intact junction, the severe overlap in 1D spectra precluded unambiguous sequence-specific assignment. However, the presence of additional resonances in the spectrum of the HJ was inferred from a comparison with the summed spectra of the individual arms. These additional resonances indicated, albiet somewhat indirectly since the specific assignments were lacking, that all bases up to and including the junction are paired (Wemmer et al., 1985). This hypothesis, consistent with analysis of the CD spectrum (Marky et al., 1987) and studies using chemical probes (Gough et al., 1986; Furlong & Lilley, 1986), is now directly and unambiguously validated by the analysis of J1 base-pairing reported here.

Imino Proton Assignments for J2. Since the base sequence of J2 is almost identical to that of J1 except for the two base pairs that are exchanged between arms III and IV (Figure 1), the chemical shifts of most resonances in these two HJs are expected to be very similar. In fact, the spectra for J1 and J2 were analyzed in parallel, making use of their differences in resonance degeneracy problems to facilitate the assignment process. However, due to the complexity of the assignment problem, we have been careful to independently check the assignments against the NOE connectivities observed in each of the NOESY spectra. Since the assignments of the imino proton resonances of J1 have been described in detail, only a brief summary of the J2 assignments is given here. Full details are provided in the supplementary materials.

Unambiguous sequence-specific assignments were obtained from the observed NOEs for 24 of the 28 imino proton resonances of J2, leaving G_{39}-C_{26}, A_{40}-T_{25}, C_{50}-G_{47}, and C_{56}-G_{41}. Since there was only one set of resonances remaining with chemical shifts characteristic of an AT base pair, these could be assigned by default to A_{40}-T_{25}. This assignment was corroborated by the large line widths observed for the A_{40}-T_{25} resonances, which is characteristic of the base pairs at the junction. The adjacent base pair in arm III (G_{39}-C_{26}) could be assigned from NOE connectivities to A_{40}-T_{25}. In retrospect, it was evident that connectivities between G_{39}-C_{26} and G_{38}-C_{27} could not be identified due to resonance degeneracies. Of the two sets of GC resonances remaining, the large line width of the imino proton resonance of the base pair at the junction (C_{56}-G_{41}) served to distinguish the C_{56}-G_{41} base pair from C_{50}-G_{47}. The assignment of C_{50}-G_{47} was confirmed by the virtual identity of the proton resonance frequencies with those found for the corresponding resonance in J1. These additional assignments brought the total number of imino proton resonances identified to 28 as expected, and showed that as for J1 the eight residues at the junction of J2 are all fully base paired.

Amino Protons. The only amino protons that have a slow enough rate of exchange with solvent to consistently exhibit discrete resolved resonances under the experimental conditions used in these studies are those from the amino group of cytosine. These proton resonances are easily distinguished by their unique chemical shifts and, as shown in Figure 3, by a strong cross-peak between each pair of protons due to a combination of strong NOE and chemical exchange. Figure 3 also shows that nearly all of the cross-peaks are well-resolved. The hydrogen bonded (NH_2-2) and non-hydrogen bonded (NH_2-1) protons are distinguished by their relative chemical shifts, with the hydrogen bonded proton resonating at lower field. As noted above, these resonances proved to be of great value in obtaining the sequence-specific assignments because both intra-base-pair and sequential $d(1,3;4NH_2)$ and $d(4NH_2;1,3)$ connectivities can be observed. In addition to the cross-peaks from the 2H, $4NH_2$, 5H, and 6H resonances, there are some broad, low-intensity cross-peaks observed to the imino proton resonances from signals resonating at 5.5–7.5 ppm that are attributed to the amino protons of adenine and guanine bases. These resonances are very broad due to rapid rotation about the C–N bonds [e.g., Boelens et al. (1985) and Rajagopal et al. (1988)]. Only a few of them exhibited identifiable NOEs that could be used for making specific assignments.

Stacking Domains. The sequence-specific assignments in Table I provide the necessary background for obtaining information on the nature of the duplex stacking arrangements of J1 and J2. As mentioned in the Introduction, all available evidence to date indicates that HJs form pairs of smoothly stacked helical duplex domains. Previous studies of J1 have indicated that arm I stacks with arm II and arm III with arm IV (Churchill et al., 1988; Chen et al., 1988). Since the experimental conditions have been carefully controlled, the differences in the chemical shifts observed for J1 and J2 can be related to the relative geometry of the stacking domains, i.e., chemical shift differences must be caused by the differences in base sequence in arms III and IV (G_{40}-C_{25} and T_{56}-A_{41} in J1 become A_{40}-T_{25} and C_{56}-G_{40} in J2) and any corresponding changes induced in the adjacent residues. If one assumes that J1 and J2 have an identical domain structure with arm III stacked on arm IV, then the differences in chemical shift should be limited to the III/IV stacking domain, as there are no changes in arms I and II. However, we find that the chemical shifts in J2 for the imino protons of the base pairs at the junction of arm I (C_8-G_{57}) and arm II (A_{24}-T_9) are also substantially different (>0.1 ppm) from J1. This observation could be explained by both HJs having the same structure with arm I stacked on arm IV and arm II on arm III, but such an explanation would be in conflict with all of the previously mentioned evidence for the J1 arm stacking geometry. The more likely explanation is that the two junctions have different arm stacking arrangements.

These two possibilities can be differentiated by NOEs between the two duplexes within a stacking domain. Once the sequence-specific assignments were made, the NOESY spectra were further analyzed for interproton contacts that span the junction points of J1 and J2. These correspond to standard sequential NOEs, but they occur between residues that are not consecutive in sequence. The observation in J1 of d_{cs}(G_{57} N1H;T_9 Me) and d_{cs}(G_{57} N1H;A_{24} 2H) NOEs indicate that arm I must stack on arm II, and observation of d_{cs}(G_{40} N1H;T_{56} N3H), d_s(G_{40} N1H;A_{41} 2H), and d_{cs}(G_{40} N1H;T_{56} Me) NOEs indicate that arm III stacks on arm IV. This result is consistent with previous studies of J1 from other laboratories on the basis of chemical footprinting experiments (Churchill et al., 1988; Chen et al., 1988). It is of importance to note that not all of the expected NOEs are observed between the stacked arms across the junction. This occurs for two reasons. First, it is logical to assume some degree of structural perturbation at the junction [e.g., von Kitzing et al. (1990)], and in fact, sequential and intraresidue NOEs between protons of the base and sugar rings indicate some deviations from idealized B form geometry (S. Chen, F. Heffron, and W. J. Chazin, in preparation). Second, the resonance line widths of the residues at the junction tend to be broad, a property attributed to partial melting of these base pairs at the high temperatures required for NMR analysis of these structures (Chen et al., 1991), which makes detection of the associated cross-peaks more difficult.

The direct identification of I/II and III/IV stacking in J1 implies that the chemical shift differences between J1 and J2 can be attributed to differences in their arm stacking arrangement. Due to a greater degree of spectral overlap problems for the crucial junction residues, there are fewer unambiguously identified NOEs that span the junction point of J2. However, the NOE between the imino protons of T_9 and T_{25} is clearly resolved indicating that in contrast to J1 arm II stacks on arm III. Thus, both chemical shift differences and NOEs are consistent in indicating that the arm stacking domains of J1 and J2 are different. The difference in the arm stacking arrangements of J1 and J2 is also reflected in their different mobilities on polyacrylamide gels (Chen et al., 1991) and suggests that their three-dimensional structures are different. The change in arm stacking arrangement and structural properties in response to changes in the sequence at the junction is consistent with the concept of HJs playing an important role in determining the outcome of recombinational events.

CONCLUDING REMARKS

Although arm stacking arrangements of immobile HJs can be identified by other techniques [e.g., Cooper and Hagerman (1987, 1989), Churchill et al. (1988), Duckett et al. (1988), and Murchie et al. (1989)], NMR studies are unique in providing the identity of the stacking geometry by a direct method. Comparisons of the chemical shifts of J1 and J2 (Table I) suggest that disturbances in response to the changes of arm stacking partners and base sequences extend at least one base pair in each arm. Our data also suggest that arm IV may be affected up to four base pairs removed from the junction point. It has been pointed out earlier by a calorimetric study of J1 that arm IV has a lower T_m (~10 K) than that of the other three arms (Marky et al., 1987). The origin of the difference in arm IV between J1 and J2 can only be clarified by further experiments and a clearer understanding of the subtleties of the sequence dependence of duplex DNA structure and dynamics.

The complete sequence-specific assignments for J1 and J2 lay the groundwork for more detailed analysis of the structure and dynamics of these HJs by NMR, which in turn will guide the design of experimental strategies to probe the molecular determinants of the sequence dependence of HJ structure. In this report, several interarm NOEs are identified that implicate differences between the arm stacking geometries adopted by J1 and J2, but further information is required to unambiguously distinguish among the various structural models for the Holliday junction. Efforts to obtain this information are already underway in this laboratory by the combination of molecular simulation and experiment. Detailed analysis of HJ structure and dynamics should provide important information for understanding the role of Holliday junctions in genetic recombination and repair.

ACKNOWLEDGMENT

We thank Dr. Mark Rance for continued support with experimental methodology and Beth Larson for help with typing the manuscript.

SUPPLEMENTARY MATERIAL AVAILABLE

Two pages of text describing in detail the assignment of the imino proton resonances of J2. Ordering information is given on any current masthead page.

REFERENCES

Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. (1985) *J. Magn. Reson. 62*, 378–386.

Chazin, W. J., Wüthrich, K., Hyberts, S., Rance, M., Denny, W. A., & Leupin, W. (1986) *J. Mol. Biol. 190*, 439–453.

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.

Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R., & Seeman, N. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4653–4656.

Clegg, R. M., Murchie, A. I. H., Zechel, A., Carlberg, C., Diekmann, S., & Lilley, D. M. J. (1992) *Biochemistry 31*, 4846–4856.

Cooper, J. P., & Hagerman, P. J. (1987) *J. Mol. Biol. 198*, 711–719.

Cooper, J. P., & Hagerman, P. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4653–4656.

Duckett, D. R., Murchie, A. I. H., Diekmann, S., Kitzing, E., Kemper, B., & Lilley, D. M. J. (1988) *Cell 55*, 79–89.

Furlong, J. C., & Lilley, D. M. J. (1986) *Nucleic Acids Res. 14*, 3995–4007.

Gough, G. W., & Lilley, D. M. J. (1985) *Nature 313*, 154–156.

Gough, G. W., Sullivan, K. M., & Lilley, D. M. J. (1986) *EMBO J. 5*, 191–196.

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.

Lu, M., Guo, Q., Seeman, N. C., & Kallenbach, N. R. (1989) *J. Biol. Chem. 264*, 20851–20854.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Marky, L. A., Kallenbach, N. R., McDonough, K. A., & Seeman, N. C. (1987) *Biopolymers 26*, 1621–1634.

Murchie, A. I. H., Clegg, R. M., Kitzing, E. V., Duckett, D. R., Diekmann, S., & Lilley, D. M. J. (1989) *Nature 341*, 763–766.

Murchie, A. I. H., Portugal, J., & Lilley, D. M. J. (1991) *EMBO J. 10*, 713–718.

Plateau, P., & Guéron, M. J. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.

Rajagopal, P., Gilbert, D. E., van der Marel, G. A., van Boom, J. H., & Feigon, J. (1988) *J. Magn. Reson. 78*, 526–537.

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., 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.

Seeman, N. C., Chen, J.-H., & Kallenbach, N. R. (1989) *Electrophoresis 10*, 345–354.

Sigal, N., & Alberts, B. (1972) *J. Mol. Biol. 71*, 789–793.

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., & Reid, B. R. (1985) *Annu. Rev. Phys. Chem. 36*, 105–137.

Wemmer, D. E., Wand, A. J., Seeman, N. C., & Kallenbach, N. R. (1985) *Biochemistry 24*, 5745–5749.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.