Articles

DNA Junctions, Antijunctions, and Mesojunctions†

Shou Ming Du, Siwei Zhang, and Nadrian C. Seeman*

Department of Chemistry, New York University, New York, New York 10003

Received March 19, 1992; Revised Manuscript Received July 15, 1992

ABSTRACT: Antijunctions and mesojunctions are new classes of multistranded DNA complexes. They represent a generalization of DNA branched junction complexes, such as the Holliday recombination intermediate. Each strand of a conventional branched junction participates in two different double helices, and this is also true for mesojunctions and antijunctions. The helix axes of conventional branched junction complexes may be drawn to converge at a point, but this convergence occurs for lines drawn perpendicular to the helix axes of antijunctions. Mesojunctions are complexes that mix these features of junctions and antijunctions. Antijunction complexes require an even number of strands. We have synthesized the mesojunction containing three strands, the two mesojunctions containing four strands, and the antijunction containing four strands; we compare them with branched junctions containing three and four strands, derived by permutations of the same sequences. Each double helix is designed to contain 1.5 turns of DNA. A tendency to oligomerize makes it difficult to capture antijunctions and mesojunctions in stable discrete complexes, in contrast to conventional branched junctions. For both three-strand and four-strand complexes, T_m is highest for conventional branched junctions. Ferguson analysis indicates similarities in the occluded surface area of junctions, antijunctions, and one four-strand mesojunction, but the other four-strand mesojunction has a much lower apparent surface area. Hydroxyl radical cleavage patterns suggest that the four-strand antijunction and the low-surface-area four-strand mesojunction form stacking domains, analogous to the behavior of conventional branched junctions. These new structures are related to replicational and recombinational intermediates and to single-stranded nucleic acid knots.

The last decade has seen a large amount of interest in multistranded DNA complexes containing more than a single double helical component. The most prominent of these complexes are the immobile (Seeman, 1982; Kallenbach et al., 1983) and partially mobile (Seeman, 1982; Chen et al., 1988) DNA branched junctions, particularly four-arm junctions, which are analogs of the recombinational Holliday (1964) junction. Structural (Cooper & Hagerman, 1987, 1989; Churchill et al., 1988; Chen et al., 1988; Duckett et al., 1988; Murchie et al., 1989), physical (Seeman et al., 1985; Wemmer et al., 1985; Marky et al., 1987; Lu et al., 1991, 1992), substrate (Mueller et al., 1988, 1990; Duckett et al., 1988), and liganding properties (Guo et al., 1989) of three-arm and four-arm junctions have been determined; five-arm and six-arm branched junctions have also been constructed and characterized (Wang et al., 1991). In addition, branched DNA junctions have been used as the building blocks of DNA geometrical objects, including a quadrilateral (Chen et al., 1989) and a hexacatenane whose helix axes have the connectivity of a cube (Chen & Seeman, 1991).

Recently, we have constructed a trefoil knot molecule from a synthetic single-stranded DNA molecule containing two double helical segments (Mueller et al., 1991). In attempting to generalize this system to a design scheme for more intricate knots (Seeman, 1992), it has become evident that branched junctions are only a single member of the class of double-helical, multistranded molecules in which each strand participates in two different double helices. The other members of this class of molecules, antijunctions and mesojunctions, are introduced here.

Figure 1a contains a diagrammatic representation of three-strand and four-strand DNA complexes. The molecular backbones are shown as solid lines, with an arrowhead denoting the 3′ end of each strand. It is convenient to visualize the new types of molecules by drawing them so that they flank a central polygon. The assembly of strands creates a central triangle or square, whose vertices are formed by the right-handed crossing of two backbones. Approximately one half-turn of duplex DNA is indicated by thin parallel lines flanking each vertex; these parallel lines are drawn to suggest base pairs between antiparallel backbones. The axis of each of these local helices is indicated by a pair of thin lines perpendicular to the base pairs. A conventional three-arm branched junction flanks a triangle at the upper left, and a four-arm branched junction flanks a square in the upper center. In both of these cases, the helix axes point radially from the centers of the polygons. The other structures illustrated are representatives of the extended class of molecules. It is simplest to begin with the structure at the upper right: None of its helix axes points toward the center of the square, rather they point in a circumferential direction about the perimeter of the square. It can be seen that four strands form a squarelike double-helical object. The helix axes are all perpendicular to the helix axes of the four-arm branched junction. We term this structure an "antijunction".

Is it possible to flank the polygon with structures in which the helix axes are not exclusively radial or tangential? The answer is yes, but these "mesojunctions" can be made only within limits. If one insists that all nucleotides be pairable, and that long nonpairing loops are forbidden, then one must

† This research has been supported by Grant GM-29554 from the NIH.

0006-2960/92/0431-10955$03.00/0

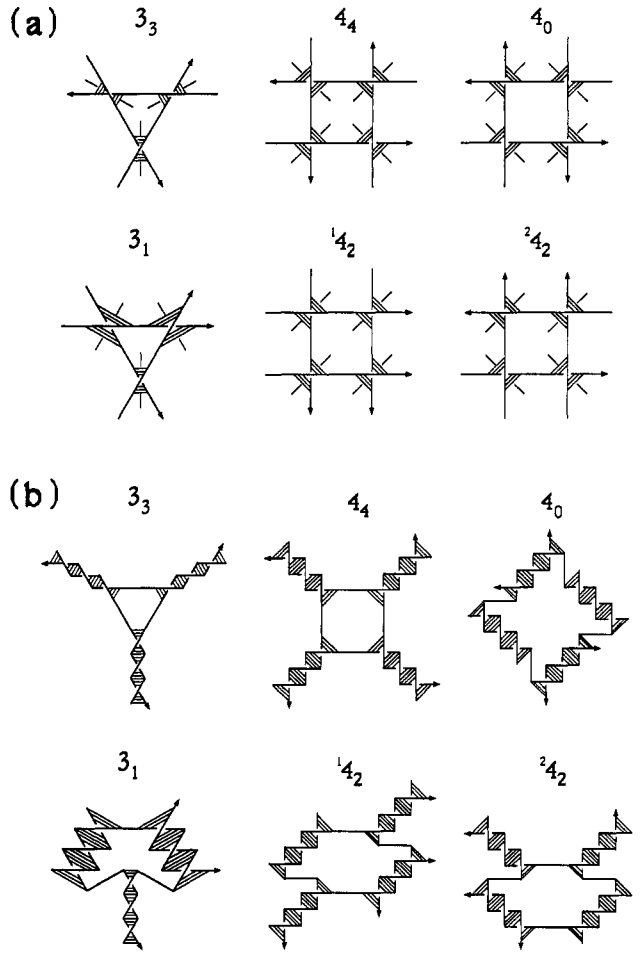

FIGURE 1: Schematic drawings of three-strand and four-strand junctions, antijunctions, and mesojunctions. (a) The helical arrangements that can flank a triangle or a square. Each polygon is flanked by strands of DNA that extend beyond the vertices in each direction. The arrowheads indicate the 3′ ends of the strands. The vertices correspond to the nodes formed by approximately a half-turn of double helical DNA. Base pairs are represented by lines between antiparallel strands. Thin lines perpendicular to the base pairs represent the axis of each helix. The complexes 3_3 and 4_4 correspond to conventional branched junctions. Their helix axes are all pointing toward the center of the triangle or square along the bisectors of the angles of the polygons "radial". The complex 4_0 is a four-strand antijunction, in which each helix axis ideally is perpendicular to the angular bisectors ("circumferential"). The complexes are named by a number, which indicates the number of double helical segments, followed by a subscript, which indicates how many of them are pointed along the angular bisectors, and preceded by a superscript (if necessary) enumerating different members of a class otherwise undifferentiated by the first two numbers. The complexes on the bottom row contain a mix of the two orientations of helix axes. The complex 3_1 contains one radial helix axis and two circumferential axes. The complexes $^1\!4_2$ and $^2\!4_2$ each contain a pair of radial helix axes and a pair of circumferential axes, but the arrangement is different between them: $^1\!4_2$ contains alternating radial and circumferential axes, whereas $^2\!4_2$ contains two radial axes on the right of the square and two circimferential axes on the left. (b) The complexes studied here. Each molecule corresponds to the one in the same position in part a. However, here we have indicated 1.5 turns of DNA per double helical segment, rather than 0.5 turns of DNA, as shown in a. The corners of the polygons have been preserved from a. This results in apparent unpaired regions between double helices; these are artifacts of this representation and are not designed into the molecules. The structures have been symmetrized somewhat, but no structural implications should be drawn. $^2\!4_2$ has been rotated 90° from a.

change the orientations of two helices simultaneously. Thus, no three-arm antijunction can be made from three strands of DNA that flank a triangle, but a three-strand mesojunction (3_1, lower left in Figure 1) is feasible. Likewise no arrangement of three helices and one helix can be made, but two different four-strand mesojunctions ($^1\!4_2$ and $^2\!4_2$, lower center and lower right, Figure 1a) can be made containing two radial and two tangential helices. The sequences of the component strands of each of the structures shown can be derived by switching the 5′ and 3′ sections of the strands that form the branched junctions. Higher order antijunctions and mesojunctions that combine radial and tangential components to different extents are possible in principle. We name the structural types by designating the number of strands, followed by a subscript indicating how many of the double helices are radial. Thus the conventional four-arm junction is identified as 4_4, and the four-strand antijunction is termed 4_0. The various mesojunctions are enumerated by a preceding superscript, indicating a serial number. Table I lists the antijunctions and mesojunctions for complexes containing as many as 12 strands. The entries in the table are derived from the formula for determining the number of different necklaces that can be made with two differently-colored beads (Constantine, 1987). The bulk of the entries in the table are in fact mesojunctions. Conventional branched junctions correspond to the unitary entries down the first column of the table. Antijunctions correspond to the unit entries on the right edge of each row for even-stranded complexes; there are no antijunctions for odd-stranded complexes.

Table I: Populations of Junctions, Antijunctions, and Mesojunctions[a]

no. of arms	no. of circumferential arms							
	0	2	4	6	8	10	12	total
3	1	1						2
4	1	2	*1*					4
5	1	2	1					4
6	1	3	3	*1*				8
7	1	3	4	1				9
8	1	4	8	4	*1*			18
9	1	4	10	7	1			23
10	1	5	16	16	5	*1*		44
11	1	5	20	26	10	1		63
12	1	6	29	50	29	6	*1*	122

[a] Conventional junctions are underlined. Antijunctions are in italics. All other structures indicated are mesojunctions. The total listed on the right refers to the total number of structures with the number of arms indicated on the left. The number of arms is equivalent to the number of strands.

We have constructed molecules corresponding to all of the radial and tangential combinations shown in Figure 1a. The structures displayed there contain a single half-turn in each of their double helical components. In fact, any number of half-turns of DNA can be accommodated in each of the double helical domains. In order to ensure the stabilities of the complexes synthesized, we have built them out of molecules that contain approximately 1.5 turns of DNA in each helical segment, rather than the half-turn shown in Figure 1a. These molecules are illustrated in Figure 1b. The three-arm and four-arm branched junctions JYG and JXG have been reported previously (Wang et al., 1991). Here, we report the construction and characterization of each of the new molecules by means of native gel electrophoresis, Ferguson (1964) analysis, and thermal denaturation; preliminary structural characterization by means of hydroxyl radical autofootprinting analysis (Tullius & Dombroski, 1985, 1986; Churchill et al., 1988) has also been performed.

MATERIALS AND METHODS

Sequence Design. The sequences of the branched junctions JYG and JXG have been designed by SEQUIN (Seeman,

1990); JXG is an extension of J1 (Seeman and Kallenbach, 1983). Strands are named with a suffix of 12 or 21, depending upon the order of the 5′ and 3′ halves being the same (12) or reversed (21) from that of the strand that is a component of the conventional branched junctions. The sequence of JYG is as follows:

strand Y112:
GGAAGTGCTCCGTAGCGATGGTGCGTGATTGC

strand Y212:
GCAATCACGCACCATCCTGCTACGACTCAACG

strand Y312:
CGTTGAGTCGTAGCAGGCTACGGAGCACTTCC

The sequence of JXG is as follows:

strand X112:
GGTAGGACCGCAATCCTGAGCACGTCTCAACG

strand X212:
CGTTGAGACGTGCTCACCGAATGCACAGTTCC

strand X312:
GGAACTGTGCATTCGGACTATGGCACTGGTGG

strand X412:
CCACCAGTGCCATAGTGGATTGCGGTCCTACC

Additional strands have been synthesized for this study; the sequences of these strands are determined by reversing the 3′ and 5′ halves of the strands above with the same two first letters in their names:

strand Y321:
GCTACGGAGCACTTCCCGTTGAGTCGTAGCAG

strand X221:
CCGAATGCACAGTTCCCGTTGAGACGTGCTCA

strand X321:
ACTATGGCACTGGTGGGGAACTGTGCATTCGG

strand X421:
GGATTGCGGTCCTACCCCACCAGTGCCATAGT

The conventional branched junctions (3_3 and 4_4) are constructed from equimolar mixtures of [Y112, Y212, Y312] and [X112, X212, X312, X412], respectively. The mesojunction 3_1 is constructed from an equimolar mixture of [Y112, Y212, Y321]. The four-strand antijunction 4_0 is formed from an equimolar mixture of [X112, X221, X312, X421]. The four-strand mesojunction $^14_2$ has the components [X112, X212, X321, X321, X421] and the four-strand mesojunction $^24_2$ has the components [X112, X221, X312, X412].

Synthesis and Purification of DNA. All DNA molecules in this study have been synthesized on an Applied Biosystems 380B automatic DNA synthesizer, removed from the support, and deprotected, using routine phosphoramidite procedures (Caruthers, 1982). DNA strands have been purified by HPLC, utilizing a Du Pont Zorbax Bio Series oligonucleotide column using a gradient of NaCl in a solvent system containing 20% acetonitrile and 80% 0.016 M sodium phosphate, as described previously (Wang et al., 1991).

Formation of Hydrogen-Bonded Complexes. Complexes are formed by mixing a stoichiometric quantity of each strand, as estimated by OD_{260}. This mixture is then heated to 90 °C for 3 min and slowly cooled to the desired temperature. Stoichiometry is estimated by native gel electrophoresis of adjacent dimers (e.g., X112 and X212). A single band on a native gel is taken to indicate a homogeneous stoichiometric complex.

Thermal Transition Profiles. Response to increased temperature is measured on a Gilford Response II UV–vis spectrophotometer at 268 nm. Complexes are formed at a concentration of 500 nM for each strand, in a solution containing 40 mM sodium cacodylate, pH 7.0, and 10 mM $MgCl_2$. Temperature is incremented at the rate of 0.1 °C/min.

Hydroxyl Radical Analysis. Individual strands of the complexes are radioactively labeled, and are additionally gel purified from a 15% denaturing polyacrylamide gel. For 4_4, 3_3, and $^14_2$, each of the labeled strands [approximately 1 pmol in 50 mM Tris·HCl (pH 7.5) containing 10 mM $MgCl_2$] is annealed to a 100-fold excess of the unlabeled complementary strands, it is annealed to a 100-fold excess of a mixture of the other strands forming the complex, it is left untreated as a control, or it is treated with sequencing reagents (Maxam and Gilbert, 1977) for a sizing ladder. As discussed below, 4_4, 3_3, and $^14_2$ are stable complexes that form readily at high strand concentrations; hence, we can be confident of examining the labeled strand in the appropriate complex when an excess of its unlabeled partners is added to it, as done in previous studies (Churchill et al. 1988). 3_1, 4_0, and $^24_2$ do not form readily at high strand concentrations; consequently, hydroxyl radical analyses are performed under the low, equimolar strand concentrations shown below to be consistent with specific complex formation. The samples are annealed by heating to 90 °C for 3 min and then cooled slowly to the appropriate temperature for complex formation (vide infra). Hydroxyl radical cleavage of the double-strand and junction samples for all strands takes place for 2 min (Tullius & Dombroski, 1985), with modifications noted by Churchill et al. (1988). The reaction is stopped by addition of thiourea. The sample is dried, dissolved in a formamide/dye mixture, and loaded directly onto a 15% polyacrylamide/8.3 M urea sequencing gel. Autoradiograms are scanned with a Hoefer GS300 densitometer in transmission mode.

Polyacrylamide Gel Electrophoresis. (*A*) *Denaturing Gels.* These gels contain 8.3 M urea and are run at 55 °C. Gels contain 10% acrylamide (19:1, acrylamide/bisacrylamide). The running buffer consists of 89 mM Tris·HCl, pH 8.0, 89 mM boric acid, and 2 mM EDTA (TBE). The sample buffer consists of 10 mM NaOH and 1 mM EDTA, containing 0.1% Xylene Cyanol FF tracking dye. Gels are run on an IBI Model STS 45 electrophoresis unit at 70 watts (50 V/cm), constant power, dried onto Whatman 3MM paper, and exposed to X-ray film for up to 15 h.

(*B*) *Native Gels.* Gels contain 12% or 15% acrylamide (19:1, acrylamide/bisacrylamide). DNA is suspended in 10–25 μL of a solution containing 40 mM Tris·HCl, pH 8.0, 20 mM acetic acid, 2 mM EDTA, and 12.5 mM magnesium acetate (TAEMg); the quantities loaded vary as noted below. The solution is boiled and allowed to cool slowly to 45 °C. Samples are then brought to a final volume of 10 μL with a solution containing TAEMg, 50% glycerol, and 0.02% each of Bromophenol Blue and Xylene Cyanol FF tracking dyes. Gels are run on a Hoefer SE-600 gel electrophoresis unit at 8 V/cm at 4 or 45 °C and exposed to X-ray film for up to 15 h or stained with Stainsall dye. Absolute mobilities (cm/h) of native gels run at 45 °C are measured for Ferguson analysis;

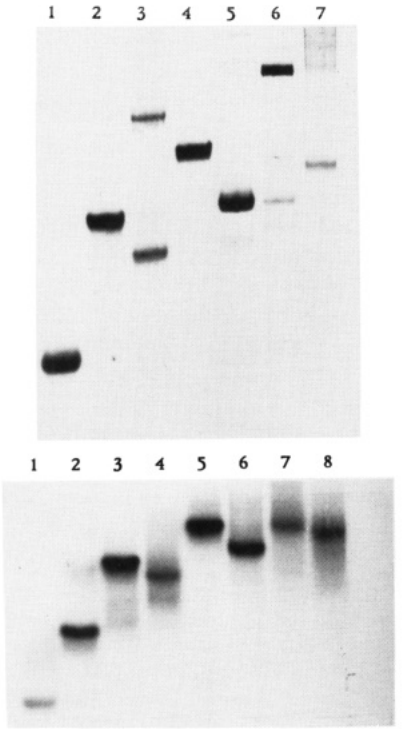

FIGURE 2: Native gels of three-strand and four-strand junctions, antijunctions, and mesojunctions. (a) Equimolar mixtures of strands at 5 μM strand concentrations, which are electrophoresed at 4 °C. This 15% polyacrylamide gel is stained with Stains-all dye. Lane 1 contains a mixture of strands Y112 and Y212. Lane 2 contains a mixture of strands Y112, Y212, and Y312, to form the 3_3 conventional branched junction. Lane 3 contains a mixture of strands Y112, Y212, and Y321, which constitute the 3_1 mesojunction. Note the presence of two bands in this lane. Lane 4 contains strands X112, X212, X312, and X412, to form the conventional 4_4 branched junction. Lane 5 contains strands X112, X212, X321, and X421, to form the $^14_2$ mesojunction. Faint bands indicating multimers are evident above the main band. Lane 6 contains strands X112, X212, X312, and X421, to form the $^24_2$ mesojunction. Note that the target complex is only a minor component in this gel and that a larger species is the major product. Lane 7 contains strands X112, X221, X312, and X421, to form the 4_0 antijunction. A large portion of the material migrates in a smear of low mobility. (b) Equimolar mixtures of strands at 100 nM strand concentrations, which are electrophoresed at 45 °C. This is an autoradiogram of a 12% polyacrylamide gel. Lane 1 contains strand Y112. Lane 2 contains the same material as lane 1 in a; likewise, lane 3 contains 3_3, lane 4 contains 3_1, lane 5 contains 4_4, lane 6 contains $^14_2$, lane 7 contains $^24_2$, and lane 8 contains 4_0. Note that in each case it is possible to obtain a single band, whose mobility is similar to that of the conventional junction with the same number of strands.

logarithms are measured to base 10.

RESULTS

Formation of the Complexes. The first question to ask about this new system is whether the structures illustrated in Figure 1 can actually form as discrete complexes. The answer to this question is affirmative, but only under appropriate conditions. In the case of branched junctions containing three to six arms, formation of stable complexes has been demonstrated on native polyacrylamide gels run under conditions optimized to stabilize nucleic acid duplexes. In particular, low temperatures and high concentrations of DNA and magnesium have been used to stabilize those molecules (e.g., Kallenbach et al., 1983; Wang et al., 1991). As illustrated in Figure 2a, application of those conditions to antijunctions and mesojunctions does not, in general, result in stable 1:1:1 or 1:1:1:1 complexes. The lane corresponding to 4_0 yields uninterpretable high molecular weight smears, and the lanes

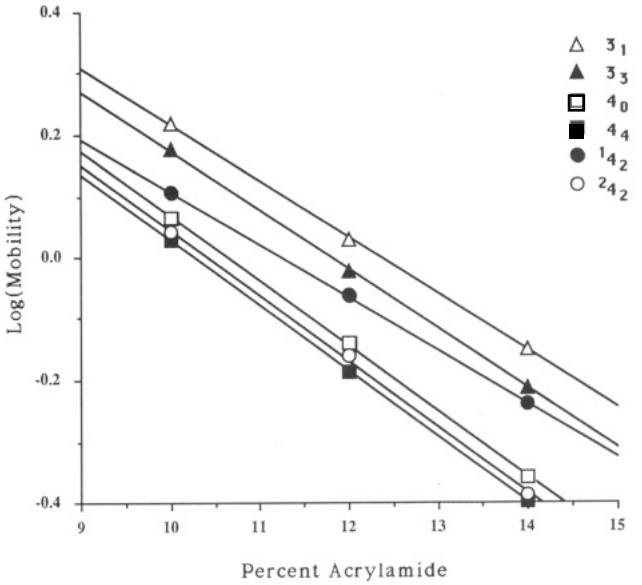

FIGURE 3: Ferguson plots of three-strand and four-strand junctions, antijunctions, and mesojunctions. Each of the species in indicated by a coded geometrial object: open triangles for the 3_1 mesojunction, filled triangles for the 3_3 conventional three-arm branched junction, open squares for the 4_0 antijunction, filled squares for the 4_4 conventional four-arm branched junction, filled circles for the $^14_2$ mesojunction, and open circles for the $^24_2$ mesojunction. The negative slopes of these plots are 0.092 and 0.097 for 3_1 and 3_3, 0.106 and 0.107 for 4_0 and 4_4, and 0.086 and 0.107 for $^14_2$ and $^24_2$. Note that $^14_2$ has a mobility similar to the four-strand complexes at the lower acrylamide concentrations used, but it migrates similarly to three-strand complexes at the higher ones.

corresponding to 3_1, $^24_2$, and, to a lesser extent, $^14_2$ contain multimers. It is necessary to use low concentrations of DNA (ca. 100 nM), and higher temperatures (45 °C), in order to block the formation of complexes containing more than a single strand of each molecular species. Figure 2b shows a gel containing each of the three-strand and four-strand complexes, including conventional branched junctions: Each is seen to run as a single band at these low concentrations and high temperatures. We assume that a single band indicates a closed discrete complex of well-defined composition. The patterns seen in Figure 2a are taken to indicate multimers or complexes with indefinite strand compositions, such as 1-2-3-4-1-2-3.... It is possible to titrate two-strand mixtures (three-strand complexes) or three-strand mixtures (four-strand complexes) with the missing strand in order to estimate the stoichiometry of the complexes formed, as done previously (Kallenbach et al., 1983; Chen et al., 1988). In each case, the stoichiometry appears to be 1:1:1 or 1:1:1:1, as appropriate (data not shown).

Ferguson Analysis. The slope of the Ferguson (1964) plot of log (mobility) as a function of polyacrylamide concentration yields information about the friction constant of the molecule (Rodbard & Chrambach, 1971). Figure 3 indicates the Ferguson plots for the molecules formed here. The slopes of 4_4, 4_0, and $^24_2$ are very similar to each other, despite their different pairing motifs. In sharp contrast to these four-strand molecules is the fourth four-strand molecule, $^14_2$, whose slope is 80% that of the other three molecules, suggesting that some of its surface area is relatively occluded. This is not unexpected from the schematic shown in Figure 1b, in which the two central helical domains could be expected to occlude one another from the medium. The two three-strand complexes differ from each other, indicating that their different connectivities result in different surface areas. Wang et al. (1991) have demonstrated that the surface area of 4_4 is smaller than expected, in comparison with three-arm, five-arm, and six-

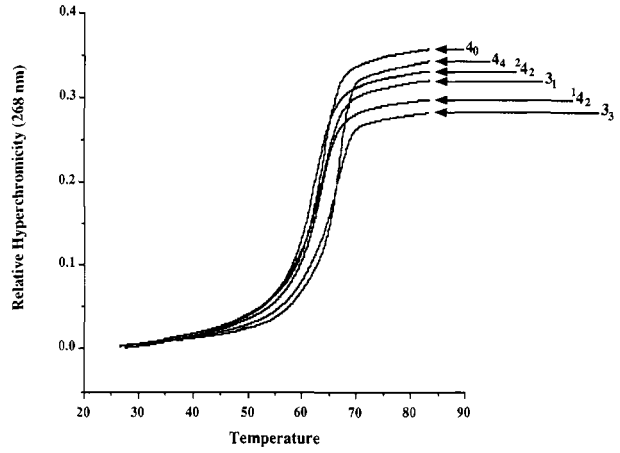

FIGURE 4: Thermal transition data for three-strand and four-strand junctions, antijunctions, and mesojunctions. The value of each point is derived by taking $OD_{268}(T)/OD_{268}(T_0) - 1]$, where $T_0 = 25$ °C. Each curve contains 601 data points, obtained at 0.1 °C intervals. The curves have been smoothed by means of a 31-point linear interpolation. Note that the conventional junctions have the highest melting points. Note that all curves have substantial premelting transitions.

arm junctions. The similarity of the slopes of 4_4, 4_0, and $^24_2$ suggests that the surfaces of these structures are occluded to a similar extent.

It is worth noting that none of the bands separates into multiple bands when run in gels containing different concentrations of acrylamide. Thus, the strands comprising each trimer or tetramer complex appear to associate in a unique fashion. Nevertheless, this finding does not rule out the possibility of a given complex equilibrating between multiple conformers during migration in the gel.

Thermal Transition Profiles. Figure 4 shows the thermal transition profiles of the six complexes. The key parts of the curves shown are above 45 °C, where we know that individual closed complexes are formed at these concentrations. The data are indicated as T_{max} (T_m), where T_{max} is the readily measured inflection point temperature and T_m is the estimated transition midpoint. The conventional four-arm junction, 4_4, is the highest, at 67.3 (66.1) °C. The least stable complex is $^24_2$, melting at 62.1 (62.0) °C, which might be predicted from inspection of Figure 1b: Stacking some of the domains of this complex implies destablization of base pairs in other domains in order to relieve strain. The other four-strand complexes are of intermediate stability, with $^14_2$ melting at 63.6 (62.5) °C and 4_0 at 64.1 (63.1) °C. The 3_1 complex is less stable than the 3_3 complex, melting at 64.2 (63.0) °C, compared with 67.1 (64.6) °C. As judged from derivative plots (data not shown), all the complexes have significant premelting transitions. Derivative plots also indicate that the cooperativity of the transition of the 4_4 complex is greater than that of any of the other complexes. Figure 4 shows clearly that the traditional junctions, 3_3 and 4_4, are more stable than the mesojunctions or the antijunction studied. Nevertheless, the similarity of all the melting temperatures indicates that all of the new complexes base pair to a significant extent.

Hydroxyl Radical Analysis. Hydroxyl radical autofootprinting has been extremely useful in providing structural information about immobile four-arm (Churchill et al., 1988), five-arm, and six-arm junctions (Wang et al., 1991), three-arm junctions (Guo et al., 1990), partially mobile four-arm junctions (Chen et al., 1988), and tethered four-arm junctions (Kimball et al., 1990). Experiments are performed by labeling a component strand of the complex and exposing it to hydroxyl radicals. Chemical attack is compared between the patterns seen when the strand is part of the complex, or when it is paired with its Watson–Crick complement in duplex DNA. Decreased susceptibility to attack at particular sites in the complex, relative to duplex, is interpreted to suggest that access to the hydroxyl radical may be limited by steric factors. Likewise, similarity to the duplex pattern at possible points of flexure is assumed to indicate helixlike structure.

Figure 5 illustrates scans of each complex in comparison with duplex DNA. We illustrate representative scans and a semiquantitative interpretation of the differences noted between the cleavage of the strand in the duplex and in the junction complex. We use a scale of three different sizes of triangles at individual sites to indicate the extent of protection. Large triangles indicate protection of 25% or greater. Intermediate triangles represent protection between 10% and 25%. Small triangles represent protection below 10%. Intermediate and strong protection are all reproducible. The weakest protection must be reproduced consistently three times before being scored. Some sites, such as position 11 on strand 2 of the $^14_2$ mesojunction (Figure 5c), show weak protection on the scan selected for illustration, but they are not reproducible, and hence are not scored.

The Four-Strand Branched Junction. Strong protection on the conventional branched junction 4_4 is seen at the two residues that flank the branch point on the crossover strands (2 and 4), as noted in previous work (Churchill et al., 1988; Kimball et al., 1990). Also noted again here are the intermediate protections four nucleotides 3′ to the branch site on the noncrossover strands (1 and 3). With 16 nucleotide pairs per arm, this junction is longer than the immobile junction used earlier, and extra protection is seen one turn 5′ to these protections on the noncrossover strands, in the two residues six nucleotides 5′ to the branch point.

The Four-Strand Antijunctions. The antijunction 4_0 shows protection at the residues that flank the bend at the middle of each strand. The protection at the strand-centers is stronger on strands 1 and 2 than on strands 3 and 4. Protection may be noted 10 nucleotides 5′ to this region, as well. Figure 6a illustrates a model consistent with these data: Rather than forming a squarelike molecule, as shown by the schematics in Figure 1, the protection upstream from the centers suggests that the structure forms two antiparallel stacking domains that occlude each other. The upstream protection data are roughly 4-fold symmetric, suggesting that neither of the two conformers in Figure 6a predominates, but that they are roughly equiprobable. The differences between the strand-center protections imply that the 2-fold symmetry of the structures shown in Figure 6a is clearly an idealization, and that subtle structural or dynamic differences destroy that symmetry. These data suggest that a quadrilateral built earlier (Chen et al., 1989) may assume the same type of structure.

The Four-Strand Mesojunction $^14_2$. This structure forms most readily of all the unusual structures described here. Whereas the rule in all the multihelical structures examined to date is that stacking is maximized where possible, it is most useful to interpret the protection data in terms of stacking models. It is not possible to form a stacking structure from this molecule in which the two helical end (radial) domains stack on the two central domains, because there are 1.5 helical turns in the central (circumferential) domains; had those domains been a single helical turn long, such a structure would have been possible.

This leaves two further possible multiple–stacking structures, in which both radial helices stack on one circumferential

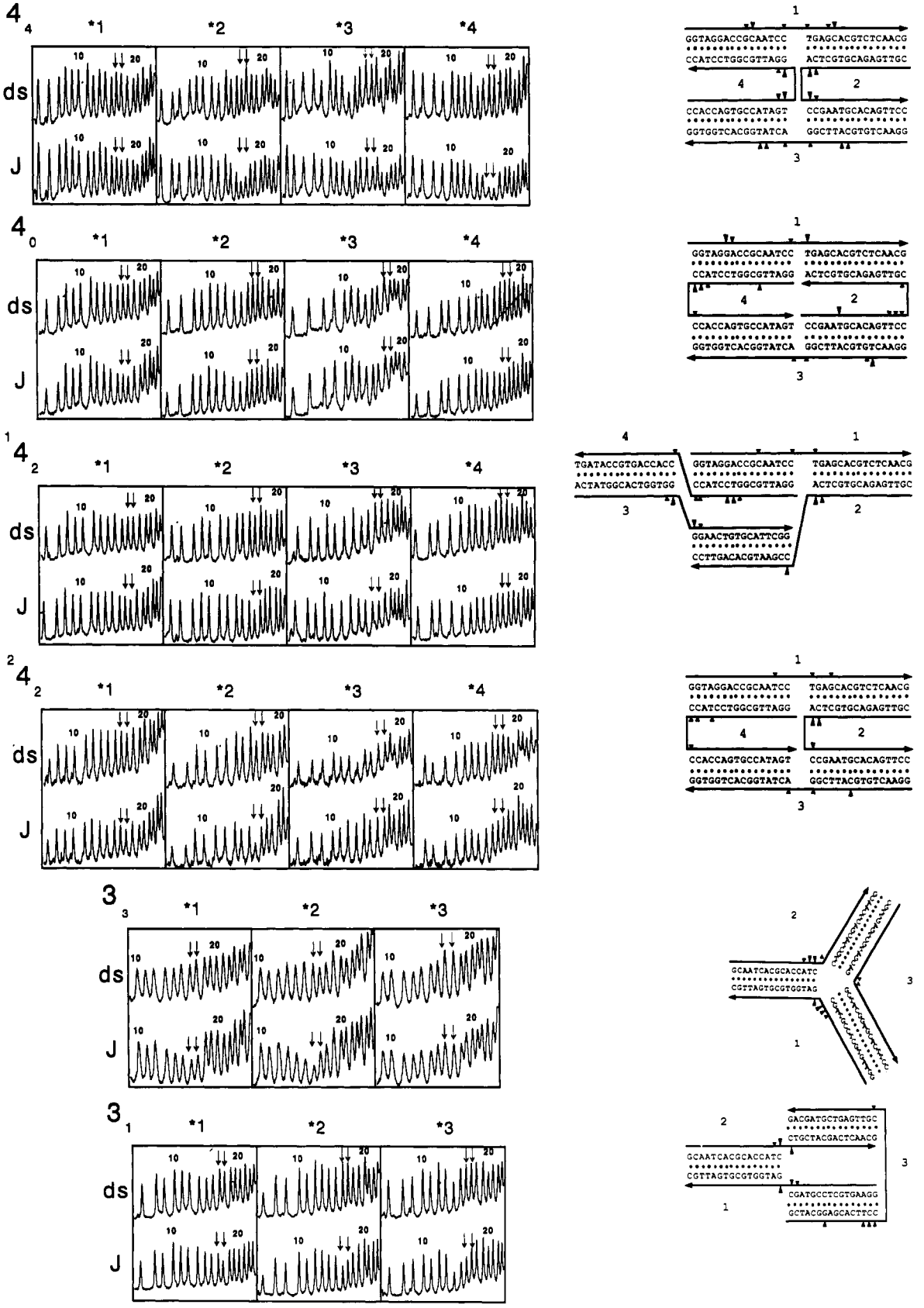

FIGURE 5: Hydroxyl radical cleavage patterns for three-strand and four-strand junctions, antijunctions, and mesojunctions. Each panel of this figure contains two parts, a direct comparison of the hydroxyl radical sensitivity of an individual strand in its complex (J) and the same strand paired to its Watson–Crick complement (ds). The central two nucleotides of each strand (16 and 17) are indicated by two arrows. Protections at individual sites are indicated in the schematic diagram below the comparison by triangular indicators, with the extent of protection indicated qualitatively by the size of the triangle: The largest triangles correspond to protection greater than 25%, relative to duplex, intermediate triangles are protection between 10% and 25%, and small triangles are protection less that 10% relative to duplex. Protection must be triply reproduced before being scored. (a) The 4_4 junction. The features seen here are similar to those noted previously for this type of complex (Churchill et al., 1988). (b) The 4_0 antijunction. (c) The $^1 4_2$ mesojunction. (d) The $^2 4_2$ mesojunction. Position 20 of strand 4 of this complex shows an enhancement, in addition to the protections noted. (e) The 3_3 junction. (f) The 3_1 mesojunctions.

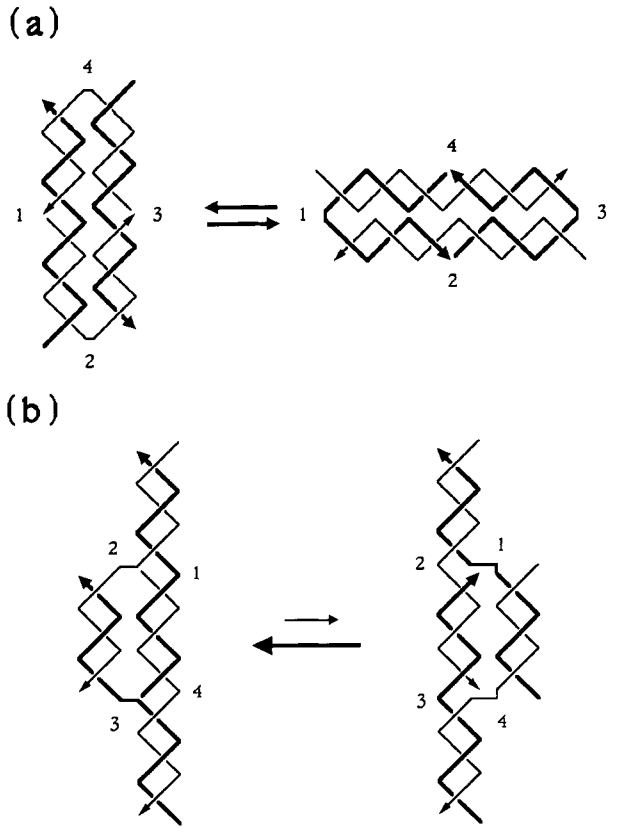

FIGURE 6: Structural models for the 4_0 antijunction and the $^{1}4_2$ mesojunction. (a) The 4_0 antijunction. Strands are numbered near their middles; strands 1 and 3 have been drawn with broader pen than strands 2 and 4. The 3′ ends of strands are indicated by the arrowheads. The hydroxyl radical protection data suggest that the molecule forms two helical domains, but that there is little bias between the two possible arrangements shown here. In the arrangement on the left, strands 1 and 3 are helical, and strands 2 and 4 bridge between domains; in the arrangement on the right, strands 2 and 4 are helical, and strands 1 and 3 bridge between domains. Both of these models have an 2-fold symmetry axis at the center of each figure perpendicular to the plane of the page; however, because protection at the strand centers differs between the two ends of each complex, we have not indicated this symmetry on the drawing. (b) The $^{1}4_2$ mesojunction. The drawing conventions are the same as in part a. The two models shown here each stack three helices to form one domain, while the fourth helix is isolated. The two structures are inequivalent; the structure on the left has two 3′ ends in its isolated helix, whereas the structure on the right contains two 5′ ends in its isolated helix. There is a 2-fold symmetry axis in each structure that is horizontal and in the plane of the page. This results in the spacing of the isolated helix on the left containing two major groove steps and one minor groove step (ca. 16 nucleotide pairs0, but the isolated helix on the right contains two minor groove steps and one major groove step (ca. 14 nucleotide pairs). The protection data suggest that the model on the left is favored, as indicated by the larger arrow in that direction. This may be due to the fact that the isolated helix on the right is strained unless it is restricted to about 14 nucleotide pairs. The inappropriate length of the isolated helix on the right is indicated by the small right-angle bends at the centers of strands 1 and 4. It is possible by sequence symmetry for strand 2 to cross strand 4 in a two-step branch migratory isomerization, thereby shortening the isolated helix, but this does not seem to be a dominant structure here.

helix, to form a triple stack and the other radial helix remains unstacked entirely (Figure 6b). These two structures are inequivalent: In one case the free ends of the unstacked domain are 3′ ends, and in the other case, they are 5′ ends; consequently, the crossover strands enter the unstacked domain in the 5′→3′ or 3′→5′ directions, respectively. The first of these structures is favored by a domain containing enough nucleotide pairs to accomodate two major grooves and a minor groove (ca. 16 nucleotide pairs), and the second is favored by a domain containing two minor grooves and a major groove (ca. 14 nucleotide pairs). Although model building with physical models (Seeman, 1988) suggests that the second structure is not nearly so easily achieved for the molecule in question, it is important to realize that the means of sequence assignment has resulted in a potential ambiguity of pairing at the crossover point near the center of strand 4. This ambiguity results in the possibility of a branch-migratory-like isomerization, which in turn could create a double crossover between strands 4 and 2, and which shortens the unique domain to 14 nucleotide pairs.

The hydroxyl radical protection data are most consistent with the first of these structures. The major loci of protection are at the centers of strands 2 and 3, the crossover points in the model. This protection is not as strong as that seen at the center of the conventional branched junction, 4_4. Additional, protection may be noted about six nucleotides 5′ to the crossover points on strands 1 and 4 (Figure 6b). This upstream protection on strands 1 and 4 is consistent with the predictions of the first model, resulting from model-predicted interhelical contacts between the isolated domain and the central helix of the other domain. Additional weak protection is seen at the centers of strands 1 and 4. This additional protection is inconsistent with the model presented, and suggests the possible presence of small amounts of the other triple-stack model, or of other models with less overall stacking.

The Four-Strand Mesojunction $^{2}4_2$. The connectivity of this molecule resembles an antiparallel tethered junction (Kimball et al., 1990) with an extremely short tether. The cleavage pattern supports this view of the molecule. The protection on strand 2 is much like that on the conventional junction 4_4, and the weak protection noted on strand 3, seven nucleotides 5′ to the middle, are similar in both molecules. Thus, these two strands may resemble their structures in 4_4. The sequence of strand 3 is the strongest stacking sequence encountered in previous studies—the GpA sequence forms the helical portion of one of the two helical domains of the four-arm J1 junction (Churchill et al., 1988), and it forms the helical portion of the only helical domain in the six-arm junction J6YLG (Wang et al., 1991). The tether formed by the middle internucleotide linkage of strand 4 is much too short to bridge between two intact helices a full 1.5 turns long. Strand 1 shows nonspecific weak protection in the vicinity of the middle, and strand 4 shows some weak protection at the middle and four nucleotides 5′ to the middle. This protection is not readily interpreted. The antiparallel nature of the helical arms in which strand 2 participates and the probable continuous helix in the vicinity of the middle of strand 3 are the only concrete features that may be derived from these data.

The Three-Strand Junction. Hydroxyl radical autofootprinting has been used previously to characterize a three-strand junction containing eight nucleotide pairs per arm (Guo et al., 1990). Those investigators have found that two helical arms appear to form a single stacking domain. The sequence of the junction studied here is different from the one studied by Guo et al., but the results are similar. Strands 1 and 2 show strong protection in the vicinity of the branch point, whereas strand 3 shows weak protection in that region. These data are consistent with a model in which one strand is not very different from helical, and two strands contain sharp bends.

The Three-Strand Mesojunction 3. This molecule corresponds to a replication fork that has been tethered at the ends of the daughter strands furthest from the branch point. Protection may be noted on strands 1 and 2 near their middle regions. These are weaker than the protection noted for the

same strands in the 3_3 structure. The third strand has protection at its middle, and also shows protection 10 nucleotides 5′ to the middle, suggesting that it lies near the helix formed by strands 2 and 3. No stacking domains appear to be formed from two helices in this structure.

DISCUSSION

Complex Formation. The results indicate that the formation of the new structures is much more problematic than formation of branched junctions. Whereas branched junctions of appropriate stoichiometry have always been seen to form readily when their arms are long enough (Kallenbach et al., 1983; Wang et al., 1991), these molecules are less tractable. One must ensure that the temperature is *high* enough and the concentration is *low* enough to avoid indefinite polymerization of the constituent molecules of the antijunction 4_0. It is worth noting that the essential features of the 4_0 molecule have been constructed before, as the edges of a DNA quadrilateral whose vertices are separated by 1.5 turns of DNA (Chen et al., 1989); the four strands there are held together pairwise in the quadrilateral by exocyclic arms, thereby converting the complex into two cyclic strands that are hextuply linked. Formation of the quadrilateral from four constituent junctions is straightforward, and no octalateral is seen.

The two four-strand mesojunctions differ substantially in their abilities to form closed complexes. $^{2}4_2$ is extremely difficult to assemble, whereas $^{1}4_2$ forms readily. Model building (Seeman, 1988) suggests that $^{2}4_2$ built with 1.5 turns of DNA per helix is very strained, whereas $^{1}4_2$ is relatively unstrained; a single turn of DNA in the circumferential helices of $^{2}4_2$ does not appear to result in a strained molecule. The 3_1 complex is also difficult to form, but model building suggests that single-turn circumferential helices would lessen the strain in that case, as well. Multimeric structures can also lessen strain caused by twist, and the dimers of $^{2}4_2$ and 3_1 are seen prominently on gels at high concentration. The successful assemblies of these new complexes shown in Figure 2b have been achieved by selecting conditions unfavorable for multimeric structures (low concentration and high temperature).

Hydroxyl Radical Footprinting. In general, the protection noted here for the new structures is weaker than that seen at the junctions in the analysis of conventional branched junctions (Churchill et al., 1988). However, weak protection, similar to that seen here, is consistently noted in those structures at a site four nucleotides 3′ to the branch point; when the motion of the structure is limited, by tethering the junctions to assume a fixed conformer (Kimball et al., 1990), protection is increased at those loci. It is possible that more tightly constrained versions of these assemblies, such as a completely immobile $^{1}4_2$ complex, will yield more conclusive data. It is possible that weak protection seen in some cases results from some distortion of the DNA structure (Tullius, 1987), rather than from occlusion of the site: Nevertheless the dominant sites of protection are consistent with occlusion at the sites expected in the models presented for 4_0 and $^{1}4_2$; model building using physical models (Seeman, 1988) and the protection data are in excellent agreement. Further structural characterization will clarify the validity of the assumptions applied to produce the preliminary structural models presented here.

Stacking. The dominant motif inferred from the assembly experiments and from the hydroxyl radical studies is that individual double helices frequently stack to form longer double helical domains. This is similar to the results seen earlier in studies of four-arm Holliday analogs (Churchill et al., 1988), three-arm branched junctions (Guo et al., 1990), and six-arm junctions (Wang et al., 1991). In particular, the $^{1}4_2$ mesojunction and the 4_0 antijunction seem to contain stacking structures. If one were to synthesize the $^{1}4_2$ mesojunction with single-turn circumferential junctions, it would be possible to form two two-helix stacked domains, rather than the three-helix stack apparently found. The stacking seen in the 4_0 antijunction suggests that the quadrilateral formed earlier may indeed be rhombus-shaped, rather than squarelike, with a short minor diagonal.

Components of Biological Structures. The surroundings of the strand centers at the base of the radial helices in both $^{1}4_2$ and 3_1 correspond to replication forks. In fact, they may be a better model for some features of that structure than the conventional three-arm branched junction, because one end of each daughter helix is free. The results obtained from hydroxyl radical attack on $^{1}4_2$ suggest that the parental helix and one daughter helix may tend to stack in a replication fork. Inasmuch as the two daughter sequences are the same, sequence stability cannot simply determine which stack (parental with leading daughter helix or with lagging daughter helix) will form, as it apparently does in the case of four-arm branched junctions (Chen et al., 1988; Duckett et al., 1988). Nevertheless, the continuous strands are different, so a preference cannot be ruled out.

The 3-helix:1-helix structure derived from hydroxyl radical studies for the $^{1}4_2$ mesojunction (Figure 6b) bears a resemblance to the double-crossover Holliday junction (Holliday, 1964); however, it contains only one, rather than two, strands crossing over between the two helical domains at each end of the unique helix. In addition, this molecule is capable of branch migration, which would produce double-strand crossovers, given the appropriate sequence and helix lengths on the circumferential helices. That such crossover structures are not seen here suggests that their formation requires greater torsional stress than exists here. Whereas this is a molecule that resembles intermediates postulated in recombination structures, its physical properties bear further examination.

Nanotechnological Structures. The 4_0 antijunction is the central structure formed by the assembly of four three-arm junctions to generate a quadrilateral. A lattice of four-arm 4_4 junctions (e.g., Seeman, 1985a,b) will consist of alternating 4_4 junctions as vertices and 4_0 antijunctions at the centers of the quadrilaterals formed. It is also possible to model 2-D lattices built with either of the four-strand mesojunctions as components. The antijunctions presented here have been derived from helices flanking a polygon. Those with 6 or more helices can also surround a polyhedron. The edges of a DNA molecule with the connectivity of a cube (Chen & Seeman, 1991) correspond to a 12 helix antijunction different from those described here. The roles of such structures in nanotechnological applications has been treated previously (Seeman, 1991). The role of antijunctions and mesojunctions in the construction of intricate single-stranded nucleic acid knots is described elsewhere (Seeman, 1992).

ACKNOWLEDGMENT

We would like to thank Dr. Richard D. Sheardy for valuable discussions and assistance with the measurement of thermal denaturation profiles on his spectrophotometer (supported by NSF grant DMB-8996232 to him), Dr. Albert S. Benight for valuable discussions of the thermal data, and Dr. Andrew T.-J. Fu for insightful discussions clarifying the nature of double crossover structures. Discussions with Dr. Thomas Tullius on the hydroxyl radical footprinting data have been extremely useful for our interpretations. We are grateful to

Dr. Nicholas Cozzarelli for providing us with a copy of the program Knotter by John Jenkins. The support of Biomolecular Imaging on the NYU campus by the W. M. Keck Foundation is gratefully acknowledged.

REFERENCES

Caruthers, M. H. (1982) in *Chemical and Enzymatic Synthesis of Gene Fragments* (Gassen, H. G., & Lang, A., Eds.) pp 71–79, Verlag Chemie, Weinheim.

Chen, J., & Seeman, N. C. (1991) *Nature 350*, 631–633.

Chen, J.-H., Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R. & Seeman, N. C. (1988) *Biochemistry 27*, 6032–6038.

Chen, J.-H., Kallenbach, N. R. & Seeman, N. C. (1989) *J. Am. Chem. Soc. 111*, 6402–6407.

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

Constantine, G. M. (1987) *Combinatorial Theory and Statistical Design*, p 235, John Wiley & Sons, New York.

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. 86*, 7336–7340.

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

Ferguson, K. A. (1964) *Metabolism 13*, 985–1002.

Guo, Q., Seeman, N. C. & Kallenbach, N. R. (1989) *Biochemistry 28*, 2355–2359.

Guo, Q., Lu, M., Churchill, M. E. A., Tullius, T. D. & Kallenbach, N. R. (1990) *Biochemistry 29*, 10927–10934.

Holliday, R. (1964) *Genet. Res. 5*, 282–304.

Kallenbach, N. R., Ma, R.-I. & Seeman, N. C. (1983) *Nature (London) 305*, 829–831.

Kimball, A., Guo, Q., Lu, M., Cunningham, R. P., Kallenbach, N. R., Seeman, N. C. & Tullius, T. D. (1990) *J. Biol. Chem. 265*, 6544–6547.

Lu, M., Guo, Q., Seeman, N. C. & Kallenbach, N. R. (1991) *J. Mol. Biol. 221*, 1419–1432.

Lu, M., Guo, Q., Mark, L. A., Seeman, N. C. & Kallenbach, N. R. (1992) *J. Mol. Biol. 223*, 781–789.

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

Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.

Mueller, J. E., Kemper, B., Cunningham, R. P., Kallenbach, N. R. & Seeman, N. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85* 9441–9445.

Mueller, J. E., Newton, C. J., Jensch, F., Kemper, B., Cunningham, R. P., Kallenbach, N. R. & Seeman, N. C. (1990) *J. Biol. Chem. 265*, 13918–13924.

Mueller, J. E., Du, S. M. & Seeman, N. C. (1991) *J. Am. Chem. Soc. 113*, 6306–6308.

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

Rodbard, D. & Chrambach, A. (1971) *Anal. Biochem. 40*, 95–134.

Seeman, N. C. (1982) *J. Theor. Biol. 99*, 237–247.

Seeman, N. C. (1985a) *J. Biomol. Struct. Dyn. 3*, 11–34.

Seeman, N. C. (1985b) *J. Mol. Graphics 3*, 34–39.

Seeman, N. C. (1988) *J. Biomol. Struct. Dyn. 5*, 997–1004.

Seeman, N. C. (1990) *J. Biomol. Struct. Dyn. 8*, 573–581.

Seeman, N. C. (1991) *DNA Cell Biol. 10*, 475–486.

Seeman, N. C. (1992) *Mol. Eng.* (in press).

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 *Progress in Clinical & Biological Research 172A: The Molecular Basis of Cancer* (Rein, R., Ed.) pp 99–108, Alan Liss Inc., New York.

Tullius, T. D. (1987) *Trends Biochem. Sci. 12*, 297–300.

Tullius, T. D. & Dombroski, B. (1985) *Science 230*, 679–681.

Tullius, T. D. & Drombroski, B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469–5473.

Wang, Y., Mueller, J. E., Kemper, B. & Seeman, N. C. (1991) *Biochemistry 30*, 5667–5674.

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