Assembly and Characterization of Five-Arm and Six-Arm DNA Branched Junctions†

Yinli Wang,‡ John E. Mueller,‡,§ Börries Kemper,‖ and Nadrian C. Seeman*,‡

Department of Chemistry, New York University, New York, New York 10003, and Institute of Genetics, University of Köln, 5000 Köln 41, Germany

Received December 26, 1990; Revised Manuscript Received March 15, 1991

ABSTRACT: DNA branched junctions have been constructed that contain either five arms or six arms surrounding a branch point. These junctions are not as stable as junctions containing three or four arms; unlike the smaller junctions, they cannot be shown to migrate as a single band on native gels when each of their arms contains eight nucleotide pairs. However, they can be stabilized if their arms contain 16 nucleotide pairs. Ferguson analysis of these junctions in combination with three-arm and four-arm junctions indicates a linear increase in friction constant as the number of arms increases, with the four-arm junction migrating anomalously. The five-arm junction does not appear to have any unusual stacking structure, and all strands show similar responses to hydroxyl radical autofootprinting analysis. By contrast, one strand of the six-arm junction shows virtually no protection from hydroxyl radicals, suggesting that it is the helical strand of a preferred stacking domain. Both junctions are susceptible to digestion by T4 endonuclease VII, which resolves Holliday junctions. However, the putative helical strand of the six-arm junction shows markedly reduced cleavage, supporting the notion that its structure is largely found in a helical conformation. Branched DNA molecules can be assembled into structures whose helix axes form multiply connected objects and networks. The ability to construct five-arm and six-arm junctions vastly increases the number of structures and networks that can be built from branched DNA components. Icosahedral deltahedra and 11 networks with 432 symmetry, constructed from Platonic and Archimedean solids, are among the structures whose construction is feasible, now that these junctions can be made.

Branched DNA junctions containing three or four arms around a single point are seen as ephemeral structures in the biological processes of replication and recombination [e.g., Kornberg (1980) and Kucherlapati and Smith (1988)]. For several years, these unstable structures, such as four-arm Holliday (1964) junctions, have been modeled by immobile junctions (Seeman, 1982; Kallenbach et al. 1983a), which lack their homologous sequence symmetry. Immobile junctions can also serve as the basis for the construction of objects on the nanometer scale (Seeman, 1982, 1985a,b), and it has been suggested that they can also be combined to form networks (Seeman, 1982) and new biomaterials (Seeman, 1985a; Robinson & Seeman, 1987; Chen et al., 1989). Assembly of complex figures is done by terminating the arms of individual junctions in cohesive ends, so that they can self-assemble and then be ligated together. Thus, the same procedures used with linear DNA in genetic engineering, to form linear or circular molecules (Cohen et al. 1973), can be used with branched molecules to form multiply connected objects (Seeman, 1982). The edges of these objects are composed of double-helical DNA, while the vertices correspond to the branch points of the junctions from which they are assembled. A key concept here is *connectedness*, which is the number of edges of a closed object or a lattice that meet at a vertex; the number of arms of the junction at a given vertex limits the maximum connectedness of that vertex. In the last several years, we have formed a series of two-connected macrocycles with three-arm DNA junctions (Ma et al., 1986) and with four-arm DNA junctions (Petrillo et al., 1988); in addition, we have constructed a specific quadrilateral from three-arm junctions (Chen et al., 1989). Recently, we have built a three-connected, 3-D object from three-arm junctions (Chen & Seeman, 1991). The helix axes of this object have the [4,3] connectivity (Wells, 1977) of a cube, or a rhombohedron.

Useful new materials from branched DNA junctions are likely to involve N-connected networks (Seeman, 1982), as well as individual closed polyhedra. Although 3-D networks can be formed from three-connected or four-connected objects (Wells, 1977), many networks involve higher connectivities (Williams, 1979). For example, the two-dimensional packing of equilateral triangles entails a six-connected network; similarly, each of the vertices of the simple cubic stick lattice, formed from three mutually perpendicular intersecting lines, is six-connected. There are a number of space-filling regular arrays that are five-connected, the simplest of which is formed by stacking hexagonal prisms in parallel columns (Williams, 1979). In addition, the icosahedron is a major polyhedral structural element found both in viruses (Caspar & Klug, 1962) and in geodesic domes [e.g., Edmondson (1987)]; the icosahedral deltahedron reflecting the simplest ($T = 1$) surface structure is five-connected, and the more complicated deltahedra ($T > 1$) are five- and six-connected.

If five-connected and six-connected materials are to be constructed from branched nucleic acid building blocks, it is necessary to explore the feasibility of building junctions with five or six arms. We have done this here, and we report the first assembly of five-arm and six-arm junctions. We find that these junctions are somewhat less stable than three-arm and four-arm junctions. Longer double-helical arms are needed

† This research has been supported by Grant N00014-89-J-3078 from the Office of Naval Research and by Grant GM-29554 from the NIH.

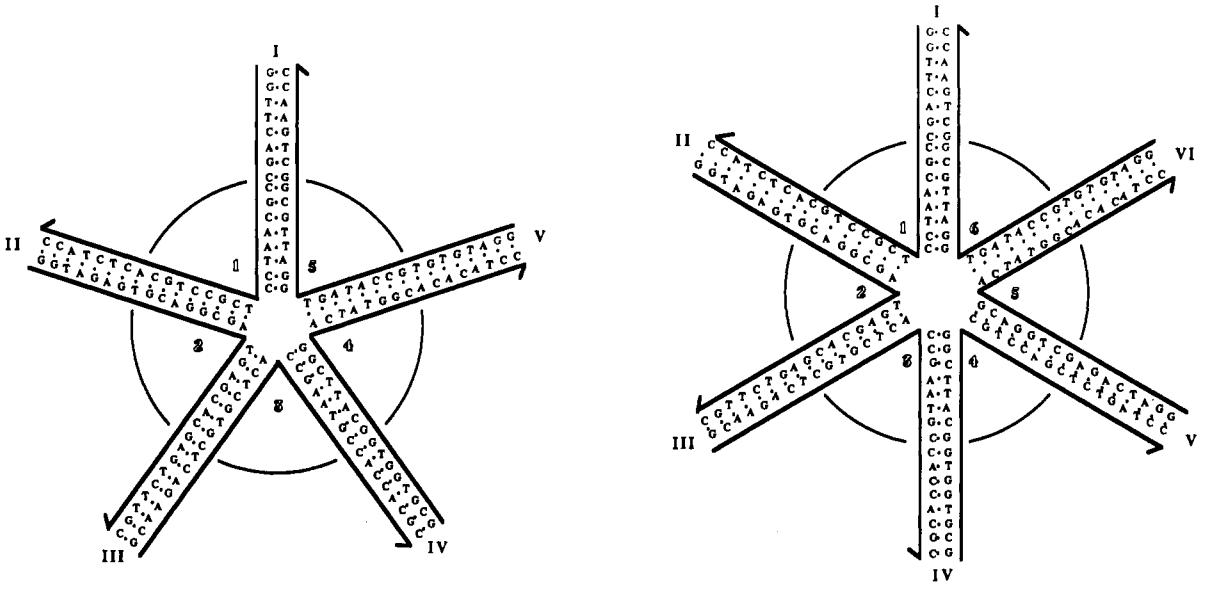
* To whom correspondence should be addressed.

‡ New York University.

§ Permanent address: Department of Biological Sciences, SUNY/Albany, Albany, NY 12222.

‖ University of Köln.

0006-2960/91/0430-5667$02.50/0



JYL5G (J5M8) JYL6G (J6M8)

FIGURE 1: Sequences of the five-arm and six-arm junctions constructed. Each of the larger junctions is shown with its complete sequence; J5YLG contains five 32-mer strands, and J6YLG contains six 32-mer strands. The smaller junctions, J5M8 and J6M8, are contained within the circles that bound the eight nucleotide pairs of these junctions that are closest to the branch points. The 3′ end of each strand is indicated by a half-arrowhead. The strand numbers are indicated by Arabic numerals, and the arm numbers are denoted by Roman numerals. Note that arms II and III of the five-arm junctions have the same sequence flanking the junction, but branch migration can only occur in the junction-flanking dinucleotide of strand 2 pairs with itself, which is extremely unlikely. This circumstance applies to arms II and III of the six-arm junctions, and it also applies to arms IV and V of the six-arm junctions. Note that strands 1, 2, 3, and 5 of J5YLG are identical with strands 1, 2, 3, and 6 of J6YLG, respectively.

to stabilize these structures than are needed to stabilize three-arm and four-arm junctions. Nevertheless, they can be assembled, and we expect that they can be used as building blocks for five-connected and six-connected objects and networks.

There is another motivation for investigating junctions with a greater number of arms flanking the branch point: A surprising feature of four-arm Holliday junctions is that they arrange their helical arms into stacked helical domains; one pair of domains appears to dominate over the other in the immobile junctions studied to date (Cooper & Hagerman, 1987, 1989; Churchill et al., 1988; Duckett et al., 1988; Murchie et al., 1989). One of the variables associated with stacking dominance is the number of arms in the junction. It is of interest to see if stacking dominance patterns are retained as the number of arms is increased from four to five or six. Accordingly, we have analyzed the stacking properties of five-arm and six-arm junctions by hydroxyl radical autofootprinting analysis (Tullius & Dombroski, 1985, 1986; Churchill et al., 1988) and by endonuclease VII susceptibility [e.g., Mueller et al. (1989)]. These experiments reveal that the five-arm junction has no dominant stacking domains but that the six-arm junction contains a single preferred domain.

Materials and Methods

Sequence Design. Branch migration may be eliminated in four-arm junctions by using all four base pairs (A·T, T·A, G·C, and C·G) in the four positions that flank the junction (Seeman, 1982). For junctions with more arms, one must assume that two identical adjacent base pairs will not undergo the branched migratory pairing–isomerization reaction (Seeman, 1982; Seeman & Kallenbach, 1988). The assumption is warranted, because branch migratory pairing–isomerization requires two successive complementary residues on one strand to pair. This pairing is not likely to occur on physical grounds, because the two glycosyl bonds cannot get far enough apart. We have used this approach for the design of five-arm and six-arm junctions. The sequences have been assigned by the program SEQUIN (Seeman, 1990). The sequences of the junctions constructed in this study are shown in Figure 1. Eight nucleotide pairs nearest the junction on four of the arms are derived from J1 (Seeman & Kallenbach, 1983; Kallenbach et al., 1983a), an extensively studied four-arm junction. These are arms I and III–V of the five-arm junctions and arms I, III, IV, and VI of the six-arm junction. We have interrupted strand 1 of J1 to make the five-arm junctions and strands 1 and 3 of J1 to make the six-arm junctions; these are the strands that appear to form continuous helices in J1 (Churchill et al., 1988), but retention of the junction-flanking bases is expected to preserve stacking preferences (Chen et al., 1988; Duckett et al., 1988).

Two other junctions are used as electrophoretic standards in this study, a three-arm junction, JYG, and a four-arm junction, JXG. The sequences of both these junctions have been designed by SEQUIN; JXG is an extension of J1. The

Chart I

JYG sequence	
strand 1:	G-G-A-A-G-T-G-C-T-C-C-G-T-A-G-C-G-A-T-G-G-T-G-C-G-T-G-A-T-T-G-C
strand 2:	G-C-A-A-T-C-A-C-G-C-A-C-C-A-T-C-C-T-G-C-T-A-C-G-A-C-T-C-A-A-C-G
strand 3:	C-G-T-T-G-A-G-T-C-G-T-A-G-C-A-G-G-C-T-A-C-G-G-A-G-C-A-C-T-T-C-C
JXG sequence	
strand 1:	G-G-T-A-G-G-A-C-C-G-C-A-A-T-C-C-T-G-A-G-C-A-C-G-T-C-T-C-A-A-C-G
strand 2:	C-G-T-T-G-A-G-A-C-G-T-G-C-T-C-A-C-C-G-A-A-T-G-C-A-C-A-G-T-T-C-C
strand 3:	G-G-A-A-C-T-G-T-G-C-A-T-T-C-G-G-A-C-T-A-T-G-G-C-A-C-T-G-G-T-G-G
strand 4:	C-C-A-C-C-A-G-T-G-C-C-A-T-A-G-T-G-G-A-T-T-G-C-G-G-T-C-C-T-A-C-C

sequences of JYG and JXG are shown in Chart I.

Synthesis and Purification of DNA. All DNA molecules used in this study are 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 with a gradient of NaCl in a solvent system containing 20% acetonitrile and 80% 0.02 M sodium phosphate. Fractions from the major peak are collected, concentrated, desalted, and evaporated. This method of purification separates 16-mers from 15-mers, but it is probably not adequate to separate 32-mers from 31-mers. Nevertheless, not much of this contaminant is made (<2% as estimated by trityl analysis), and we detect no impurity in any of our experiments.

Formation of Hybrid Molecules. Junctions 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. Errors in stoichiometry are estimated by native gel electrophoresis of adjacent dimers. A single band on a native gel is taken to indicate a homogeneous stoichiometric complex.

Hydroxyl Radical Analysis. Individual strands of the junctions are radioactively labeled and are additionally gel purified from a 10% denaturing polyacrylamide gel. Each of the labeled strands [approximately 10 pmol in 50 mM Tris·HCl (pH 7.5) containing 10 mM $MgCl_2$] is annealed to an excess of the unlabeled complementary strand, it is annealed to an excess of a mixture of the other four or five strands, it is left untreated as a control, or it is treated with sequencing reagents (Maxam & Gilbert, 1977) for a sizing ladder. The samples are annealed by heating to 90 °C for 3 min and then slowly cooled to 4 °C. Hydroxyl radical cleavage of the double-strand and junction samples for all strands takes place at 4 °C 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 10% polyacrylamide/8.3 M urea sequencing gel. Autoradiograms are scanned with a Hoefer GS300 densitometer in transmission mode.

Enzymatic Reactions. (*A*) *Kinase Labeling.* One microgram of an individual strand of DNA is dissolved in 10 μL of a solution containing 55 mM Tris·HCl, pH 7.6, 1 mM spermidine, 10.0 mM $MgCl_2$, 15 mM dithiothreitol (DTT), and 0.2 mg/mL nuclease-free bovine serum albumin (BSA) (BRL) and mixed with 1–2 μL of 1.25 mM [γ-^{32}P]ATP (10 mCi/mL) and 1–2 units of polynucleotide kinase (Boehringer) for 45 min–2 h at 37 °C. The reaction is stopped by desiccation or ethanol precipitation of DNA.

(*B*) *Endonuclease VII Digestions.* Endonuclease VII is prepared as described previously (Kemper & Garabett, 1981). Ten picomoles of five-arm or six-arm junction is dissolved in 15 μL of a solution containing 40 mM Tris–acetate (pH 8.1), 20 mM sodium acetate, 2 mM EDTA, and 12.5 mM $MgCl_2$ (TAEMg). The solution is brought to 90 °C for 3 min and allowed to cool slowly. Samples are then brought to a volume of 30 μL in a solution containing 40 mM Tris·HCl (pH 8.1), 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol, and bovine serum albumin (100 μg/mL). Fifteen microliters of this solution is treated at 4 °C with 1000 units of endonuclease VII. Reactions are stopped after 4 h by dessication.

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/bis(acrylamide)]. The running buffer consists of 100 mM Tris·HCl, pH 8.3, 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 tacking dye. Gels are run on an IBI Model STS 45 electrophoresis unit at 70 W (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 10 or 20% acrylamide [19:1 acrylamide/bis(acrylamide)]. Three to six micrograms of DNA is suspended in 10–25 μL of a solution of TAEMg. The solution is boiled and allowed to cool slowly to room temperature. Samples are then brought to a final volume of 11–27 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 10 V/cm overnight in a 4 °C cold room and exposed to X-ray film for up to 15 h or stained with Stainsall dye. Mobilities are measured relative to the Xylene Cyanol FF tracking dye.

Results

Assembly of Junctions. We have shown previously that it is possible to assemble four-arm and three-arm junctions with double-helical arms containing eight (Kallenbach et al., 1983a; Ma et al., 1986) or even six nucleotide pairs (Kallenbach et al., 1983b); Guo et al. (1990) have recently characterized a three-arm junction containing eight nucleotide pairs per arm. Thus, as a first try, it is not unreasonable to use eight nucleotide pairs per arm to stabilize five-arm and six-arm junctions. Accordingly, junctions J5M8 and J6M8 (Figure 1) have been synthesized, and assembly has been attempted. The sequence of J5M8 has been selected by inserting a new double-helical arm in the middle of the first strand of the well-characterized four-arm junction, J1 (Kallenbach et al., 1983a; Seeman et al., 1985; Wemmer et al., 1985; Marky et al., 1987; Copper & Hagerman, 1987, 1989; Churchill et al., 1988, Seeman et al., 1989); J6M8 contains a further arm inserted into the third strand of J1.

It is possible to mix the contiguous pairs of J5M8 and J6M8 with perfect 1:1 stoichiometry (data not shown). Figure 2 shows the result of combining the individual strands in equimolar mixtures containing from one to five or six strands. The higher mobility products seen in the junction lanes indicate that the junctions break down on the gel, unlike junctions containing fewer arms. To demonstrate that the gel illustrates dissociation, rather than imperfect stoichiometry, the junction bands have been excised from gels, reequilibrated, and then reelectrophoresed. The dissociation persists, indicating the fundamental instability of the complex (data not shown).

In response to this result, we have doubled the lengths of all the arms and have synthesized the junctions J5YLG and J6YLG (Figure 1). Figure 3 contains the same information about equimolar mixtures for these junctions as does Figure 2 for the shorter junctions. In this case, there is only a single band seen in the lanes containing all the strands. We take this finding to indicate that these junctions cohere as a single species under the conditions of electrophoresis. Thus, doubling the lengths of the arms, from 8 to 16, is sufficient to stabilize complexes containing five-arm and six-arm junctions.

Ferguson Analysis. The two junctions reported here are the third and fourth species of immobile junctions that have been described. It is worthwhile to compare the electrophoretic behavior of all these types of junctions: We have determined the Ferguson (1964) plots of these species by comparing the mobilities of junctions containing three through six arms, as

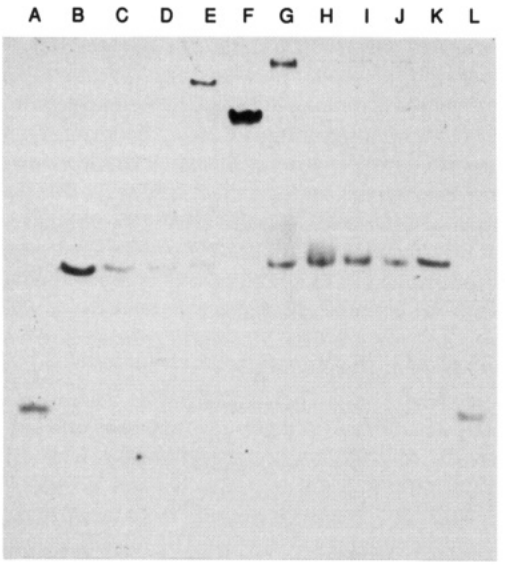


FIGURE 2: Electrophoresis of equimolar ratios of J5M8 and J6M8. The gel shown here is a 20% native gel stained with Stainsall dye. Lane F contains 4 μg of the well-characterized junction J1 (Kallenbach et al., 1983a), containing 16 nucleotides per strand, which is used as a marker here. Lanes A and L contain 4 μg of strand J5M8.3, lane B contains 2 μg of J5M8.4 + J5M8.5, and lane K contains 2 μg each of J6M8.4 + J6M8.5. Lanes C–J (except F) contain 1 μg each of the following components: (C) J5M8.1 + J5M8.2 + J5M8.3; (D) J5M8.1 + J5M8.2 + J5M8.3 + J5M8.4; (E) all strands of J5M8; (G) all strands of J6M8; (H) J6M8.1 + J6M8.2 + J6M8.3 + J6M8.4 + J6M8.5; (I) J6M8.1 + J6M8.2 + J6M8.3 + J6M8.4; (J) J6M8.1 + J6M8.2 + J6M8.3. Note in particular that the full junctions in lanes E and G are unstable, with significant amounts of dimer present.

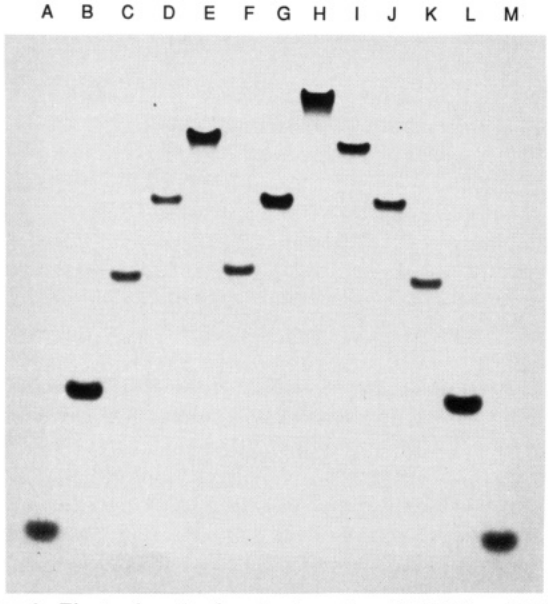


FIGURE 3: Electrophoresis of equimolar ratios of J5YLG and J6YLG. The gel shown here is a 10% native gel stained with Stainsall dye. Lane F contains a three-arm junction, JYG, with 16 nucleotide pairs per arm, and lane G contains a four-arm junction, JXG, with 16 nucleotide pairs per arm. Lanes A and M contain 4 μg of strand J5YLG.3, lane B contains 2 μg of J5YLG.4 + J5YLG.5, and lane L contains 2 μg of each of J6YLG.4 + J6YLG.5. Lanes C–K (except F and G) contain 1 μg each of the following components: (C) J5YLG.1 + J5YLG.2 + J5YLG.3; (D) J5YLG.1 + J5YLG.2 + J5YLG.3 + J5YLG.4; (E) all strands of J5YLG; (H) all strands of J6YLG; (I) J6YLG.1 + J6YLG.2 + J6YLG.3 + J6YLG.4 + J6YLG.5; (J) J6YLG.1 + J6YLG.2 + J6YLG.3 + J6YLG.4; (K) J6YLG.1 + J6YLG.2 + J6YLG.3. Note that in contrast to the analogous Figure 2, all species migrate as single bands.

a function of polyacrylamide concentration. All four junctions contain 16 nucleotide pairs per arm.

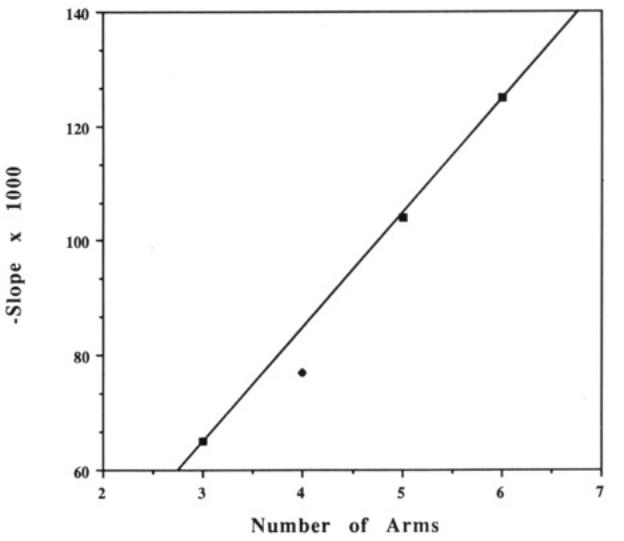


FIGURE 4: Ferguson plot analysis of three-arm, four-arm, five-arm, and six-arm junctions. The Ferguson parameters of the junctions, given as (intercept, –slope), are as follows: J3G (0.506, 0.065), J4G (0.533, 0.074), J5G (0.644, 0.104), J6G (0.761, 0.125). These increasingly negative slopes are plotted as a function of the number of arms of the junction. The line was fitted excluding JXG, whose position was then plotted. The increment for JXG is about 40% lower than expected, on the basis of the relationship between the other three junctions.

The mobility, M, of a molecule as a function of total gel concentration, T, may be described by the well-known relationship (Rodbard & Chrambach, 1970)

$$\log M = \log M_0 - K_R T$$

where M_0 is the free mobility and K_R is the retardation coefficient. Rodbard and Chrambach (1971) have shown that K_R is an approximately linear function of the surface area of the electrophoresing species. If one models an N-arm junction as a group of N cylinders emanating from a central site, the surface area of a series of junctions containing an increasing number of arms of identical length will increase linearly, as cylinders are added. Figure 4 illustrates K_R as a function of N. The linear dependence seen, when $N = 3$, 5, and 6, is in agreement with this model. The lower value of K_R for the four-arm junction is in agreement with experiments suggesting that its arms partially occlude each other (Cooper & Hagerman, 1987, 1989; Churchill et al., 1988; Chen et al., 1988; Duckett et al., 1988; Murchie et al., 1989), thereby decreasing its overall effective surface area. About 40% of the surface area of one incremental arm is occluded in the four-arm junction, relative to the exposed area of the other junctions.

Hydroxyl Radical Analysis. The structure of four-arm junctions has been analyzed in solution by means of hydroxyl radical chemical attack experiments (Churchill et al., 1988), in which the hydroxyl radicals have been generated by Fenton chemistry involving Fe(II)EDTA^{2-} (Tullius & Dombroski, 1985, 1986). The strategy of the experiments is to compare the chemical attack pattern of each strand when it is part of a junction, with the pattern obtained when the strand is complexed with its normal Watson–Crick complement. Experiments with four-arm junctions indicate that the patterns of two noncontiguous strands are the same in both cases, while the other two strands exhibit protection at the site of the junction. From these experiments it has been concluded that strands with the same patterns in both pairing environments probably have double-helical conformations near the junction, while the other strands form crossover structures (Churchill et al., 1988). The arms shared by helical strands are thought

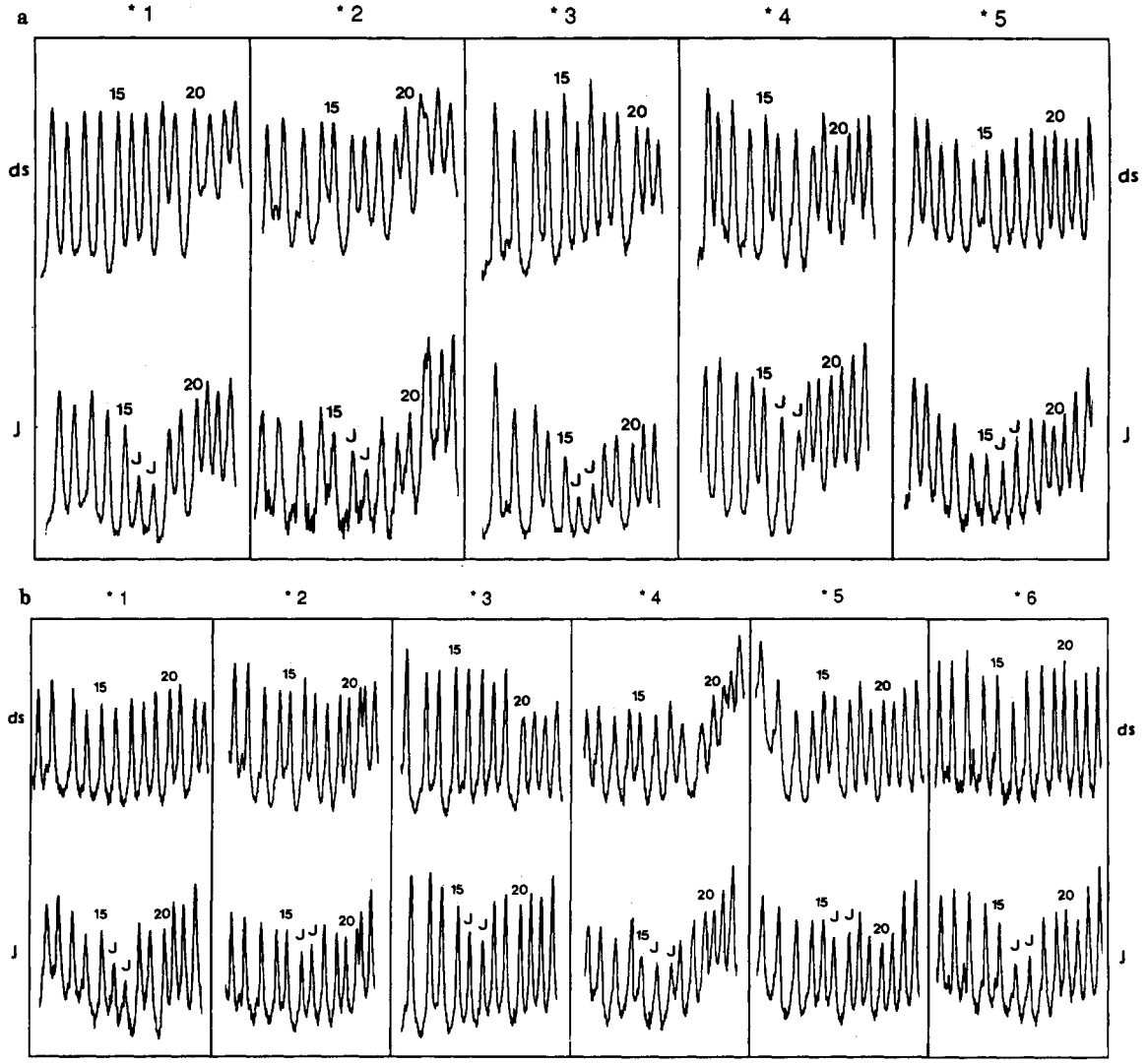


FIGURE 5: Hydroxyl radical cleavage of five-arm and six-arm junctions. (a) The cleavage pattern of J5YLG is shown as a densitometer trace of the denaturing gel on which the products of the reaction are run. The labeled strand is indicated above each of the five panels, and the numbers indicate the residue number. These positions have been calibrated by the Maxam–Gilbert (1977) G-reaction (Maxam & Gilbert, 1977). The upper portion of each panel (ds) corresponds to each strand in a double-helical environment, while the lower portion (J) indicates the same results when the strand is incorporated into the junction. The two nucleotides that flank the junction (16 and 17) are indicated by a "J" to facilitate interpretation. Note that, in every case, the junction is protected to some extent, relative to the double strand. (b) This figure illustrates the same experiment conducted on J6YLG. The same conventions apply as in (a). Note that strand 5 does not appear to be protected relative to the double strand, in contrast to the rest of the strands of J6YLG.

to form a continuously stacked helical domain within the junction structure, so that the four-arm junction contains two helical domains.

Figure 5a illustrates the same comparative experiment for the five-arm junction. None of the strands in the five-arm junction shows the same chemical attack pattern as it does in the duplex. Each is somewhat protected in the vicinity of the junction, suggesting that no helical domains dominate within the structure. It is clear that all five arms of a single molecule cannot simultaneously be members of a two-arm stacking domain. The chemical analysis performed here samples the entire population of molecules and cannot distinguish readily between an average of several different domain structures and structures containing no domains. Thus, it is possible that all pairs of adjacent arms sometimes form stacked helical domains, but none of the five possible domains is strongly preferred over any of the others.

A somewhat different result is seen for the six-arm junction, whose chemical attack pattern is shown in Figure 5b. Strand 5 shows approximately the same cleavage pattern both in the double-stranded environment and in the junction. This finding suggests that arms V and VI form a double-helical stacking domain, but that none of the other pairs of adjacent arms do so to the exclusion of possible domains formed by the adjoining pair. In particular, there is very little or no stacking between arms IV and V or between arms VI and I; likewise, there is either little stacking or no stacking dominance between arms I and II, II and III, or III and IV.

Endonuclease VII Cleavage. Endonuclease VII is an enzyme known to cleave unusual DNA structures (Kemper et al., 1990), in particular branched junctions (Mizuuchi et al., 1982). We and others have shown that it cleaves the crossover strands of four-arm Holliday junction analogues (Mueller et al., 1988; Duckett et al., 1988), and not the helical strands. Recently, we have shown that the enzyme requires more than eight nucleotide pairs in junction arms that are cleaved, and we have modeled some of the features of the molecule (Mueller et al., 1990). The endonuclease VII cleavage pattern of the five-arm and six-arm junctions provides an analysis complementary to the hydroxyl radical experiments. The gel illustrating the five-arm junction cleavage pattern is shown in Figure 6a. Each strand is primarily cleaved two or three

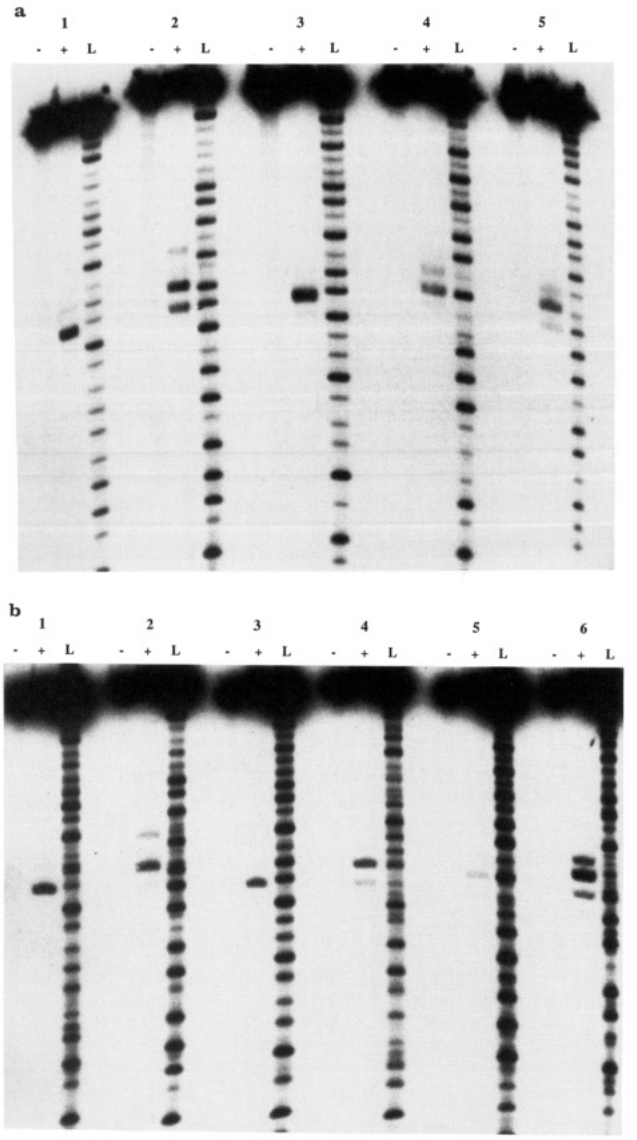


FIGURE 6: Cleavage of five-arm and six-arm junctions by endonuclease VII. (a) This is a denaturing gel showing the cleavage pattern of J5YLG when the junction is treated with endonuclease VII. The gel is divided into five sets of three lanes each, and the number of the labeled strand is shown above each set. The lanes are indicated as untreated (−), treated (+), and a sequencing ladder (L). The ladder is a Maxam–Gilbert sequencing ladder that has been calibrated against known standards to account for the difference in cleavage position between the chemical and enzymatic reactions. The positions of cleavage are as follows: (1) 18; (2) 18, 19, 20; (3) 18; (4) 18, 19, 20; (5), 18, 19, 20. (b) This gel indicates the same experiment performed on J6YLG. Note that there is virtually no cleavage of strand 5. The positions of cleavage are as follows: (1) 18; (2) 18, 19, 20; (3) 18; (4) 18, 19, 20; (5) 19; (6) 18, 19, 20. The results of this experiment are in agreement with arms V and VI (see Figure 1) forming a dominant stacking domain.

residues 3′ to the junction, in agreement with earlier studies of the enzyme (Jensch & Kemper, 1986). Although their patterns differ somewhat, none of the strands is cleaved particularly better or worse than the others. This result suggests no well-defined structural features for the complex, other than the presence of the junction.

The gel illustrating the six-arm junction cleavage pattern is seen in Figure 6b. Extensive cleavage of all strands except 5 may be noted from this gel, again primarily two or three residues 3′ to the junction. Strand 5 is hardly cleaved at all. Since the enzyme has been shown to cleave structures that are not in helical conformations, one is led to the conclusion that strand 5 largely adopts a helical conformation. This finding is in agreement with the results of the hydroxyl radical analysis; it supports the notion that arms V and VI form a helical domain but that this is the only dominant domain within the structure.

Discussion

Stability and Structure of Five-Arm and Six-Arm Junctions. We have demonstrated that it is possible to construct five-arm and six-arm junctions. However, the cost in free energy of building these larger junctions is greater than the cost associated with three-arm or four-arm junctions. This is seen in the requirement for longer arms to generate a single band on polyacrylamide gels under the conditions chosen. We have tried to stabilize J5M8 on gels with the maximum amount of spermine in the running buffer that does not precipitate the material; the junction does not remain intact under these conditions either. We have not yet tried to determine the minimum-length arms necessary to stabilize five-arm and six-arm junctions. When eight nucleotide pairs per arm proved insufficient, we have merely doubled their lengths.

As first suggested for branched junctions by Sigal and Alberts (1972), it is possible for any two adjacent arms to form a helical stacking domain. A surprising result of the hydroxyl radical (Churchill et al., 1988) and other (Cooper & Hagerman, 1987, 1989) studies of the four-arm junction J1 (and its extended-arm variants) is that two of the stacking domains predominate strongly over the other two, although it is possible to change this structure by imposing tethered constraints (Kimball et al., 1990). It has also been shown that the sequence directly flanking the junction determines this dominance (Chen et al., 1988; Duckett et al., 1988). In the five-arm junction, we find no evidence for a dominant stacking domain. From model building (Seeman, 1989), there is no apparent impediment to forming a single stacking domain in this structure, although the presence of two simultaneous stacking domains appears very strained. By contrast, the six-arm junction appears to contain a single prominent stacking domain. Model building suggests that three simultaneous stacking domains can exist in this structure, but evidence for only one has been obtained. The stacking sequence in this domain, a G·C pair stacked 5′ → 3′ on an A·T pair (the first base is on the continuous helical strand), forms one of the dominant domains in J1, but it is present in the five-arm junction without dominating. The other stacking sequence that dominates in J1, C·G stacked on T·A, is present in both the five-arm and the six-arm junctions without appearing to dominate.

Macromolecular Construction with Five-Arm and Six-Arm Junctions. The ability to construct five-arm and six-arm junctions means that a large number of structures can be constructed from branched DNA molecules that were previously impossible. Williams (1977) has tabulated the space-filling networks with 432 symmetry that can be formed from Platonic and Archimedean solids. These structures can, in principle, be formed from the helix axes of DNA molecules; the symmetry is much lower if one considers the entire DNA molecule. Of the nineteen networks listed, only four can be constructed from branched components with four arms. However, only four of these networks require components with more than six arms. The left side of Figure 7 illustrates a five-connected space-filling network formed from repetition of octahedra and a truncated cube. The right side of Figure 7 illustrates the conceptually simplest six-connected network, one generated by the packing of cubes.

In addition to periodic networks, closed polyhedral DNA structures, such as deltahedra based on icosahedral symmetry,

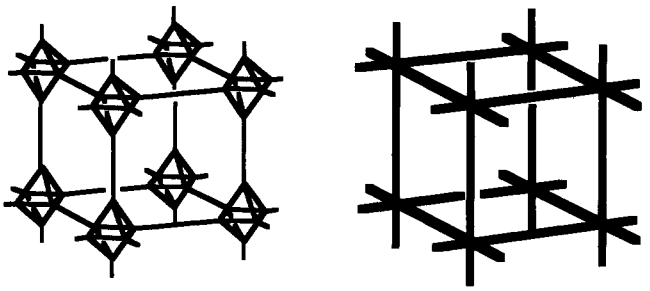


FIGURE 7: Five-connected and six-connected networks. The left side of this drawing illustrates a five-connected network. The polyhedra that compose it are regular octahedra whose centers are placed at the corners of a cube. To accommodate the octahedra, the cube has been transformed into a truncated cube. The size of the octahedra determines the symmetry of the faces of the truncated cube; when the edges of the octahedron are the same length as the space between them, the face of the truncated cube is a regular octagon. The centers of the cubes and octahedra are loci of 432 symmetry, although DNA structures would be of lower symmetry because their base sequences would not be symmetric. Note that every vertex of this network has five edges meeting there. The simplest conceptual network, the six-connected cubic lattice, is shown on the right side of this drawing. If the object depicted were ligated into a lattice, only one "box" out of eight would correspond to this object, and the others would be formed by connections between the boxes. The object shown has, ideally, 432 symmetry, but a lower symmetry box and lattice could be formed by making the edges and/or the connections between edges of unequal lengths.

are now possible. These triangle-based structures are likely to be flexurally stiff [e.g., Edmondson (1987)]; this is important, because of the flexibility noted for some constructions involving DNA junctions (Ma et al., 1986; Petrillo et al., 1988). Conversely, the unexpected torsional stiffness found when assembling a cube from three-arm junctions (Chen & Seeman, 1991) suggests that torsion-absorbing relief might be necessary in some constructions; the insertion of extra, unused, arms at vertices might provide this relief. For example, relief could be provided for four-connected objects by the use of five-arm junctions, with the extra arm functioning as a shock absorber; the easy synthesis of a quadrilateral from three-arm junctions (Chen et al., 1989) is possibly a precedent for the utility of an extra arm. Finally, it is useful to point out that it is now feasible to consider making low-symmetry constructions predicated on six-arm junctions, such as the DNA framework of a recently proposed biochip (Robinson & Seeman, 1987).

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.

Caspar, D. L. D., & Klug, A. (1962) *Cold Spring Harbor Symp. Quant. Biol. 27*, 1–24.

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.

Cohen, S. N., Chang, A. C. Y., Boyer, H. W., & Helling, R. B. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3240–3244.

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

Copper, 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.

Edmondson, A. C. (1987) *A Fuller Explanation: The Synergetic Geometry of H. Buckminster Fuller*, pp 232–243, Birkhäuser, Boston.

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

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.

Jensch, F., & Kemper, B. (1986) *EMBO J. 5*, 181–189.

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

Kallenbach, N. R., Ma, R.-I., Wand, A. J., Veeneman, G. H., van Boom, J. H., & Seeman, N. C. (1983b) *J. Biomol. Struct. Dyn. 1*, 158–168.

Kemper, B., & Garabett, M. (1985) *Eur. J. Biochem. 115*, 123–131.

Kemper, B., Pottmeyer, S., Solaro, P., & Kosak, H. (1990) *Structure and Methods* (Sarma, R. H., & Sarma, M. H., Ed.) Vol. 1., pp 215–229, Adenine Press, New York.

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.

Kornberg, A. (1980) *DNA Replication*, W. H. Freeman, New York.

Kucherlapati, R., & Smith, G. R. (1988) *Genetic Recombination*, American Society for Microbiology, Washington, DC.

Ma, R.-I., Kallenbach, N. R., Sheardy, R. D., Petrillo, M. L., & Seeman, N. C. (1986) *Nucleic Acids Res. 14*, 9745–9753.

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.

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

Petrillo, M. L., Newton, C. J., Cunningham, R. P., Ma, R.-I., Kallenbach, N. R., & Seeman, N. C. (1988) *Biopolymers 27*, 1337–1352.

Robinson, B. H., & Seeman, N. C. (1987) *Protein Eng. 1*, 295–300.

Rodbard, D., & Chrambach, A. (1970) *Proc. Natl. Acad. Sci. U.S.A. 65*, 970–977.

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., & Kallenbach, N. R. (1983) *Biophys. J. 44*, 201–209.

Seeman, N. C., & Kallenbach, N. R. (1988) in *Molecular Structure* (Stezowski, J., Ed.) pp 189–194, Oxford University Press, Oxford.

Seeman, N. C., Maestre, M. F., Ma, R.-I., & Kallenbach, N. R. (1985) in *The Molecular Basis of Cancer* (Rein, R., Ed.) Progress in Clinical and Biological Research 172A, pp 99-108, Alan 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*, 789-791.
Tullius, T. D., & Dombroski, B. (1985) *Science 230*, 679-681.
Tullius, T. D., & Dombroski, B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469-5473.
Wells, A. F. (1977) *Three-dimensional Nets and Polyhedra* John Wiley & Sons, New York.
Wemmer, D. E., Wand, A. J., Seeman, N. C., & Kallenbach, N. R. (1985) *Biochemistry 24*, 5745-5749.
Williams, R. (1979) *The Geometrical Foundation of Natural Structure*, Dover, New York.

Synthesis and Biological Evaluation of a Fluorescent Analogue of Folic Acid†

Terence P. McAlinden,‡ John B. Hynes,§ Shirish A. Patil,§ G. Robbin Westerhof,∥ Gerrit Jansen,∥ Jan H. Schornagel,∥ Suresh S. Kerwar,⊥ and James H. Freisheim*,‡

Department of Biochemistry and Molecular Biology, Medical College of Ohio, P.O. Box 10008, Toledo, Ohio 43699-0008, Department of Pharmaceutical Sciences, Medical University of South Carolina, Charleston, South Carolina 29425-2303, Department of Oncology, Free University Hospital, De Boelelaan 1118, 1081 HV Amsterdam, The Netherlands, and Medical Research Division, American Cyanamid Company, Lederle Laboratories, Pearl River, New York 10965

Received January 8, 1991; Revised Manuscript Received March 13, 1991

ABSTRACT: A fluorescein derivative of the lysine analogue of folic acid, N^{α}-pteroyl-N^{ϵ}-(4′-fluoresceinthiocarbamoyl)-L-lysine (PLF), was synthesized as a probe for dihydrofolate reductase (DHFR) and a membrane folate binding protein (m-FBP). Excitation of PLF at 282 nm and at 497 nm gave a fluorescence emission maximum at 518 nm. Binding of PLF to human DHFR or human placental m-FBP results in approximately a 20-fold enhancement in the magnitude of the fluorescence emission, suggesting that the ligand interacts with a hydrophobic region on these proteins. Additional evidence suggests that an energy transfer may occur between the pteridine and the fluorescein moieties. PLF binds to the active site of human DHFR since methotrexate (MTX) competes stoichiometrically and the denatured enzyme in the presence of PLF did not exhibit fluorescent enhancement. The dissociation constant for the fluorescein derivative with respect to human DHFR is 115 nM as compared to 111 nM for folic acid. The K_i value for the competitive inhibition of human DHFR by the fluorescent analogue of folic acid is 2.0 μM compared to 0.48 μM for folic acid. PLF was reduced to N^{α}-(7,8-dihydropteroyl)-N^{ϵ}-(4′-fluoresceinthiocarbamoyl)-L-lysine (H_2PLF) and assayed by the enzymatic conversion to the tetrahydro derivative. The K_m value for human DHFR for the dihydrofolate analogue is 2.0 μM. The K_D value for H_2PLF to human DHFR is 47 nM as compared to 44 nM for dihydrofolate. The K_D values for both H_2PLF and PLF indicate that the fluorescein moiety does not significantly affect folate binding in enzyme binary complexes. PLF also binds to a m-FBP on a human leukemia CCRF/CEM-FBP overproducing cell line, resulting in a fluorescence enhancement that is completely abolished in the presence of excess folic acid. There is an absence of fluorescence enhancement when PLF is added to CCRF/CEM cells or CCRF/CEM-7A cells, which overproduce the carrier protein involved in the uptake of reduced folates and MTX, indicating the specificity of PLF toward the m-FBP. The binding of PLF to CCRF/CEM-FBP cells is similar to that of folic acid as demonstrated by [^{3}H]folic acid competition for binding and by a K_D determination of 1.6 nM. CCRF/CEM-FBP cells can be separated from parental CCRF/CEM and from CCRF/CEM-7A cells by fluorescence-activated cell sorting using this fluorescent analogue of folic acid.

Fluorescent analogues of the classical folic acid antagonist methotrexate (MTX)[1] have previously been synthesized and evaluated as probes of dihydrofolate reductase (DHFR)[1] structure and function (Freisheim et al., 1986). X-ray crystallographic studies on DHFR from *Lactobacillus casei* and *Escherichia coli* (Matthews et al., 1977, 1978) and human sources (Oefner et al., 1988; Davies et al., 1990) reveal certain invariant residues in the primary sequence of DHFR which have important functions in substrate and inhibitor binding. An invariant arginine residue in bacterial (Arg 57) and vertebrate DHFRs (Arg 70) is involved in a charge interaction with the α-carboxylate group of the L-glutamate of MTX. The

† This research was supported, in part, by grants from the National Institutes of Health to J.H.F. (CA41461) and to J.B.H. (CA25014) and from the Dutch Cancer Society to G.J. and J.H.S. (IKA-89-34). J.H.F. is the recipient of the Harold and Helen McMaster Chair in Biochemistry and Molecular Biology.

* Correspondence should be addressed to this author.
‡ Medical College of Ohio.
§ Medical University of South Carolina.
∥ Free University Hospital.
⊥ Lederle Laboratories.

[1] Abbreviations: PLF, N^{α}-pteroyl-N^{ϵ}-(4′-fluoresceinthiocarbamoyl)-L-lysine; DHFR, dihydrofolate reductase; FBP, folate binding protein; MTX, methotrexate [(4-amino-4-deoxy-N^{10}-methylpteroyl)glutamic acid]; H_2PLF, N^{α}-(7,8-dihydropteroyl)-N^{ϵ}-(4′-fluoresceinthiocarbamoyl)-L-lysine; H_4PLF, N^{α}-(5,6,7,8-tetrahydropteroyl)-N^{ϵ}-(4′-fluoresceinthiocarbamoyl)-L-lysine; MTX-F, 1-[[(4-amino-4-deoxy-N^{10}-methylpteroyl)-γ-L-glutamyl]amino]-5-[(4′-fluoresceinthiocarbamoyl)amino]pentane; DNS-Cl, 5-(*N*,*N*-dimethylamino)-1-naphthalenesulfonyl chloride; Tris, tris(hydroxymethyl)aminoethane; NADPH, reduced nicotinamide adenine dinucleotide phosphate; FAB-MS, fast atom bombardment mass spectrum; DMF, dimethylformamide; DEPC, diethyl phosphorocyanidate; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HBSS, HEPES balanced saline solution; HPLC, high-performance liquid chromatography; TLC, thin-layer chromatography; NMR, nuclear magnetic resonance; APA, 4-amino-4-deoxy-N^{10}-methylpteroic acid.

0006-2960/91/0430-5674$02.50/0