Article No. mb981991 *J. Mol. Biol.* (1998) **282**, 59–70

JMB

Sequence Dependence of Branch Migratory Minima

Weiqiong Sun, Chengde Mao, Furong Liu and Nadrian C. Seeman*

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

The Holliday junction is a central intermediate in the process of genetic recombination. The position of its branch-point can relocate through an isomerization known as branch migration. This migration occurs because the branch-point is flanked by homologous symmetry. All attempts at modeling the kinetics of branch migration have relied on the assumption that branch migration minima are sequence-independent. We have tested that assumption here, using a competition assay based on symmetric immobile branched junctions; these are junctions that cannot undergo branch migration, despite the fact that they are flanked by homology. The assay used is predicated on the non-association of strands displaced in the assay; we have tested this assumption, and have performed our experiments under conditions where we know that it is true. We have measured the free energy of relocating a branched junction from a fixed non-homologous sequence to all possible dimeric symmetric sequences. We find that the assumption of sequence-independence is often valid, but that it is not universally true. We find that the flanking sequences can have a marked effect on the free energy measured, both for extensions of symmetry and for reversals of flanking nucleotides. We have varied the temperature in our experiments, and have derived both enthalpies and entropies for the different sequences. The entropies are largely unfavorable, whereas the enthalpies are largely favorable; regardless of the signs of these quantities, we see that this is another system where enthalpy-entropy compensation is operative.

Keywords: Holliday junctions; branch migration; symmetric immobile junctions; branched DNA; genetic recombination

*Corresponding author

Introduction

Genetic recombination occurs universally in living organisms, from viruses to humans. A great deal of DNA metabolism is clearly devoted to ensuring the stability required for its role as genetic material, but recombination is a process that affords the flexibility necessary for the adaptation of species to a changing environment. The basic molecular feature of recombination is the interaction of two pieces of DNA to yield new genetic material that may incorporate segments of both interacting molecules, either directly or in the form of copied information. The resulting DNA may show such alterations as insertions, deletions, changes of sequence, rearrangements or exchanges of flanking markers. The Holliday (1964) junction intermediate is a central paradigm in the molecular mechanism of this process. It has been shown to be an authentic intermediate in site-specific recombination (Hoess *et al.*, 1987; Kitts & Nash, 1987; Nunes-Duby *et al.*, 1987), and it is believed to be involved in homologous recombination (DasGupta *et al.*, 1981).

The Holliday junction consists of four strands, arranged in four double-helical arms. Figure 1 illustrates the way in which the Holliday junction can participate in genetic recombination. At the first stage (I, in Figure 1), DNA strands containing homologous sequences, but different flanking markers, are aligned. The allelic flanking markers are A and α, B and β. The Holliday junction is a four-stranded intermediate (II) in which two strands have (effectively, but not mechanistically) been nicked and religated so that they fuse their sequences and form a branch-point; the other two strands remain intact and are not involved directly in branching in the equilibrium structure (Lilley & Clegg, 1993; Seeman & Kallenbach, 1994). The strands that join the double-helical domains are called the crossover strands, and the other two strands are called the

E-mail address of the corresponding author:
ned.seeman@nyu.edu

0022–2836/98/360059–12 $30.00/0

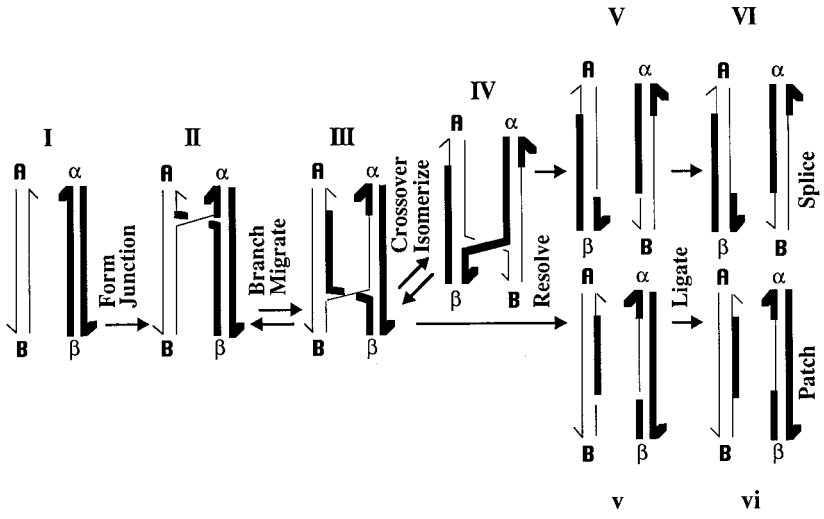

Figure 1. The formation and resolution of the Holliday structure in genetic recombination. The process is shown proceeding from the left to the right. Each of the possible stages is labeled with capital or small Roman numerals. In the first stage, I, two homologous double-helices of DNA align with each other. The two strands of each duplex are indicated by the two pairs of lines terminated by half arrows, which indicate the 3′ ends of the strands. Strands are distinguished by their thickness. Each of these two homologous regions carries a flanking marker, A and B in the strands on the left, and α and β on the right. After the first step, the homologous pairs have formed a Holliday intermediate, II, by exchanging strands. Note that the two crossover strands are composite strands with both a thick and a thin portion formed through any of a number of possibilities. The parallel representation of the Holliday junction is shown. The homologous 2-fold sequence symmetry of this structure permits it to undergo the iterative isomerization process, branch migration; movement in the direction indicated results in structure III. The Holliday intermediate may or may not undergo the crossover isomerization process to produce structure IV, in which the crossover and non-crossover strands are switched. Note that this process has meaning only if the Holliday intermediate is 2-fold symmetric, rather than 4-fold symmetric. Although indicated as separate, the crossover isomerization process could be a feature of branch migration (Mueller *et al.*, 1988). If crossover isomerization occurs an odd number of times, resolution by cleavage of the crossover strands yields structure V, but structure v results if crossover isomerization occurs an even number of times (including 0) before cleavage. Ligation of v generates a patch recombinant, vi; this is a pair linear duplex DNA molecules containing heteroduplex DNA because of branch migration, but which have retained the same flanking markers. Ligation of VI yields splice recombinant molecules, which have exchanged flanking markers.

helical strands. Homology permits the Holliday junction to undergo branch migration, an isomerization that relocates the branch-point (III; e.g. see Hsieh & Panyutin, 1995). In addition to branch migration, another isomerization appears possible, the crossover isomerization reaction. In crossover isomerization, the crossover strands and the helical strands switch functions (IV). Thus, the fused strands now take on the role of the helical strands, and the intact strands become the crossover strands. The isomerization is spontaneous (Li *et al.*, 1997; the position of the equilibrium has been examined in junctions flanked by both symmetric (Zhang & Seeman, 1994) and asymmetric (Miick *et al.*, 1997) sequences. Resolution of the unisomerized Holliday junction (III) by a resolvase that cleaves the crossover strands (v), such as endonuclease VII (Mueller *et al.*, 1988), yields patch recombinants (vi) upon ligation; patch recombinants can lead to gene conversion. However, resolution of the isomerized Holliday junction (IV) by the same resolvase (V) produces splice recombinants (VI), corresponding to the exchange of flanking markers.

Branch migration complicates the physical characterization of branched junctions, because it can lead to a heterogeneous population of molecules. However, branch migration is consequence of the junction being flanked by homologous (two-fold) sequence symmetry, as illustrated by Figure 2. Elimination of homologous symmetry in synthetic molecules produces immobile branched junctions whose branch-points are fixed (Seeman, 1982; Kallenbach *et al.*, 1983). It is also possible to fix the branch-point of a symmetric junction, by coupling its motion to that of an immobile junction within a double crossover context (Zhang *et al.*, 1993); these junctions are known as symmetric immobile junctions.

The branch-point in a Holliday junction is clearly a pivotal site. There are at least four transformations of this point, three of which are illustrated by Figure 1. (1) The formation of the branch-point from two duplex molecules; (2) branch migration from one position to another; (3) crossover isomerization; and (4) interconversion of parallel and antiparallel junctions (not shown). In the past, we have used immobile branched junctions to measure the thermodynamics of branch-point formation (Lu *et al.*, 1992) and of the interconversion of parallel and antiparallel conformers (Lu *et al.*, 1991). Most recently, we have used symmetric immobile junctions to estimate the thermodynamics of crossover isomers for symmetric sequences (Zhang & Seeman, 1994). The remaining key transformation that lacks thermodynamic characterization is branch migration, which we describe here.

It is clear from Figure 1 that branch migration is a key step in Holliday-mediated recombination. The position to which the branch point migrates determines the extent of heteroduplex DNA, potentially determining the products of the recombination event; it is evident that the heteroduplex

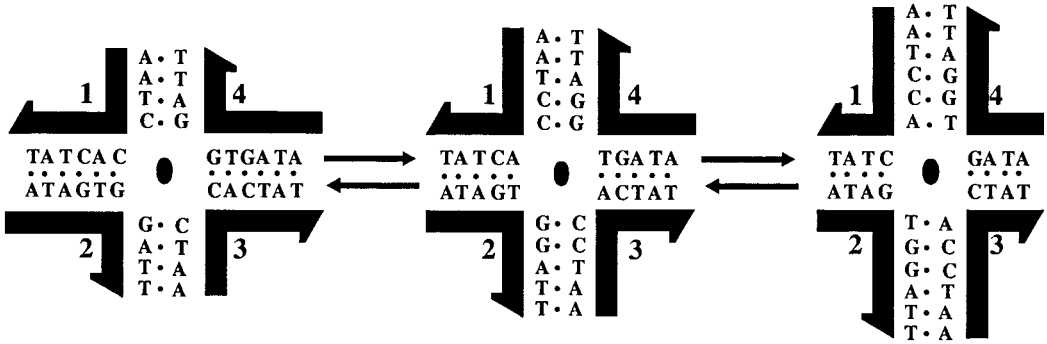

Figure 2. Branch migration. The thick black lines correspond to strand backbones, with the half-arrowheads pointing in the 5′ to 3′ direction of the molecule. The lens-shaped object in the center of each branched structure indicates the 2-fold sequence-symmetry of the branched junction: strand 1 has the same sequence as strand 3, and strand 2 has the same sequence as strand 4. In the central molecule, only the five nucleotides flanking the central branch on each side are shown. The central molecule is shown to branch migrate in each direction. To the right, the base-pairs in the horizontal arms re-pair to join the vertical arms; to the left, the base-pairs in the vertical arms re-pair to form the horizontal arms. The figures on the left and the right are free to migrate again, as they, too, are 2-fold symmetric. Repeated migration eventually results in the production of two linear molecules from each branched molecule.

regions in products vi and VI of Figure 1 are a consequence of the branch migration that converts their precursor, species II, to species III. It is known that this reaction may be spontaneous (e.g. see Thompson *et al.*, 1976), but it may also be enzyme-catalyzed (Lloyd & Sharples, 1993; Tsaneva *et al.*, 1993). A number of investigators have measured the kinetics of spontaneous double-stranded branch migration in elegant systems that have used a similar approach: branched molecules are assembled; at a fixed time they are freed to migrate; the resolution of the branched species to linear species is monitored, and the step rate is estimated from this observation by modeling a random walk with an absorbing barrier. In all experiments (Thompson *et al.*, 1976; Warner *et al.*, 1979; Gellert *et al.*, 1983; Johnson & Symington, 1993; Panyutin & Hsieh, 1994) and simulations (Robinson & Seeman, 1987) of which we are aware, the probability of migration from one migratory minimum to another is assumed to be independent of sequence, as shown in the top portion of Figure 3. Here, we have asked whether that assumption is warranted, or whether the sequence can affect the migratory minima, as shown at the bottom of Figure 3.

We have measured the relative free energies of branched junctions flanked by all possible symmetric dinucleotide sequences. In order to do this, we have used symmetric immobile junctions (Zhang *et al.*, 1993) predicted on DAO double crossover molecules (Fu & Seeman, 1993); the topology of a DAO double crossover molecule is illustrated at the top of Figure 4. We have made these measurements by means of a competition experiment illustrated in the lower portion of Figure 4. Species B is a double crossover molecule that contains two branched junctions flanked by the central nucleotides of the well-characterized branched junction J1 (Seeman & Kallenbach, 1983); its favored crossover isomer is well known (Churchill *et al.*, 1988). Strands 1 and 3 are the crossover strands of the upper junction, and strands 2 and 4 are the crossover strands of the lower junction. One helical turn (ten nucleotide pairs) from the lower junction on each helical domain is a sym-

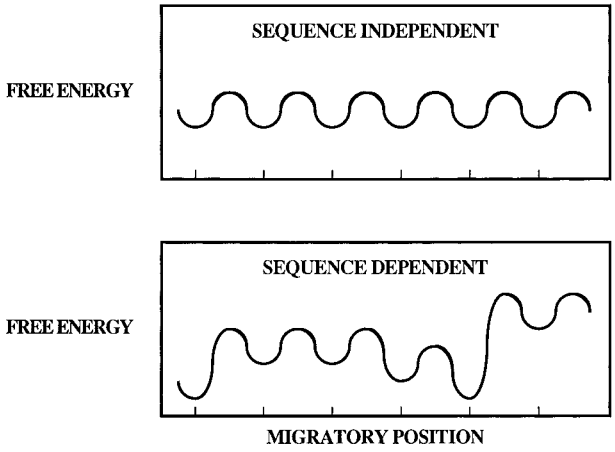

Figure 3. Models for branch migration. Two qualitative models for branch migration are presented here. In each case, we show six adjacent branch migratory minima. The minima represent the positions that correspond to intact junctions; if, as is believed (Lilley & Clegg, 1993; Seeman & Kallenbach, 1994), the nucleotides flanking an intact junction are paired, then these minima correspond to those structures. The maxima represent the barriers over which the system must climb in order to get to the next position. The panel at the top represents a sequence-independent model of the branch migration, in which every minimum represents the same free energy, and every barrier is also the same. The bottom panel represents a sequence-dependent model, in which some sequences produce a junction with lower free energy than others. We have made no attempt to alter the barrier heights, although fixed barrier heights, relative to minima, result in inherently higher barriers.

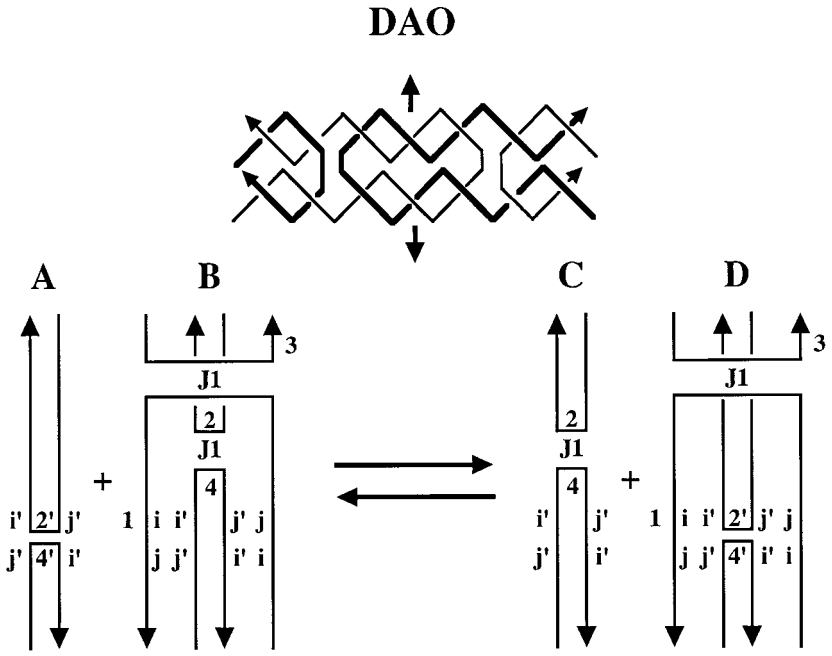

Figure 4. The experimental system used to measure free energy differences in symmetric immobile branched junctions. The top panel illustrates the topology of the DAO molecule, a double-crossover molecule in which the crossovers are separated by an odd number of double helical half-turns. In the molecule shown, they are separated by three half-turns. In the molecule shown, they are separated by three half-turns. The molecule contains four strands, two drawn with thin lines and two drawn with thick lines; their 3′ ends are indicated by arrowheads. The vertical arrows indicate the 2-fold symmetry axis that relates the thick strands to the thin strands. The bottom panel illustrates the experiment performed here in order to measure the sequence-dependence of the free energy of branch-points. The unwrapped representation of DAO double-crossover molecules is shown, featuring a conventionally drawn antiparallel junction between the inner strands, and a crossover that spans the molecule between outer strands; this form is easier to use when sequence information must be indicated. Species B is a conventional double-crossover molecule, in which each of the branched junctions is flanked by the J1 sequence (Seeman & Kallenbach, 1983). The strand numbers are shown; strands 1 and 3 from one J1 junction and strands 2 and 4 form the other one. A double-helical turn away is a potentially symmetric site, flanked by the sequence 5′-i,j-3′. Species A represents two strands, 2′ and 4′, that do not associate in solution. Their combined sequence is the same as in strands 2 and 4, but their crossover-point is different, occurring at the sequence 5′-j′,i′-3′. When A and B are equilibrated together, strands 2′ and 4′ can displace strands 2 and 4 from B, to produce species D and species C, which is just the two non-interacting strands 2 and 4. Species D is a symmetric immobile junction, because the lower junction is flanked by the sequences 5′-i,j-3′ and 5′-j′,i′-3′ that exhibit local 2-fold (homologous) symmetry. The reaction shown corresponds to moving the lower junction from the J1 sequence to the $[5'\text{-i,j-}3', 5'\text{-i,j-}3']_2$ symmetric sequence.

metric sequence, 5′-i,j-3′, paired by normal Watson-Crick interactions with its complement, 5′-j′,i′-3′; in this notation, i and j represent any of the four conventional deoxynucleotides. The sequence 5′-i,j-3′ occurs twice on strand 1, and its complement is seen twice on strand 4. This double crossover is equilibrated with strands 2′ and 4′. These strands are drawn stacked for clarity but, in fact, they are designed not to interact with each other or themselves. The sequence of the 5′ portion of strand 4′ and the 3′ portion of strand 2′ is exactly the same as the sequence of the 5′ portion of strand 4 and the 3′ portion of strand 2. However, the discontinuity in their sequences (the ultimate crossover site) is in a different position. Likewise, the sequence of the 5′ portion of strand 2′ and the 3′ portion of strand 4′ is the same as the 5′ portion of strand 2 and the 3′ portion of strand 4. Thus, strands 2′ and 4′ can pair with strands 1 and 3 just as strands 2 and 4 can. If this occurs, double crossover D will replace double crossover B, and independent strands C (strands 2 and 4) will replace independent strands A (strands 2′ and 4′). This reaction corresponds to moving the lower junction from a site flanked by the asymmetric J1 sequence to a junction flanked by the symmetric sequence 5′-i,j-3′ and its complement 5′-j′,i′-3′.

The double crossover molecules B and D are stable species without detectable isomers (Fu & Seeman, 1993), so other substructures that could affect the thermodynamics of this system must derived from interactions involving 2, 2′, 4 and 4′. The equilibrium is represented by the ΔG for the reaction:

$$\Delta G^{ij} = \Delta G^{\circ ij}_{C} + \Delta G^{\circ ij}_{D} - (\Delta G^{\circ ij}_{A} + \Delta G^{\circ ij}_{B})$$

Given the readily testable assumption that 2, 2′, 4 and 4′ form no substructures in solution, ΔG^{ij} measures the cost of moving the junction from the J1 sequence to the symmetric immobile sequence 5′-i,j-3′. Likewise:

$$\Delta G^{J1} = \Delta G^{\circ J1}_{C} + \Delta G^{\circ J1}_{D} - (\Delta G^{\circ J1}_{A} + \Delta G^{\circ J1}_{B})$$

represents the free energy of moving the junction from the J1 sequence flanked by strands 2 and 4 to the J1 sequence flanked by strands 2′ and 4′. This is our baseline measurement, accounting for the free energy difference that results from moving the junction from one position to the other. The difference between two reactions:

$$\Delta^{J1}G^{ij} = \Delta G^{ij} - \Delta G^{J1}$$

places the ΔG measurement on an absolute scale.

The cost of moving the branch-point to sequence 5′-i,j-3′ from sequence 5′-k,l-3′ is given by:

$$\Delta\Delta G^{ij,kl} = \Delta^{J1}G^{ij} - \Delta^{J1}G^{kl}$$

The sequence-independent model illustrated at the top of Figure 3 requires that $\Delta\Delta G^{ij,kl}$ is zero for all sequences 5′-i,j-3′ and 5′-k,l-3′. We have measured $\Delta^{J1}G^{ij}$ for all 16 dinucleotides. Within the limits of our measurements, we find that the sequence-independent model at the top of Figure 3 is often true, but that there are clear exceptions; hence, the sequence-dependent model appears to be the most appropriate to use for modeling branch migration.

Results

DNA sequence design

The DNA molecules used in this work have been designed using the program SEQUIN (Seeman, 1990). The sequence of the molecule used is shown in Figure 5. The sequence naming convention is that the sequence 5′-i,j-3′ refers to the second version of the sequence on strand 1, so that it is flanked by 5′-G- and -C-3′, to yield 5′-G-i,j-C-3′. The distinction between this version and the first version (5′-C-i,j-G-3′) becomes relevant for certain controls discussed below.

Characterization of the double-crossover complex

The first issue to address is the formation and characterization of the two double-crossover complexes, B and D. Figure 6(a) contains an autoradiogram of a typical non-denaturing gel that illustrates the components of the reaction for which i = A and j = T. Lanes 1 to 3 contain complexes labeled in strand 3, and lanes 4 to 6 contain components labeled in strand 1. Lanes 1 and 4 contain the complex of strands 1, 2, 3 and 4. Lanes 2 and 5 contain the complex of strands 1, 2′, 3 and 4′. Note that the difference in the positions of the branch-points renders these species readily separable on this gel. Lanes 3 and 6 contain equilibrated mixtures of all six strands. Note that both species are visible in these lanes.

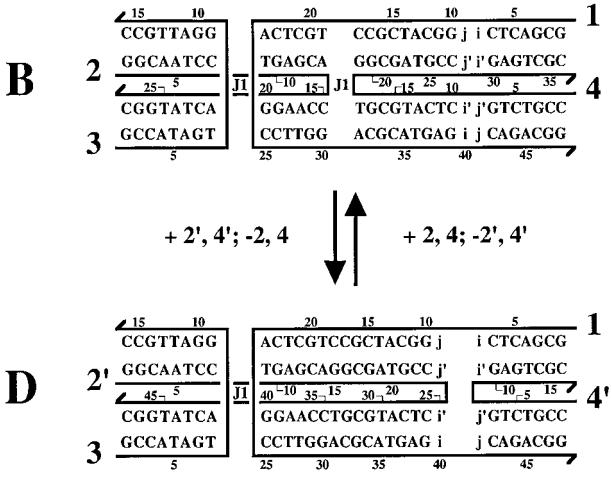

Figure 5. The sequence of the molecules used. The molecules used in this study are illustrated here. Species B and D are the same as in Figure 4, and are shown in the unwrapped representation. Half-arrowheads represent the 3′ ends of strands. Strand numbering is indicated in large numbers and sequence positions are indicated with smaller numbers. The variable positions 5′-i,j-3′ and 5′-j′,i′-3′ are indicated in both molecules. The reaction is indicated in which strands 2 and 4 in B are replaced by 2′ and 4′ to yield D.

Non-association control experiments

The key assumption in this experiment is that the single-strand components 2, 2′, 4, and 4′ neither associate with each other nor enter into complex secondary structures when they have been displaced from complexes B or D. Figure 6(b) illustrates the lack of association between these components (i = A and j = T) at a series of concentrations, ranging from 300 nM to 50 nM. We find no association between the individual strands for any of the sequences in the concentration range between 100 nM and 50 nM. Consequently, we have performed these experiments at a concentration of 88 nM. In addition, we have characterized the thermal denaturation profiles for the individual strands. In order to do this at concentrations similar to those used in the other experiments, we have used perpendicular denaturing gradient gel electrophoresis (Fischer & Lerman, 1979, 1983) to compare the denaturation profile of the molecules used here with an appropriate length of oligo(dT), which we assume to be unstructured. Figure 6(c) illustrates the differences in mobilities of the strands of two of the molecules studied here, relative to oligo(dT). The cooperative melting of the hairpin loop is shown for comparison. Within experimental error, the differences in mobility are consistently linear plots, showing no evidence of secondary structural features; none of the melts shows cooperativity at the concentrations used (Figure 6(c)). These data lead us to conclude that we have designed the sequences in such a fashion that the strands contain no detectable secondary structure, either individually or when mixed with the other three strands. Hence, strands 2, 2′, 4 and 4′ participate in no detectable secondary interaction, and it is valid to assume that their contributions as species A and C (Figure 4) to the free energy of the system are equivalent in each experiment.

Hydroxyl radical analysis

We have used hydroxyl radical autofootprinting previously to characterize unusual DNA molecules, including branched junctions (Churchill *et al.*, 1988; Chen *et al.*, 1988; Wang *et al.*, 1991), antijunctions and mesojunctions (Du *et al.*, 1992; Wang & Seeman, 1995), and double crossovers (Fu

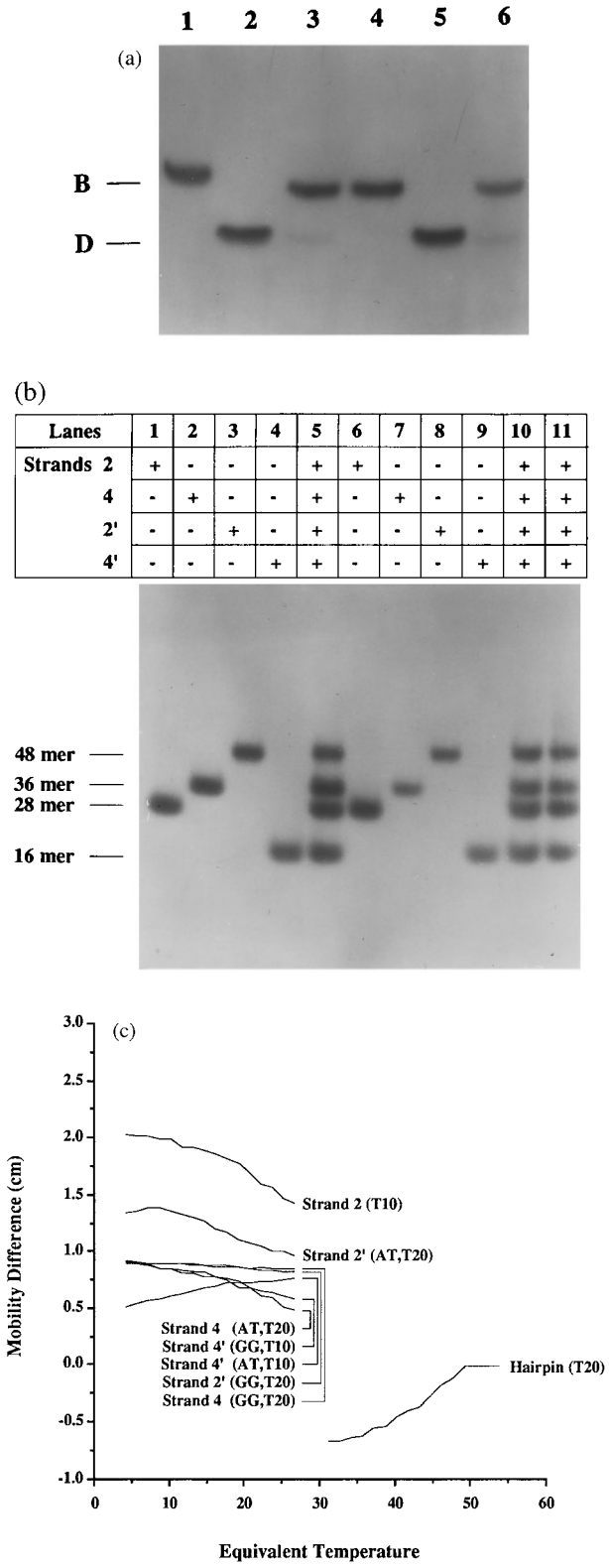

Figure 6. Experiments characterizing this system. (a) A typical competition reaction. This is an autoradiogram of a non-denaturing gel showing the components of the competition reaction. B and D indicate the two double-crossover molecules. Lanes 1 to 3 contain labeled strand 3, and lanes 4 to 6 contain labeled strand 1. Lane 1 contains the four-strand B complex, composed of strands 1, 2, 3 and 4. Lane 2 contains the four-strand D complex, composed of strands 1, 2′, 3 and 4. Lane 3 contains the competition reaction with all six strands; strands 1 and 3 are present in equimolar concentration, 65 nM, as are the concentrations of strands 2, 4, 2′ and 4′, which are 88 nM. Lanes 4 to 6 contain the same pattern. The 5′-i,j-3′ sequence here is 5′-AT-3′. (b) Non-association controls. This autoradiogram of a non-denaturing 8% polyacrylamide gel is designed to illustrate the fact that strands 2, 4, 2′ and 4′ do not associate with each other in solution. The contents of each lane are indicated above it. As in (a), the 5′-i,j-3′ sequence here is 5′-AT-3′. Lanes 1 to 4 contain single strands, and lane 5 contains all four strands. The same pattern is used in lanes 6 to 10, and lane 11 contains all four strands. The strand concentrations in lanes 1 to 5 are 300 nM, in lanes 6 to 10, 100 nM, and in lane 11, 50 nM. It is clear that there is no association visible in the lanes containing multiple species. No band is present on the gel that indicates a dimer or other slower migrating species. No association is seen for any species below 100 nM. (c) Denaturing gradient gel electrophoresis experiments demonstrating the lack of secondary structure in isolated single strands. Each of the plots represents the difference in electrophoretic mobility between one of the strands used in these experiments and a molecule of oligo(dT). Oligo(dT), assumed to be unstructured, has been used as a standard, to control for differential mobility caused by features of the denaturing gradient gel, rather than the secondary structure of the molecule. Each strand is listed, followed by a parenthetical notion indicating both the system (i = A, j = T or i = G, j = G) for which the results are displayed and the molecule with which it is compared. The results are all linear, within experimental error. The cooperative melting in this system of a molecule containing a hairpin loop (5′-TGCCGATCCTTTTTGGATC-3′) is shown as a positive control.

& Seeman, 1993, Zhang *et al.*, 1993; Zhang & Seeman, 1994; Fu *et al.*, 1994; Li *et al.*, 1997). These experiments are performed by labeling a component strand of the complex and exposing it to hydroxyl radicals. The key feature noted in these analyses is decreased susceptibility to attack when comparing the pattern of the strand as part of the complex, relative to the pattern derived from linear duplex DNA. Decreased susceptibility is interpreted to suggest that access by the hydroxyl radicals may be limited by steric factors at the sites where it is detected. Likewise, similarity to the duplex pattern at points of potential flexure is assumed to indicate that the strand has adopted a helical structure in the complex, whether or not it is required by the secondary structure. In previous studies of junctions, double crossovers and meso-junctions, protection has been seen particularly at the crossover sites, but also at non-crossover sites where strands from the two domains appear to occlude each other's surfaces from access by hydroxyl radicals (Churchill *et al.*, 1988; Fu & Seeman, 1993; Fu *et al.*, 1994; Du *et al.*, 1992; Wang & Seeman, 1995). Hence, protection is a reliable indicator of the crossover isomer; strands seen to be protected relative to duplex are the crossover strands of the junction.

Figure 7 illustrates a typical set of hydroxyl radical protection experiments for species B and D (i = T and j = A). For each double-crossover molecule, the protection pattern expected for the species is seen. Relative to the double-helical baseline structure, the two nucleotides forming the crossover-point are protected from attack, whereas the nucleotides on the non-crossover strands show no protection. These patterns confirm the ability of the complexes 1, 2, 3, 4 and 1, 2′, 3, 4′ to form symmetric immobile junctions, with patterns similar to those reported previously for these types of molecules.

Free energies, enthalpies and entropies

Figure 8(a) illustrates the final values of $\Delta^{J1}G^{ij}$ for all 16 nucleotides, at 277 K (4°C), 289 K (16°C) and 295 K (22°C). Each experiment has been repeated five or more times. We have used the measurements at different temperatures to estimate the values of $\Delta^{J1}H^{ij}$ and $T\Delta^{J1}S^{ij}$; these values are illustrated in Figure 8(b) and (c), respectively. They are derived from the fundamental relationship, $\Delta G = \Delta H - T\Delta S$. Plotting ΔG as a function of temperature yields ΔH as the intercept and ΔS as the slope. The

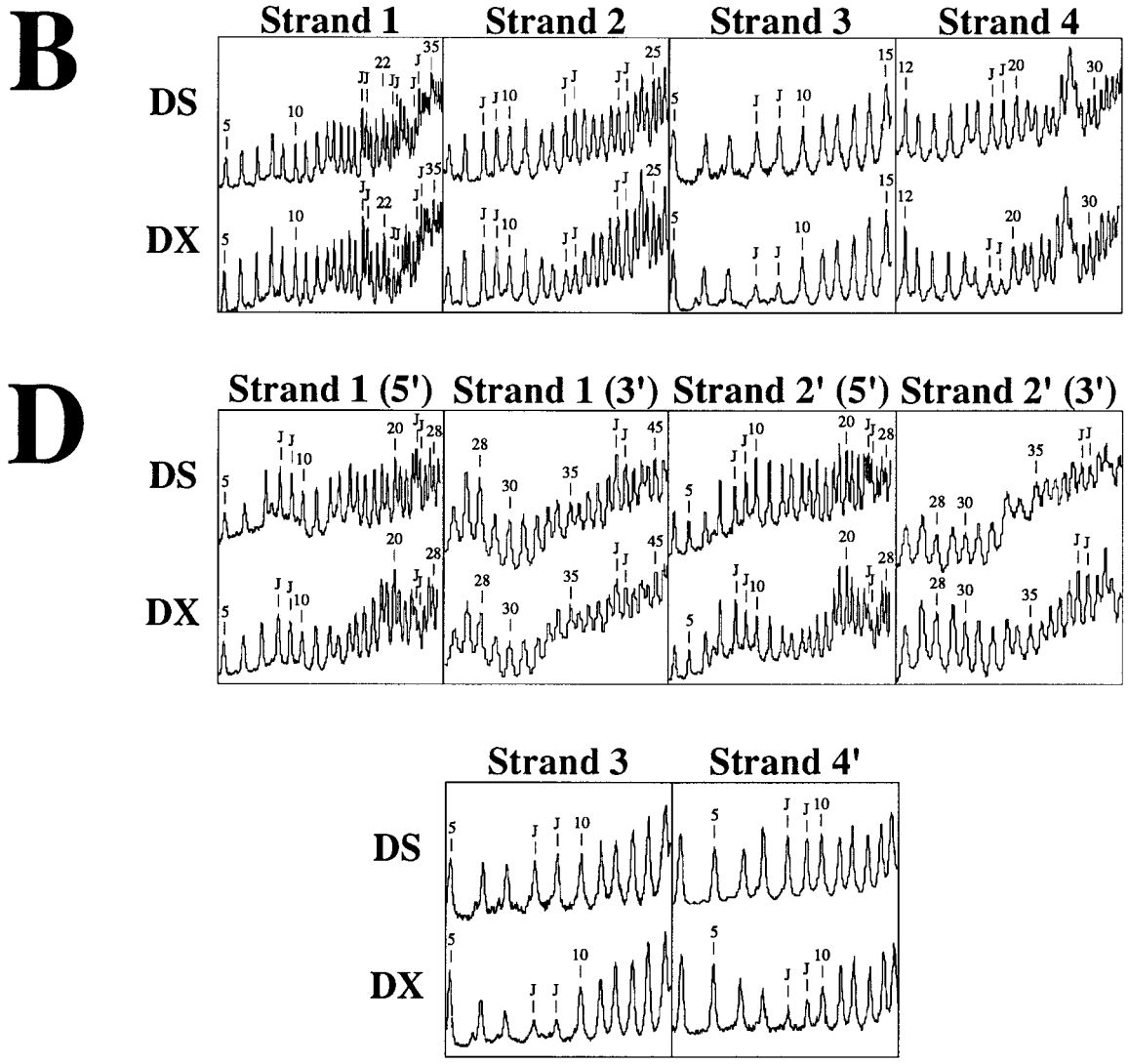

Figure 7. Hydroxyl radical autofootprinting of symmetric immobile junctions. Each of the portions (B or D) shows hydroxyl radical attack patterns for the molecules used in these experiments. Each plot represents quantification through a lane on a sequencing gel; the area of each peak is proportional to cleavage at that site. DX refers to a particular strand in the double-crossover molecule, and DS refers to the same strand when complexed with its Watson-Crick complement. Nucleotide numbers are indicated, and correspond to the numbers in Figure 5. The positions of branch-points are indicated by the symbol J, regardless of whether the position is on a crossover strand or on a helical strand. The region of crossover sites in strands 2′ and 3 in D requires two different gels to visualize clearly; these are labeled as being on the 5′ or 3′ ends of the strands. In each case, protection is seen at the junctions in the crossover strands indicated in Figure 5.

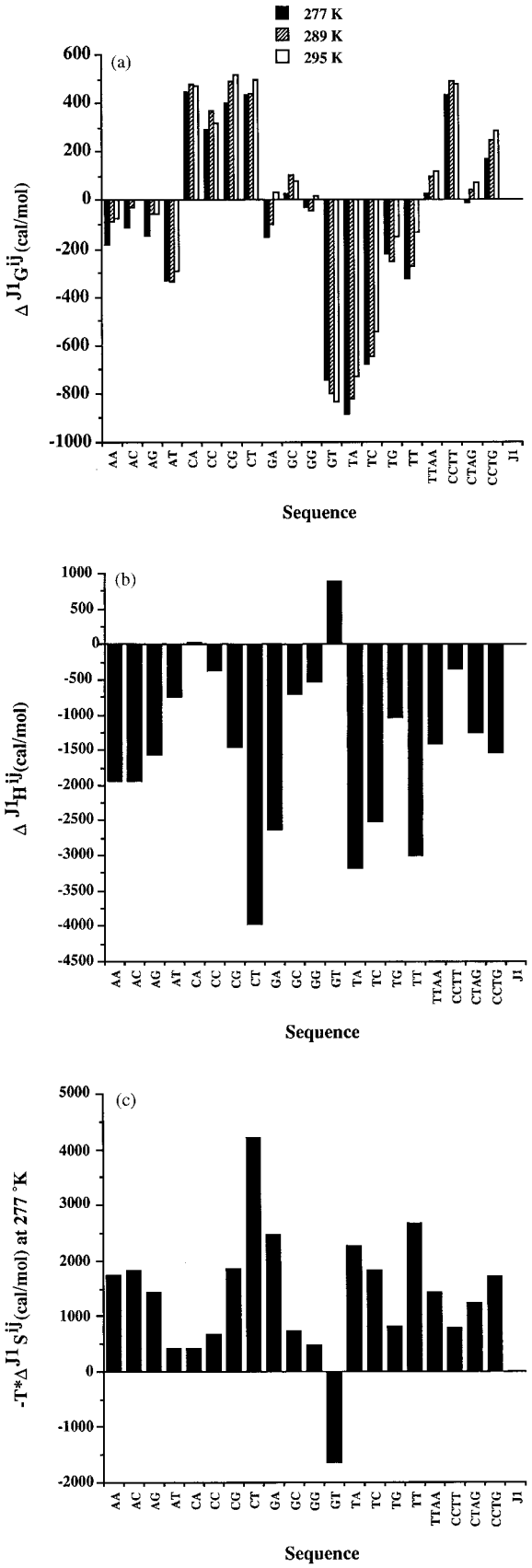

Figure 8. Thermodynamic quantities for sequences flanking symmetric immobile junctions. In each case, the individual sequences corresponding to i and j are shown along the bottom of the plot. At the right are the controls that involve variation of the flanking sequences. (a) Free energies. The quantities $\Delta^{J1}G^{ij}$ are shown in cal/mol. We have made these measurements at three different temperatures, at 277 K (filled bars), at 289 K (shaded bars) and at 295 K (open bars). It is clear that many of the values are similar to each other, but that some differ by as much as 1.3 kcal/mol (TA-CA). (b) Enthalpic contributions to the free energy. The quantities $\Delta^{J1}G^{ij}$ are shown in cal/mol. They have been derived from the dependence of the free energies on temperature. Note that all are negative, with the exceptions of GT and CA, the latter being insignificantly different from zero. (c) Entropic contributions to the free energy. The quantities $-T\Delta^{J1}S^{ij}$ are shown in cal/mol at 277 K. With the exception of GT, all of these quantities are positive. We have checked GT several extra times, and continue to get the same dependence of its free energy on temperature.

errors of these estimates are necessarily rather large.

The 4°C values of $\Delta^{J1}G^{ij}$ can be divided into three groups: Those clearly positive (between 286 and 433 calories), those clearly negative (−680 to −888 calories) and those in the middle (−333 cal to +24 calories). Curiously, the CN sequences all have strongly positive values of $\Delta^{J1}G^{ij}$, and they are the only strongly positive values that we note. The trend among the strongly negative sequences is more ambiguous, being represented by GT, TA and TC.

Our average estimated standard deviations for ΔG^{ij} are 80 cal, 4 to 16% of the measured values, which range from 0.5 to 2 kcal/mol. The derived quantities $\Delta^{J1}G^{ij}$ have larger standard deviations, because they are combined with the errors in the measurement of ΔG^{J1}; these standard deviations average 114 (4°C), 149 (16°C) and 159 calories (22°C). With these errors in mind, a significant difference, say 3σ, between two values of $\Delta^{J1}G^{ij}$ would be about 350 calories. The differences between the clearly negative and the clearly positive values certainly qualify as significantly different. Some values between these two extreme groups and the middle group are also significantly different.

In addition to the measurements described above, we have done two sets of control experiments involving the flanking sequences. In the first case, we have taken two sequences and expanded their symmetry to cover four nucleotides. Thus, we have flanked TA with another T and A at both positions, replacing the asymmetric J1 sequences, and generating the symmetric sequence TTAA. Similarly, we have flanked CT with C and T, so that the symmetric sequence CCTT has been produced. The effect of this change in flanking environment is almost negligible on $\Delta^{J1}G^{ij}$ for CT, but $\Delta^{J1}G^{ij}$ for TA changes dramatically from the clearly negative group in the middle group.

In a second set of controls, we have reversed the flanking sequences of the J1 junction, on both of

these sequences. Thus, the GATC and CATG sequences have been replaced with CATG and GATC, respectively. This alteration again has a major effect on $\Delta^{J1}G^{ij}$ for AT but, this time, $\Delta^{J1}G^{ij}$ for CT is diminished markedly.

Discussion

The sequence-dependent model for branch migration

Our key finding is that the various symmetric dimer sequences that can flank a junction may exhibit markedly different stabilities. Another way to phrase this observation is that the cost of placing a branch-point in a DNA molecule is not sequence-independent in all cases. Consequently, the cost of moving a branch-point within a mobile Holliday junction is not necessarily sequence-independent, and this observation should be taken into account in modeling the kinetics of branch migration. It is clear that measurements of activation barriers (e.g. see Thompson *et al.*, 1976) are really estimates of average activation barriers. Within the environment where we have made our observations, it appears, for example, that a branch point in a 5′-CTA-3′ sequence would be a lot more likely to be found between the T and the A than between the C and the T. Thus, for systems of known sequence, kinetic modeling could eventually be made more precise by taking the thermodynamics associated with the sequence onto account. It is clear that the data we report here are only the first approximation to the numbers that are needed for accurate modeling of branch migratory kinetics. It is clear also that the sequence-independent model is valid for many of the possible sequences, but that it is not universally true.

Caveats of interpretation in this system

The measurements that we have reported here necessarily require that a number of constraints be placed on the system to ensure its stability; these constraints may, in turn, lead to a series of artifacts. It is important that we point them out. The controls that we have run indicate that expansion of the symmetry flanking the branch-point can have a dramatic effect on the stability of a sequence, as seen when the TA sequence is expanded to TTAA. However, this effect is not a general rule, as seen in the comparison of CT and CCTT. Likewise, changing the asymmetric environment of the symmetric sequence used can have a marked effect on the stability of the sequence. Thus, reversal of sequences flanking both TA and CT changes their stabilities significantly. The effect of altering the flanking sequences appears to be so extensive that this study must be regarded as preliminary, at best. Study of symmetric tetrameric or longer sequences will be necessary before the sequence-dependence of branch-point minima can be known accurately.

We attempted initially to perform this study in symmetric immobile junctions whose sequences were constrained to be parallel, rather than antiparallel. We were unable to get this system to provide consistent results, and consequently we switched to the antiparallel system. Little is known about the local structure of the junction during spontaneous branch migration, although it has been suggested that it is tetragonal in enzyme-catalyzed branch migration (Rafferty *et al.*, 1996; Yu *et al.*, 1997). Spontaneous branch migration is modeled most readily within a parallel context (Sigal & Alberts, 1972; Robinson & Seeman, 2987). Distortion of the junction from a tetragonal or parallel conformation to an antiparallel conformation may have led to systematic errors.

Immobile junctions are most stable in the presence of Mg^{2+}, and we have performed our measurements in a solution containing 10 mM Mg^{2+}. Nevertheless, Panyutin *et al.* (1995) have shown that the rate of branch migration is increased by a factor of about 1000 when Mg^{2+} is replaced by Na^{+}. The cellular rate is more closely approximated by the rate seen in solutions containing sodium, rather than magnesium. Thus, we may have made our measurements in a context that overly stabilizes the junction.

Antiparallel symmetric immobile junctions are constrained to have their helix axes coplanar. Nevertheless, unconstrained junctions are known to adopt a conformation in which the helix axes are disposed at an angle of about 60° to each other (Lilley & Clegg, 1993). This constraint may well lead to misestimation of thermodynamic parameters that describe the junction. For example, we measured the thermodynamics of sequence-dependent crossover isomerization, and decided that we had underestimated the values of crossover preference by a factor of at least 3 (Zhang & Seeman, 1994); this suggestion has been borne out in recent measurements by Chazin, Millar and colleagues (Miick *et al.*, 1997).

Features of the observations

Figure 9 shows a plot of ΔH *versus* $-T\Delta S$ at 4°C. It is clear that this system, like many other nucleic systems (e.g. see Petruska *et al.*, 1988), demonstrates enthalpy-entropy compensation (Lumry & Rajender, 1970). The enthalpies are almost all favorable, with the exception of CA and GT but, excepting GT, the entropies are all unfavorable. The free energies that we have measured are, therefore, relatively small numbers that represent the differences of two large numbers. Consequently, it is not surprising that they are often very sensitive to the local environment of the base-pair.

We have reported previously measurements of crossover isomer preferences for the symmetric dimers that flank the branch-point (Zhang & Seeman, 1994). Some of the branch-points studied here correspond to the pairs compared in that study, so it is useful to see whether the two studies

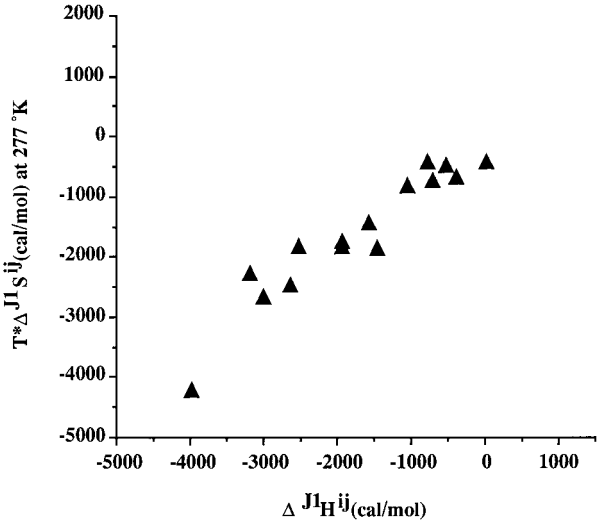

Figure 9. Enthalpy-entropy compensation in this system. This is a plot of the values of $\Delta^{J1}H^{ij}$, plotted on the abscissa, *versus* $-T\Delta^{J1}S^{ij}$, at 277 K, plotted along the ordinate. Only the points corresponding to the dinucleotides are plotted here. The correlation is over 90%. It is clear that this is yet another system in which enthalpy-entropy compensation is operative.

have produced similar results. The $\Delta\Delta G$ values measured previously compare with those measured here as follows, where the sequence given is the one preferred over its complement to be found on the helical strand, previous work, and this work: TT/AA, −634, −139; GG/CC, −442, −311; GT/AC, −4, −624; TG/CA, −144, −666; AG/CT, −128, −577; TC/GA, −588, −524. The sign of the preference agrees in each case, although the GT/AC sign comparison is clearly meaningless, and the estimates show the worst divergence. The consistency we note suggests that the qualitative picture we derive here is similar to that noted before: we suggested that complementary sequences that compete for positions on helical and crossover strands do not have large free energy barriers between then. We see this observation borne out here, because the largest differences in this work are not between dimers with complementary sequences. The extrema of our observations here are CN (positive), and GT, TA and TC (negative). None of the dimer pairs in the opposite extrema is complementary.

Materials and Methods

Synthesis, purification and labeling of DNA

All DNA molecules have been synthesized on an Applied Biosystems 380B automatic DNA synthesizer, removed from the support, and deprotected using routine phosphoramidite procedures (Caruthers, 1985). All strands were purified by polyacrylamide gel electrophoresis. Radioactive ^{32}P-containing phosphate groups were added as described (Li *et al.*, 1997).

Formation of hydrogen-bonded complexes

Complexes were produced by mixing a stoichiometric quantity of each strand, as estimated by A_{260}. The complexes were formed by heating the samples to 90°C and cooling slowly to 4°C in TAEMg buffer (40 mM Tris (pH 7.5 at 25°C), 20 mM acetic acid, 2 mM EDTA, and 12.5 mM magnesium acetate). Exact stoichiometry was determined, if necessary, by titrating pairs of strands designed to hydrogen bond together, and visualizing them by non-denaturing gel electrophoresis; absence of monomer was taken to indicate the endpoint.

Non-denaturing polyacrylamide gel electrophoresis

These gels contain 15% (w/v) acrylamide (19:1 acrylamide to bisacrylamide). DNA was suspended in 30 µl of TAEMg, and 10 µl was loaded onto the gel. The solution was boiled and allowed to cool slowly to 4°C. From 1 µl to 4 µl of non-denaturing tracking dye containing buffer, 50% (v/v) glycerol and 0.02% (w/v) each of bromophenol blue and xylene cyanol FF tracking dyes were added to the samples prior to loading onto gels. Gels were run on a Hoefer SE-600 gel electrophoresis unit at 11 V/cm at 4°C, dried onto Whatman 3MM paper and exposed to X-ray film for up to 15 hours. Autoradiograms were analyzed on a BioRad GS-525 Molecular Imager.

Thermal denaturation profiles by perpendicular denaturing gradient gel electrophoresis

With 100% denaturing conditions corresponding to 6.7 M urea and 40% formamide, 10% polyacrylamide gels in TAEMg buffer were prepared with five different denaturant concentrations, 0, 21, 42, 63 and 84%; 84% is the highest possible concentration, because higher concentrations precipitate at 4°C, where the gels were to be run. A volume (3.6 ml) of each solution was added to five different tubes, to which polymerization catalysts (ammonium persulfate and TEMED) were then added. The solutions were withdrawn from the tubes into the same 25 ml serological pipette (from low concentration to high concentration, because of the different densities). Mixing was accomplished by introducing several air bubbles slowly at the bottom of the pipette. Mixing was monitored by the addition of a small amount of tracking dye to tubes 2 and 4. The gel was then cast by adding the contents of the pipette to the space between glass plates. After polymerization, the gel was turned 90°, and wells were prepared with the gel mixture lacking denaturant. After polymerization of the wells, the DNA sample (the experimental strand and a dT_{10} or dT_{20} marker) was loaded, and the gel run in TAEMg buffer, at 15 V/cm. The gel was exposed to X-ray film for up to 15 hours.

Hydroxyl radical analysis

Individual strands of the complexes were radioactively labeled, and additionally gel-purified from a 10% to 20% (w/v) denaturing polyacrylamide sequencing gel. Each of the labeled strands (approximately 1 pmol in 10 µl of TAEMg) was complexed with its linear duplex complement or double crossover-forming complement, left untreated as a control, or treated with sequencing reagents (Maxam & Gilbert, 1977) for sizing ladder. Hydroxyl radical cleavage of the double-strand,

and double-crossover-complex samples for all strands took place at 4°C for one minute and 40 seconds (Tullius & Dombroski, 1985), with modifications noted by Churchill *et al*. (1988); the final concentrations were 1 mM, L-ascorbic acid, 0.1 mM Fe(II)EDTA^{2-}, 4.4 mM H_2O_2. The reaction was stopped by addition of thiourea to 10 mM. The sample was precipitated with ethanol, dried, dissolved in a formamide-dye mixture, and loaded directly onto a 10% to 20% polyacrylamide/8.3 M urea sequencing gel. Autoradiograms were analyzed on a BioRad GS-525 Molecular Imager.

Acknowledgments

This research has been supported by grants GM-29554 from the National Institute of General Medical Sciences, N00014-98-1-0093 from the Office of Naval Research, and NSF-CCR-97-25021 from the National Science Foundation. We thank Dr Albert S. Benight for valuable suggestions concerning this work.

References

Caruthers, M. H. (1985). Gene synthesis machines: DNA chemistry and its uses. *Science,* **230**, 281–285.

Chen, J.-H., Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R. & Seeman, N. C. (1988). Construction and analysis of monomobile DNA junctions. *Biochemistry,* **85**, 6032–6038.

Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R. & Seeman, N. C. (1988). A Holliday recombination intermediate is twofold symmetric. *Proc. Natl Acad. Sci. USA,* **85**, 4653–4656.

DasGupta, C., Wu, A. M., Kahn, R., Cunningham, R. P. & Radding, C. M. (1981). Concerted strand exchange and formation of Holliday structures by *E. coli* RecA protein. *Cell,* **25**, 507–516.

Du, S. M., Zhang, S. & Seeman, N. C. (1992). DNA junctions, antijunctions and mesojunctions. *Biochemistry,* **31**, 10955–10963.

Fischer, S. G. & Lerman, L. S. (1979). Length-independent separation of restriction fragments in two-dimensional gel electrophoresis. *Cell,* **16**, 191–200.

Fischer, S. G. & Lerman, L. S. (1983). DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels. *Proc. Natl Acad. Sci. USA,* **80**, 1579–1583.

Fu, T.-J. & Seeman, N. C. (1993). DNA double crossover molecules. *Biochemistry,* **32**, 3211–3220.

Fu, T.-J., Kemper, B. & Seeman, N. C. (1994). Endonuclease VII cleavage of DNA double crossover molecules. *Biochemistry,* **33**, 3896–3905.

Gellert, M., O'Dea, M. H. & Mizuuchi, K. (1983). Slow cruciform transitions in palindromic DNA. *Proc. Natl Acad. Sci. USA,* **80**, 5545–5549.

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

Hoess, R., Wierzbicki, A. & Abremski, K. (1987). Characterization of intermediates in site-specific recombination. *Proc. Natl Acad. Sci. USA,* **84**, 6840–6844.

Hsieh, P. & Panyutin, I. G. (1995). DNA branch migration. In *Nucleic Acids and Molecular Biology* (Eckstein, F. & Lilley, D. M. J., eds), vol. 9, pp. 42–65, Springer-Verlag, Berlin.

Johnson, R. D. & Symington, L. S. (1993). Crossed-stranded DNA structures for investigating the molecular dynamics of the Holliday junction. *J. Mol. Biol.* **229**, 812–820.

Kallenbach, N. R., Ma, R.-I. & Seeman, N. C. (1983). An immobile nucleic acid junction constructed from oligonucleotides. *Nature,* **305**, 829–831.

Kitts, P. A. & Nash, H. A. (1987). Homology dependent interactions in phage λ site-specific recombination. *Nature,* **329**, 346–348.

Li, X., Wang, H. & Seeman, N. C. (1997). Direct evidence for Holliday junction crossover isomerization. *Biochemistry,* **36**, 4240–4247.

Lilley, D. M. J. & Clegg, R. M. (1993). The structure of the four-way junction in DNA. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 299–328.

Lloyd, R. G. & Sharples, G. J. (1993). Dissociation of synthetic Holliday junctions by *E. coli* RecG protein. *EMBO J.* **12**, 17–22.

Lu, M., Guo, Q., Seeman, N. C. & Kallenbach, N. R. (1991). Parallel and antiparallel Holliday junctions differ in structure and stability. *J. Mol. Biol.* **221**, 1419–1432.

Lu, M., Guo, Q., Marky, L. A., Seeman, N. C. & Kallenbach, N. R. (1992). Thermodynamics of DNA chain branching. *J. Mol. Biol.* **223**, 781–789.

Lumry, R. & Rajender, S. (1970). Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: a ubiquitous property of water. *Biopolymers,* **9**, 1125–1227.

Maxam, A. M. & Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.

Miick, S. M., Fee, R. S., Millar, D. P. & Chazin, W. J. (1997). Crossover isomer bias is the primary sequence-dependent property of immobilized Holliday junctions. *Proc. Natl Acad. Sci. USA,* **94**, 9080–9084.

Mueller, J. E., Kemper, B., Cunningham, R. P., Kallenbach, N. R. & Seeman, N. C. (1988). T4 endonuclease VII cleaves the crossover strands of Holliday junction analogs. *Proc. Natl Acad. Sci. USA,* **85**, 9441–9445.

Nunes-Duby, S. E., Matsumoto, L. & Landy, A. (1987). Site-specific recombination intermediates trapped with suicide substrates. *Cell,* **50**, 779–788.

Panyutin, I. G. & Hsieh, P. (1994). The kinetics of spontaneous DNA branch migration. *Proc. Natl Acad. Sci. USA,* **91**, 2021–2025.

Panyutin, I. G., Biswas, I. & Hsieh, P. (1995). A pivotal role for the structure of the Holliday junction in DNA branch migration. *EMBO J.* **14**, 1819–1826.

Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C. & Tinoco, I., Jr (1988). Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proc. Natl Acad. Sci. USA,* **85**, 6252–6256.

Rafferty, J. B., Sedilnikova, S. E., Hargreaves, D., Artymiuk, P. J., Baker, P. J., Sharples, G. J., Mahdi, A. A., Lloyd, R. G. & Rice, D. W. (1996). Crystal structure of DNA recombination protein RuvA and a model for its binding to the Holliday junction. *Science,* **274**, 415–421.

Robinson, B. H. & Seeman, N. C. (1987). Simulation of double-stranded branch point migration. *Biophys. J.* **51**, 611–616.

Seeman, N. C. (1982). Nucleic acid junctions and lattices. *J. Theor. Biol.* **99**, 237–247.

Seeman, N. C. (1990). *De novo* design of sequences for nucleic acid structural engineering. *J. Biomol. Struct. Dynam.* **8**, 573–581.

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

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

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

Thompson, B. J., Camien, M. N. & Warner, R. C. (1976). Kinetics of branch migration in double-stranded DNA. *Proc. Natl Acad. Sci. USA*, **73**, 2299–2303.

Tsaneva, I. R., Müller, B. & West, S. C. (1993). RuvA and RuvB proteins of *E. coli* exhibit DNA helicase activity *in vitro*. *Proc. Natl Acad. Sci. USA*, **90**, 1315–1319.

Tullius, T. D. & Dombroski, B. A. (1985). Iron(II)EDTA used to measure the helical twist along any DNA molecule. *Science*, **230**, 679–681.

Wang, H. & Seeman, N. C. (1995). Structural domains of DNA mesojunctions. *Biochemistry*, **34**, 920–929.

Wang, Y., Mueller, J. E., Kemper, B. & Seeman, N. C. (1991). The assembly and characterization of 5-arm and 6-arm DNA branched junctions. *Biochemistry*, **30**, 5667–5674.

Warner, R. C., Fishel, R. A. & Wheeler, F. C. (1979). Branch migration in recombination. *Cold Spring Harbor Symp. Quant. Biol.* **43**, 957–968.

Yu, X., West, S. C. & Egelman, E. H. (1997). Structure and subunit composition of the RuvAB-Holliday junction complex. *J. Mol. Biol.* **266**, 217–222.

Zhang, S. & Seeman, N. C. (1994). Symmetric Holliday junction crossover isomers. *J. Mol. Biol.* 658–668.

Zhang, S., Fu, T.-J. & Seeman, N. C. (1993). Construction of symmetric, immobile DNA branched junctions. *Biochemistry*, **32**, 8062–8067.

Edited by I. Tinoco

(Received 24 April 1998; accepted 10 June 1998)