

# Branch Migration Through DNA Sequence Heterology

Indranil Biswas, Akira Yamamoto and Peggy Hsieh\*

Genetics and Biochemistry  
National Institute of Diabetes  
and Digestive and Kidney  
Diseases, National Institutes of  
Health, Bethesda  
MD 20892-1810, USA

Branch migration of a DNA Holliday junction is a key step in genetic recombination. Previously, it was shown that a single base-pair heterology between two otherwise identical DNA sequences is a substantial barrier to passage of a Holliday junction during spontaneous branch migration. Here, we exploit this inhibitory effect of sequence heterology to estimate the step size of branch migration. We also devise a simulation of branch migration through mismatched base-pairs to arrive at the underlying molecular basis for the block to branch migration imposed by sequence heterology. Based on the observation that two adjacent sequence heterologies exert their effects on branch migration more or less independently, we conclude that the step size of branch migration is quite small, of the order of one or two base-pairs per migratory step. Comparison of branch migration experiments through a single base-pair heterology with simulations of a random walk through sequence heterology suggests that the inhibition of branch migration is largely attributable to a thermodynamic barrier arising from the formation of unpaired or mispaired bases in heteroduplex DNAs.

© 1998 Academic Press Limited

**Keywords:** recombination; Holliday junction; branch migration; mismatches

\*Corresponding author

## Introduction

General genetic recombination and the recombinational repair of DNA damage are highly conserved processes that play essential roles in regulating the diversity and integrity of the genome. A central intermediate in genetic recombination is the Holliday junction, the structure that marks the point of exchange between two recombining DNA duplexes. When the Holliday junction is flanked by sequence homology, branch migration can occur involving the stepwise breakage and reformation of hydrogen bonds as one DNA strand is exchanged for another. In this way, branch migration influences the amount of genetic information transferred from one DNA molecule to another during recombination.

We have studied the kinetics of spontaneous or uncatalyzed branch migration *in vitro* using DNA substrates in which annealing of two duplexes *via* complementary single-strand tails generates four-strand intermediates having a Holliday junction that commences spontaneous branch migration from a defined point (reviewed by Hsieh & Panyutin, 1995). Based on our findings, we conclude that spontaneous branch migration is not a viable mechanism for heteroduplex DNA formation during recombination. First, since homologous recombination involves the exchange of

information between similar, but not necessarily identical DNA sequences, branch migration must be able to accommodate sequence homology. We have shown that sequence heterology as small as a single base-pair is sufficient to impose a significant block to spontaneous branch migration *in vitro* (Panyutin & Hsieh, 1993). Second, spontaneous branch migration under near physiological conditions, i.e. in the presence of free  $Mg^{2+}$ , is much too slow to account for heteroduplex DNA formation within biologically relevant time-scales. This is due to the random walk nature of spontaneous branch migration, in which the junction has an equal probability to move forwards or backwards at each step as well as an intrinsically slow step-rate of several steps per second at 37°C (Panyutin & Hsieh, 1994). Similar rates of spontaneous branch migration have been attained using a variety of approaches (Fujitani & Kobayashi, 1995; Kirby *et al.*, 1997; Muller *et al.*, 1992; Mulrooney *et al.*, 1996).

The slow rate-step for spontaneous branch migration in the presence of  $Mg^{2+}$  is attributable to the pivotal role played by the structure of the crossover point in determining the rate of branch migration. In the presence of divalent metal ions like  $Mg^{2+}$ , the Holliday junction assumes a conformation in which base stacking is retained through

the crossover point (reviewed by Lilley & Clegg, 1993; Seeman & Kallenbach, 1994). Base stacking in the Holliday junction is a barrier to movement, and disruption of base stacking in the junction can restore rapid rates of branch migration (Panyutin *et al.*, 1995).

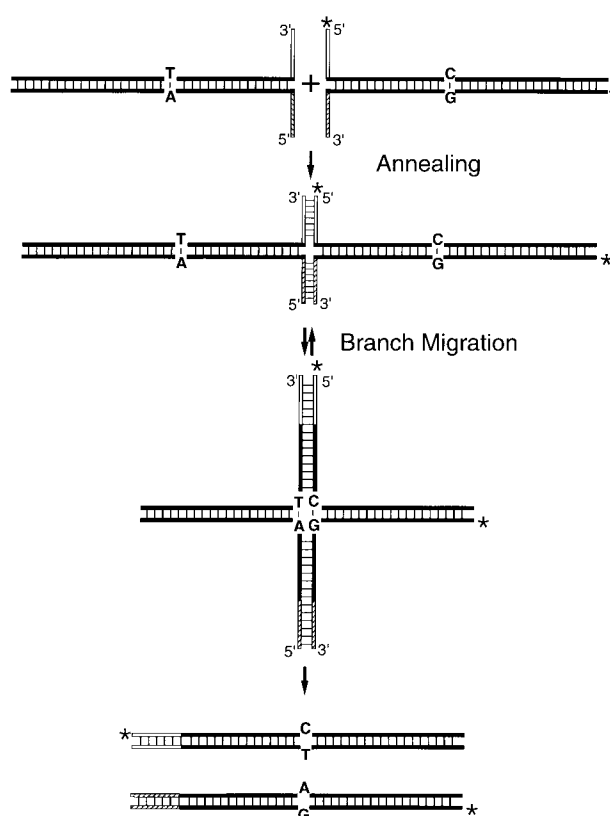
In *Escherichia coli*, branch migration is promoted by the RuvAB and RecG proteins (reviewed by Adams & West, 1995; Lloyd & Low, 1996; Shinagawa & Iwasaki, 1996). The RuvAB protein complex is targeted to Holliday junctions *via* RuvA-junction interactions, while RuvB, and ATP-dependent DNA helicase, provides the motive force for branch migration. When RuvA protein binds to the Holliday junction, it unfolds the junction and removes base stacking at the crossover point, thereby alleviating a conformational constraint to branch migration (Parsons *et al.*, 1995a). Rapid, unidirectional branch migration is achieved through the opposing action of two RuvB hexameric rings each tethered to RuvA that act as molecular motors pulling DNA into the junction and extruding heteroduplex DNA through the RuvB rings (reviewed by West, 1996).

Despite recent advances in our understanding of the parameters that affect branch migration of Holliday junctions, our knowledge of the molecular transactions involved in Holliday junction movement remains incomplete. Here, we exploit the inhibitory effect of base-pair mismatches on spontaneous branch migration to ask questions about the step size of branch migration and the underlying molecular basis for the barrier imposed by sequence heterology.

## Results

### Branch migration through a single mismatch

Previously, we demonstrated that a single base-pair heterology was sufficient to block spontaneous branch migration using branch migration substrates derived from circular, single-strand M13 phage DNA (Panyutin & Hsieh, 1993). Due to secondary structure in the phage DNA that affected both the annealing and branch migration steps, a more extensive study of branch migration through sequence heterology was not possible. Here, we examine the effect of sequence heterology on branch migration using substrates designed to study the kinetics of spontaneous branch migration (Figure 1; Panyutin *et al.*, 1995; Panyutin & Hsieh, 1994). Each substrate contains a 61 bp homologous duplex region and two heterologous, single-strand tails each 20 nt long. Annealing of the two duplex substrates (one of which is  $^{32}\text{P}$ -labeled) *via* the complementary single-strand tails results in the formation of a four-stranded Holliday junction that can initiate spontaneous branch migration from a defined point. When the Holliday junction reaches the distal end of the duplex region, spontaneous branch migration terminates in the irreversible dissociation of the Holliday junction intermediate into



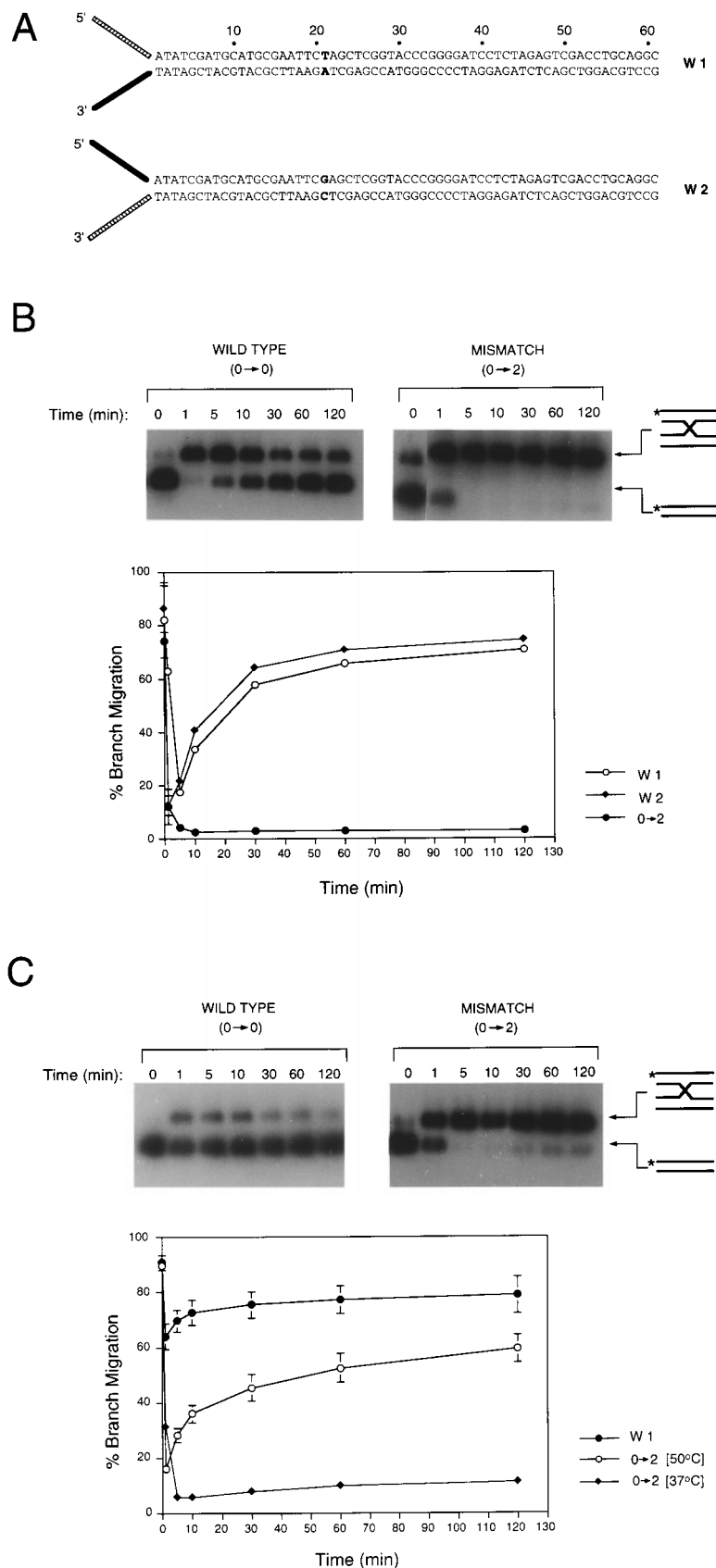
**Figure 1.** Scheme for spontaneous branch migration through a single base-pair heterology. Two branch migration substrates each having heterologous single-strand tails of 20 bases (white and hatched) and identical duplex regions with the exception of one base-pair as indicated are rapidly annealed to form an intermediate containing a Holliday junction at a defined position. The junction commences branch migration in a random walk. The products of complete branch migration are two heteroduplexes each having a mismatched base-pair. The asterisk (\*) denotes a  $^{32}\text{P}$  label.

two heteroduplex products. The formation of the junction intermediate and its conversion to heteroduplex products can be followed by gel electrophoresis.

Branch migration through a single base-pair heterology was examined in the presence or absence of  $\text{Mg}^{2+}$  using two branch migration substrates differing at position 21, one has a G:C base-pair while the other has an A:T base-pair (Figure 2A). Branch migration assays were initiated by mixing the two substrates on ice (corresponding to time zero in Figure 2B) followed by incubation at the indicated temperature and time. Annealing of the substrates to form the Holliday junction intermediate was completed in two to five minutes. Branch migration through a single base-pair heterology was essentially blocked in the presence of  $\text{Mg}^{2+}$  at  $37^\circ\text{C}$  (Figure 2B). Under these conditions, the intrinsic rate of branch migration is quite slow, of the order of several steps per minute (Panyutin & Hsieh, 1994). In the absence of  $\text{Mg}^{2+}$  (Figure 2C),

branch migration through a single base-pair heterology at 37°C was still blocked, despite an almost thousandfold increase in the rate of branch migration. Demonstration that sequence heterology

does not constitute an absolute block to branch migration was achieved by raising the temperature of the reaction to 50°C. As we have previously shown, the rate of spontaneous branch migration



**Figure 2.** Branch migration through a single base-pair heterology. A, Two branch migration substrates, W1 and W2, having a single base-pair heterology at position 21 are denoted in bold. The heterologous, single-stranded tails are indicated by black and stippled lines at the left end of each substrate. B, (top) Branch migration assays with substrates for wild-type 0 → 0 (left) or mismatch 0 → 2 (right) at 37°C in Tris-EDTA containing 10 mM MgCl<sub>2</sub>. (bottom) Time-courses of branch migration for two homologous sequences, W1 or W2, and the mismatch 0 → 2. C, (top) Branch migration assays with substrates for wild-type 0 → 0 (left) or mismatch 0 → 2 (right) at 50°C in Tris-EDTA containing 100 mM NaCl. (bottom) Time-courses of branch migration for the homologous sequence W1 at 50°C and the mismatch 0 → 2 at 37 or 50°C in Tris-EDTA plus 100 mM NaCl. Data are the average of at least two independent experiments. Quantification is of the bottom band corresponding to substrate at time zero and to duplex products at later time-points.

is accelerated approximately tenfold for each 10°C increase in temperature (Panyutin & Hsieh, 1994). Thus, branch migration of a Holliday junction through 60 bp containing a single base-pair heterology within a time-frame of a couple of hours required a 10,000-fold increase in the intrinsic rate of branch migration.

Control experiments using completely homologous duplex regions (either W1 or W2 sequences) follow the expected kinetics for spontaneous branch migration. The step time for branch migration derived from an examination of the data shown in Figure 2B and C, and a computer simulation of a one-dimensional random walk is approximately 180 to 200 ms at 37°C in the presence of 10 mM Mg<sup>2+</sup>, in excellent agreement with previously determined rates of spontaneous branch migration under these conditions (Panyutin *et al.*, 1995; Panyutin & Hsieh, 1994).

There are several experimental limitations in our study of branch migration through sequence heterology. First, as we observed previously (Panyutin *et al.*, 1995) and as is apparent here, the extent of branch migration reactions using synthetic deoxyoligonucleotide substrates saturates between 60% and 80%. We attribute this to the use of relatively long (80 nt) synthetic DNA substrates and the requirement for the annealing of two strands to create heterologous duplex substrates, since the extent of branch migration reactions is routinely in excess of 95% in the case of substrates generated from restriction endonuclease fragments (Panyutin *et al.*, 1995; Panyutin & Hsieh, 1994). Presumably, a subpopulation of the synthetic deoxyoligonucleotides bear modifications such that they are unable to complete the branch migration reaction even after lengthy incubations. In addition, imperfect annealing of two strands may, in some cases, result in barriers to branch migration. Second, the positioning of sequence heterology is limited to the middle of the duplex regions. Placement of heterologies near the origin masks their effect on branch migration for the reasons described below. In a similar vein, the effect of mismatches is also masked when heterologies are positioned near the distal end of the duplex. At temperatures of 37°C or higher, the last 10 bp or so of the duplexes do not participate in branch migration due to their inherent instability. Third, the large block to branch migration posed by mismatches requires that assays be conducted under conditions of extremely rapid branch migration, i.e. in the absence of Mg<sup>2+</sup> and at elevated temperatures.

### The effect of base-pairing on the rate of branch migration through sequence heterology

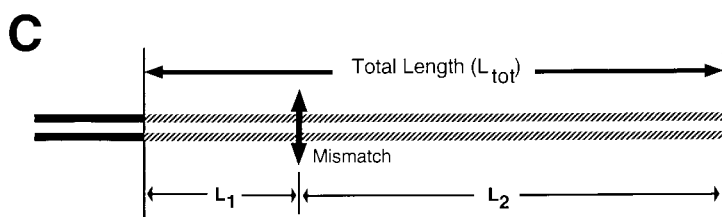
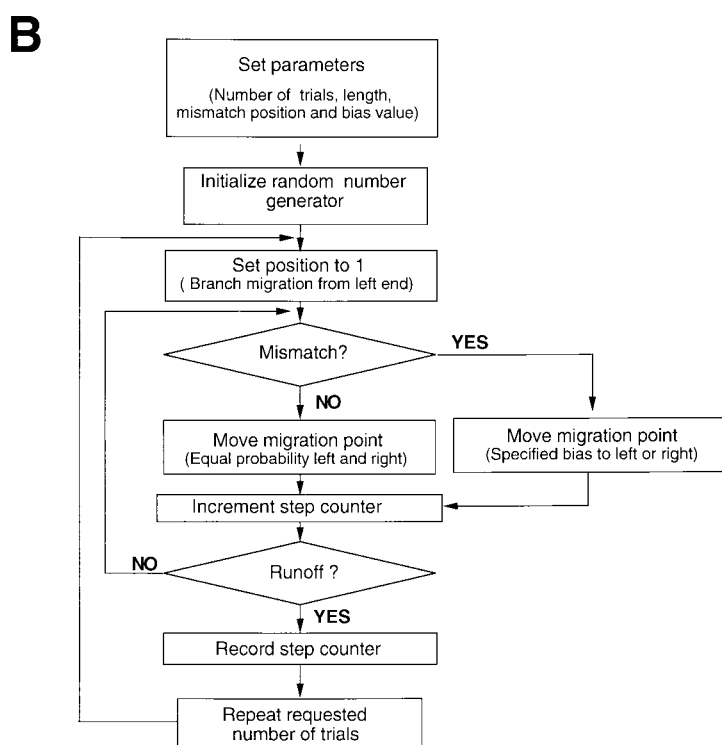
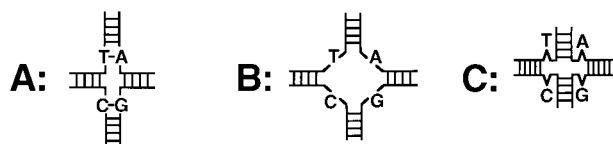
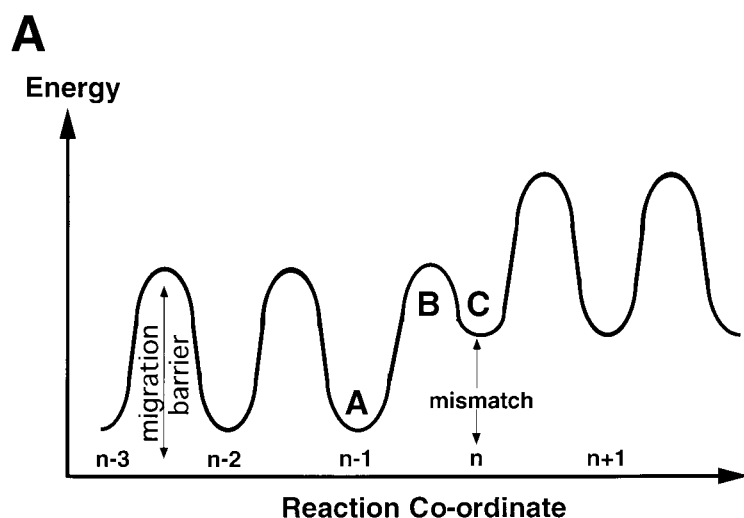
What is the underlying basis for the very large barrier to spontaneous branch migration imposed by short sequence heterologies? One scenario first advanced by Robinson & Seeman (1987) is that the barrier to branch migration can be attributed primarily to the loss of base-pairing in the heterodu-

plex DNAs formed by branch migration through sequence heterology (see Figure 3A). In this reaction scheme, it is assumed that transition state intermediates at each step of branch migration are identical, or nearly so, and that the step-time for branch migration is the same at each step of the random walk with one notable exception. If a single base-pair heterology resides at position  $n$ , then migration of the junction from  $n - 1$  to  $n$  (corresponding to A and C, respectively, in Figure 3A) will result in the conversion of an intermediate having two completely base-paired duplexes to one having two heteroduplexes each with two mismatched bases. These mismatched bases destabilize the heteroduplexes due to loss of base-pairing and stacking resulting in a higher ground state for junctions having mismatches as opposed to those with perfectly paired duplex arms (denoted by the vertical "mismatch" arrow). Correspondingly, the step-time for branch migration is shorter going from  $n$  to  $n - 1$  (i.e. traveling to the left from  $n$  in the drawing) compared to going from  $n - 1$  to  $n$ .

A second explanation for the barrier is that the presence of mismatches in some way destabilizes the "open" intermediate (B in Figure 3A), such that the activation energy barrier in the step going from  $n - 1$  to  $n$  is substantially raised. Unfortunately, in the absence of more information on the rate-limiting step in spontaneous branch migration and the nature of the intermediate corresponding to the "open" junction (see Hsieh & Panyutin, 1995), it is difficult to assess the effect of sequence heterology on the activation barrier. However, we can assess to what extent the free energy difference,  $\Delta\Delta G^\circ$ , between fully base-paired duplexes and duplexes having mismatched base-pairs, accounts for the block to branch migration posed by sequence heterology.

We compared a computer simulation of branch migration through sequence heterology with the experimentally determined time-course for branch migration at 50°C in 100 mM NaCl through a single base-pair heterology shown in Figure 2B. The computer simulation is very similar in most aspects to the simulation used previously to measure the kinetics of spontaneous branch migration (Panyutin *et al.*, 1995; Panyutin & Hsieh, 1994) and models branch migration as a one-dimensional random walk with a reflecting barrier at the origin (left) corresponding to the heterologous "tail" region and an absorbing barrier at the right representing runoff or complete branch migration leading to the irreversible formation of two heteroduplex products. The simulation differs from its predecessor in one significant aspect, however, in that it allows one to introduce biases in the probability to move forwards (towards the right as drawn) or backwards (to the left) at any given position as a consequence of sequence heterology (see Figure 3B and C).

The computer simulation was modeled after the substrates used in branch migration assays and counts the number of migratory steps required to



**Figure 3.** Simulation of spontaneous branch migration through sequence heterology. A, Schematic reaction coordinate for branch migration (adapted from Robinson & Seeman, 1987). A, B and C represent, respectively, the junction at  $n - 1$  with base-pairing intact, the "open" intermediate and the junction at  $n$  corresponding to the formation of two mismatched base-pairs. B, scheme of the computer simulation of branch migration through sequence heterology. C, The position of a mismatch heterology in duplex regions (see the text).



achieve runoff (RO). The total length of branch migration was set to 60 bp and a mismatch was positioned at  $n = 21$ . A bias in the direction of movement of the junction was introduced at the position of the mismatch with the magnitude of the bias being varied from 1:1; that is, a random walk with equal probability to move left or right, to a bias of  $22.5 \times 10^3:1$  to move to the left as opposed to the right. Each simulation carries out 5000 trials and registers the number of branch migration steps required to achieve runoff for each attempt. The fraction of trials achieving runoff after the number of steps corresponding to either five or ten minutes of branch migration is shown in Figure 4. Since the step-time for branch migration at 50°C in 100 mM NaCl is between 30 and 40  $\mu$ s (Panyutin & Hsieh, 1994), five minutes of branch migration under these conditions corresponds to  $0.75 \times 10^7$  to  $1.0 \times 10^7$  steps, whereas ten minutes of branch migration corresponds to  $1.5 \times 10^7$  to  $2.0 \times 10^7$  steps. The fraction of Holliday junctions that have completed branch migration after five or ten minutes at 50°C in 100 mM NaCl is 48% and 60%, respectively (see Figure 2C). The extents of the reactions have been normalized relative to the homologous substrates in Figure 2B. The intersections of the experimental data and the branch migration simulations occur at a bias of approximately  $1 \times 10^4$  for movement of the junction from  $n$  to  $n - 1$  versus  $n - 1$  to  $n$ . As we discuss below, a bias of this magnitude strongly supports the notion that the loss of base-pairing is largely responsible for the barrier of branch migration imposed by sequence heterology.

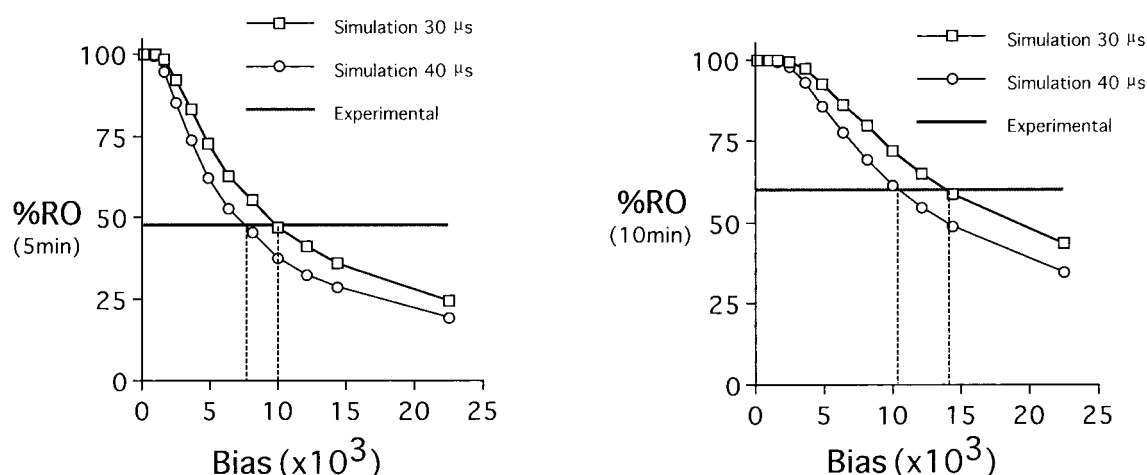
The distribution of crossover points in the case of a single mismatch was determined in the computer simulation. The location of the junction was determined after  $1 \times 10^5$  steps, a value that constitutes only 1% of the median number of steps required for RO at a bias of  $10^4$  at the mismatch

position. The distribution of junctions along the duplex was determined after 1000 such trials. For values of the bias at a mismatch between  $8.1 \times 10^3$  and  $1.2 \times 10^4$ , >99% of branch migration steps were consumed in a random walk in the region between the origin and the mismatch corresponding to region  $L_1$  in Figure 3C. From these results, the number of junctions that reside at the mismatch is proportional to  $1/(L_1 \times \text{bias})$ . Thus, the inhibitory effect of a mismatch is dependent on the position of the mismatch with respect to the origin, i.e. the length of  $L_1$ .

### Estimation of the step-size: branch migration through repair:mismatch heteroduplexes

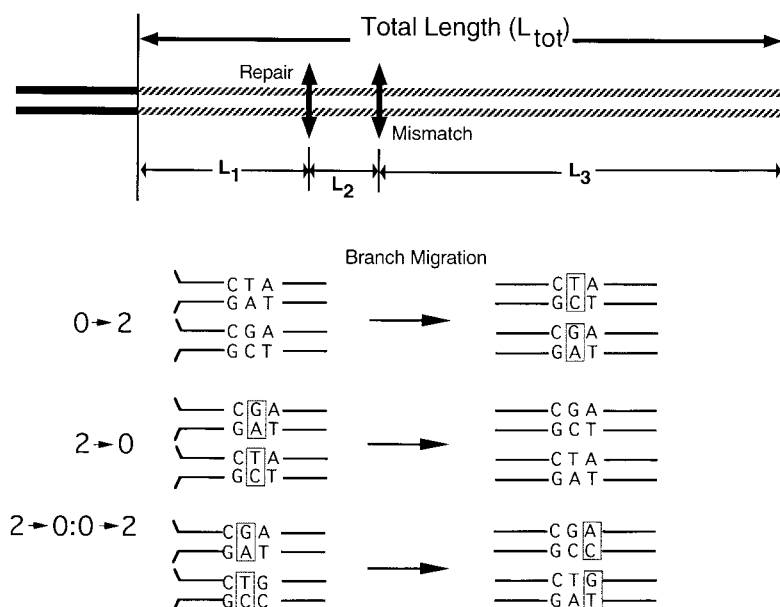
How many base-pairs does the Holliday junction traverse in a single step? To estimate the step-size of branch migration, we carried out branch migration assays using substrates in which two single base-pair heterologies were separated by varying distances. We reasoned that the step-size of migration should constitute the shortest distance between two heterologies that results in their being treated as independent events.

The findings concerning the block to spontaneous branch migration imposed by sequence heterology indicate that when a Holliday junction encounters sequence heterology, the equilibrium is shifted to favor branch migration in the direction that preserves base-pairing, i.e. away from the mismatch. A prediction is that branch migration through a pre-existing mismatch to restore perfect base-pairing, a process referred to here as repair ( $2 \rightarrow 0$ ), should proceed quite readily. In fact, branch migration in the case of repair would be expected to be somewhat faster than in the case of wild-type substrates, because once the Holliday junction moves through the mismatches in the case of repair, migration in the reverse direction is

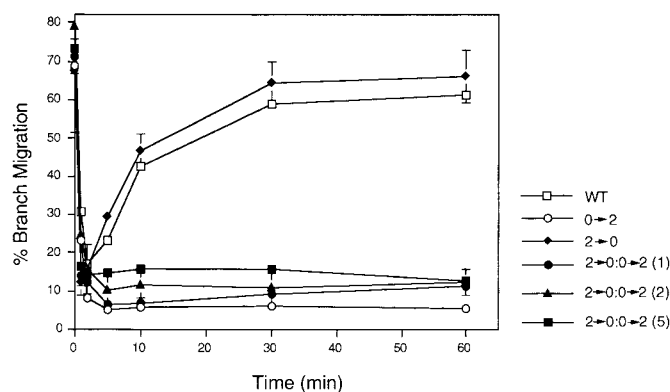


**Figure 4.** Determination of the magnitude of the barrier to spontaneous branch migration imposed by mismatches. The percentage RO for simulated branch migration through sequence heterology is plotted as a function of the size of the bias at the position of the mismatch. The continuous line corresponds to the experimentally determined fraction of RO after five minutes (left) or ten minutes (right) of branch migration.

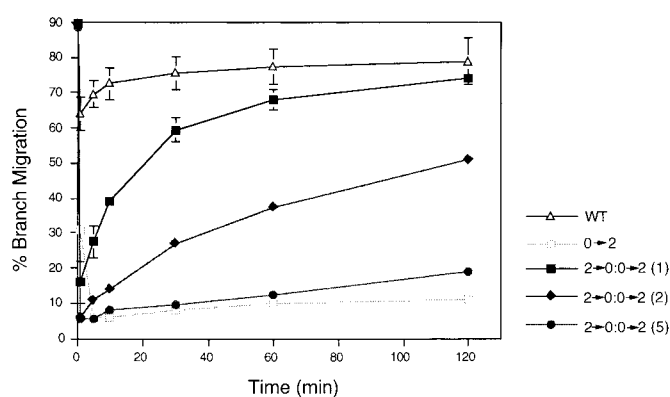
A



B



C



**Figure 5.** Branch migration through repair:mismatch heterologies. A, Scheme for branch migration with mismatch (0  $\rightarrow$  2), repair (2  $\rightarrow$  0) or repair:mismatch (2  $\rightarrow$  0:0  $\rightarrow$  2) substrates. Boxed bases denotes mismatches. The 3 nt sequence shown here resides at positions 20 to 22 of the corresponding substrates shown in Figure 2. B, Branch migration of mismatch (0  $\rightarrow$  2), repair (2  $\rightarrow$  0) or repair:mismatch (2  $\rightarrow$  0:0  $\rightarrow$  2) substrates. In the case of repair:mismatch assays, the sequence heterologies were located at adjacent positions (1), or were two or five bases apart. Assays were carried out at 37°C in Tris/EDTA containing 10 mM MgCl<sub>2</sub>. C, Branch migration of completely homologous duplexes (WT), mismatch (0  $\rightarrow$  2) or repair:mismatch 2  $\rightarrow$  0:0  $\rightarrow$  2 (1) substrates at 37°C in Tris/EDTA containing 100 mM NaCl. Repair:mismatch heterologies were separated by 1, 2 or 5 bp. Data are the average of at least two independent experiments.

unfavorable due to the loss of base-pairing when mismatches are reformed.

The effect of a repair event on spontaneous branch migration was tested using 61 bp substrates shown in Figure 5A. Two branch migration substrates, one bearing a G:A mismatch and the other having an opposing C:T mismatch were paired in branch migration assays carried out at 37°C in the presence of 10 mM MgCl<sub>2</sub>. As can be seen in Figure 5B, the time-course of branch migration was slightly faster for the repair (2 → 0) substrates compared to the corresponding wild-type substrates.

The repair event provided us with means to more rigorously determine the step-size of spontaneous branch migration. Branch migration assays were carried out in which a repair event was followed by a mismatch event located at varying distances, one, two or five base-pairs away (2 → 0:0 → 2). As can be seen from Figure 5A, the repair event acts as a reflecting barrier to the right whereas the mismatch event is a reflecting barrier to the left resulting in the junction being trapped in a random walk in the interval  $L_2$  spanning the region between the two sequence heterologies. If the interval  $L_2$  is relatively short, the time-course for RO in the case of a repair:mismatch should be faster than a mismatch alone, but slower than a repair event (or wild-type). The inhibitory effect of the mismatch should be evident when the repair

and mismatch events are separated by a distance that exceeds the step-size of branch migration. If the repair and mismatch events are encompassed within the same step, there is no net change in hydrogen bonding and the inhibitory effect of the mismatch should be masked. Computer simulations of a random walk through the repair:mismatch scenario indicated that as  $L_2$  lengthens, the time required for RO also increases in a manner roughly proportional to the square of  $L_2$  (data not shown).

Branch migration with various repair:mismatch substrates was carried out at 37°C in the presence of 10 mM MgCl<sub>2</sub>,  $\tau = 300$  ms (Figure 5B). It is evident that placement of a mismatch event adjacent to the repair event results in a substantial block to branch migration reminiscent of that observed for a mismatch alone under these conditions (see Figure 2B). Once again, however, branch migration through mismatches in the presence of Mg<sup>2+</sup> is too slow to assess the effect of distance between the repair and mismatch events.

A more sensitive measure of the extent to which the repair and adjacent mismatch events act independently during branch migration was obtained by increasing the intrinsic rate of spontaneous branch migration almost a thousandfold by removing Mg<sup>2+</sup> (Figure 5C). When branch migration reactions involving wild-type (WT), mismatch (0 → 2) or repair:mismatch (2 → 0:0 → 2) substrates were repeated at 37°C in Tris/EDTA plus 100 mM NaCl,  $\tau = 300$   $\mu$ s, it is evident that the repair:mismatch reactions were slower than wild-type, reflecting the inhibitory effect of the mismatch, but faster than a mismatch alone, reflecting the enhancement of branch migration contributed by the repair event. The initial slopes of the data shown in Figure 5C for repair:mismatch events at adjacent residues (1), or separated by 2 or 5 bp roughly approximate the expected difference in the rate of RO due to increasing the length of the  $L_2$  interval. Thus, the experimentally determined initial rate of RO decreased by fourfold and 25-fold, respectively, when the mismatch initially adjacent to the repair was moved 2 or 5 bp away. These rates of RO were two- to fivefold slower than expected from the simulations (decreases of two- and fivefold, respectively for the +2 and +5 cases). These discrepancies between the experiment and the simulation could reflect a slower step-time for branch migration in the case of repair:mismatch, perhaps due to the presence of two sequence heterologies in close proximity. Another explanation is that the slower experimentally determined rate of RO reflects a slow step in the annealing or initiation steps that is apparent at conditions of low RO. Nevertheless, the data clearly demonstrate that a repair and a mismatch situated at adjacent residues exert their effects on branch migration more or less independently, despite no net change in base-pairing. We conclude that the step-size is very small, on average 1 or 2 bp.

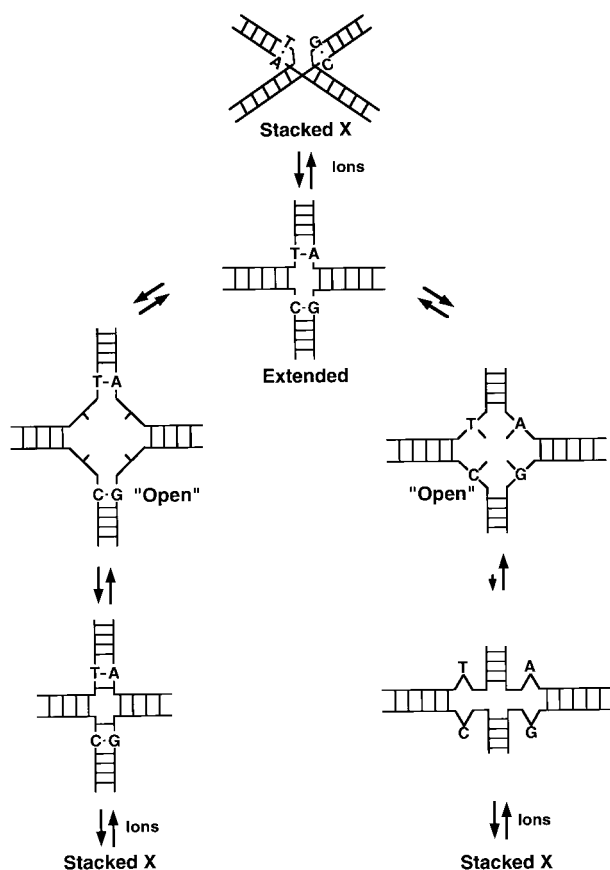


Figure 6. Scheme for spontaneous branch migration. Details are described in the text.



## Discussion

Previously, it was shown that spontaneous branch migration of a Holliday junction recombination intermediate is unable to accommodate sequence heterologies as short as one base-pair under near physiological conditions (Panyutin & Hsieh, 1993). Since homologous recombination involves the exchange of similar, though not necessarily identical sequences, this finding provided one explanation for the requirement for proteins involved in recombination and the recombinational repair of DNA damage to facilitate branch migration. Here, we exploit the inhibitory effect of sequence heterology to determine that the step-size of spontaneous branch migration is, on average, one to two base-pairs. In addition, a computer simulation of spontaneous branch migration through sequence heterology in combination with measurements of the time required to branch migrate through a single base-pair heterology reveal that the loss of base-pairing is the underlying molecular basis for the barrier to branch migration imposed by mismatches.

The barrier to spontaneous branch migration imposed by sequence heterology is substantial. Branch migration through a mismatch within the time-frame required for branch migration through completely homologous duplexes necessitated a 10,000-fold acceleration in the intrinsic rate of branch migration. Since sequence heterology is a reflecting barrier, i.e. the junction migrates in a random walk throughout the region of homology, we conclude that while sequence heterology is a barrier to spontaneous branch migration, it is not the case that the junction is trapped for any appreciable length of time in the vicinity of a mismatch. The data suggest that mismatches do not pose any significant conformational constraints on junction isomerizations that occur during the migratory step, though as discussed below, they do affect other aspects of junction conformation such as the choice of stacking partners or the extent of junction folding in the presence of  $Mg^{2+}$ .

What is the underlying molecular basis for the barrier to branch migration imposed by sequence heterology? One candidate is the destabilization of the heteroduplex DNA containing mismatched bases due to the loss of hydrogen bonding and stacking interactions in the vicinity of the mismatches. Using a computer simulation of a random walk, we introduced a bias in the random walk at the position of the mismatched bases. The best fit of this simulation with the experimentally measured time-course of branch migration corresponded to a bias of approximately  $10^4$  in favor of movement in the direction that restores perfect base-pairing. In other words, a Holliday junction that migrates through a single base heterology to create two mismatched duplexes very rapidly reverses direction to restore base-pairing. Thus, the number of times a junction resides at position  $n$  in the energy diagram in Figure 3A is very small, due

to the rapid reversal of branch migration dictated by the difference in ground states of perfectly paired *versus* mismatched duplexes. As a result, the junction spends most of its time in a random walk in the region between the origin and the sequence heterology. Only rarely will a junction migrate to distal positions, and even then, the random walk nature of branch migration dictates that the junction may return to regions proximal to the heterology (there is no barrier on its return).

An examination of the free energy differences between fully base-paired and mismatched duplexes suggests that the destabilization of duplex DNA by mismatched bases is the major contributor to the barrier to spontaneous branch migration imposed by sequence heterology. Physical studies of perfectly paired and mismatched duplexes indicate that the  $\Delta\Delta G^\circ$  for the formation of two mismatches, A-C and a G-T, from perfectly paired DNA duplexes is in the range of 6 to 8 kcal/mol (Aboul-ela *et al.*, 1985; Gaffney & Jones, 1989; Otokiti & Sheardy, 1997; Tibanyenda *et al.*, 1984; Werntges *et al.*, 1986). Since  $\Delta G^\circ = -RT \ln K$ , we calculate a value of approximately  $0.4 \times 10^{-5}$  to  $9 \times 10^{-5}$  for the equilibrium constant,  $K$ , for the formation of two mismatched heteroduplex DNAs by branch migration through sequence heterology at 50°C. Since the bias that was determined based on branch migration experiments and the computer simulation is of the order of  $1 \times 10^{-4}$ , we conclude that destabilization of the heteroduplex DNAs containing mismatched bases is largely responsible for the barrier to branch migration. We postulate that sequence heterology does not significantly increase the activation energy of the branch migration step that creates mismatched DNA, though this has not been shown directly. Since base-pairs in the vicinity of the junction crossover must be unpaired in the open intermediate, it seems unlikely that mismatches will destabilize the intermediate resulting in a higher activation energy.

How many base-pairs constitute one migratory step? In this study, we have positioned sequence heterologies at two adjacent positions in otherwise homologous duplex regions. The first heterology is a repair event consisting of pre-existing mispaired bases on each of two duplex substrates. The second is a mismatch heterology in which a G:C base-pair resides in one duplex while an A:T base-pair resides in the corresponding position in the second duplex. Branch migration through the repair heterology restores perfect base-pairing, while branch migration through the mismatch heterology creates two mismatches. There is no net change in the extent of hydrogen bonding if branch migration progresses through both heterologies.

The observation that two adjacent sequence heterologies each exert their effect on branch migration more or less independently, together with the initial observation that a single-base-pair heterology poses a substantial block to spontaneous branch migration under physiological con-

ditions (Panyutin & Hsieh, 1993) are indicative of a very small step-size, of the order of one or two base-pairs. Dynamic calculations of productive electrostatic forces that might operate in branch migration led Robinson & Seeman (1987) to propose that the step-size for spontaneous branch migration would be quite small. Our findings provide experimental support for that idea. We emphasize that the step-size, like determinations of the rate of branch migration, are indirect measurements and represent an average. The size of any particular step is likely to be influenced by parameters such as sequence and secondary structure as well as temperature.

Based on previous work from a number of laboratories (reviewed by Hsieh & Panyutin, 1995; see also Kirby *et al.*, 1997; Mulrooney *et al.*, 1996) and the data presented here, a scheme for spontaneous branch migration is shown in Figure 6. Under physiological conditions, the rate-limiting step is the conformational change of the Holliday junction from stacked X conformation to one approximating an extended conformation in which the arms are unfolded and base-stacking is lost at the crossover point. Branch migration involves the breakage and reformation of Watson–Crick hydrogen bonds as one DNA strand is exchanged for another. We envision this to occur within an intermediate having an “open” conformation, in which at least two base-pairs have been opened, though the exact conformation of this intermediate is unknown. Reformation of base-pairs with new partners results in migration of the junction 1 or 2 bp at a time. It is important to note the distinction between the step-size of branch migration and the number of base-pairs that are opened during the course of each migratory step. While our data provide support for the notion that the Holliday junction traverses only a few base-pairs per step, it is possible that a larger number of base-pairs are transiently open to facilitate junction movement. When sequence heterology is encountered such that branch migration may result in mismatches, the equilibrium lies in the direction of the restoration of base-pairing. In those rare instances when mismatched bases are present at the crossover, the propensity of these junctions to isomerize to the stacked X conformation in the presence of  $Mg^{2+}$  will vary. In addition to inhibiting branch migration, mismatched base-pairs can affect the extent of junction folding into a stacked X conformation as well as the choice of stacking partners (Duckett & Lilley, 1991). In summary, the random walk nature of spontaneous branch migration, the slow step-rate under physiological conditions, the small step-size and the large barrier to migration posed by small sequence heterologies explain why proteins are required to promote this step *in vivo*. In eukaryotes, these proteins may also have to contend with chromatin structure. We have recently shown that a nucleosome, the basic repeating unit of chromatin, is a barrier to spontaneous branch migration (Grigoriev & Hsieh, 1997).

In *E. coli*, rapid, unidirectional branch migration is promoted by the RuvAB and RecG proteins (reviewed by Lloyd & Low, 1996; Shinagawa & Iwasaki, 1996; West, 1996). RuvA is a junction-specific binding protein that deforms the Holliday junction and serves, *via* interactions with RuvB protein, to tether the DNA motors residing in an opposing pair of RuvB rings to the junction. An important function of these proteins is to facilitate branch migration through heterologous DNA sequences. Both the RuvAB and RecG proteins do so in an ATP-dependent manner (Iype *et al.*, 1994; Lloyd & Sharples, 1993; Parsons *et al.*, 1995b).

RecA protein also promotes strand exchange through sequence heterology in duplex-duplex reactions in an ATP-dependent manner (Hahn *et al.*, 1988; Morel *et al.*, 1994). Similarly, strand exchange reactions involving RecA-mediated invasion of a duplex by a single-strand DNA bearing heterologous sequences are also ATP-dependent (Kim *et al.*, 1992; Rosselli & Stasiak, 1991) suggesting that in both three and four-strand reaction, RecA overcomes the unfavorable energy of mismatch formation by hydrolyzing ATP. In addition to ATP-dependent mechanisms, some proteins may utilize the energy of DNA binding to promote heteroduplex formation between similar but not identical sequences. For example, an “antiproof-reading” activity of RecA protein is attributable to a substantial reduction in the free energy of mismatch formation mediated within a synaptic complex of a RecA-DNA nucleoprotein (Malkov *et al.*, 1997).

Studies of model branch migration reactions have provided new insights into a fundamental step in recombination. Understanding the effect of heterologous DNA sequence on this process is of vital importance for illuminating the molecular details of recombination. In addition, such studies are important for understanding the molecular mechanisms involved in the regulation of recombination of evolutionarily divergent sequences by the DNA mismatch repair pathway (reviewed by Kolodner, 1996; Matic *et al.*, 1996; Modrich & Lahue, 1996).

## Materials and Methods

### Preparation of branch migration substrates

Deoxyoligonucleotides were synthesized by automated  $\beta$ -cyanoethyl phosphoramidite DNA synthesis (Glen Research Corp.) on a model 308B synthesizer (Automated Biosystems, Inc.) and purified by denaturing urea/polyacrylamide (National Diagnostics) gel electrophoresis. Duplex substrates formed by annealing two complementary oligonucleotides were purified on native polyacrylamide (Bio-Rad) gels. Duplex substrates were 5'-end labeled with phage T4 DNA polynucleotide kinase (New England BioLabs) and [ $\gamma$ - $^{32}P$ ]ATP (New England Nuclear). For each branch migration reaction, two duplex substrates, S1 and S2, were used. The homologous duplex regions of each substrate are shown in Figure 2. Sequences of the 20 nt heterologous tails are:

S1 5' ACC ATG CTC GAG ATT ACG AG and 3' CTA GGA TCC GGA GTG CAT AA; and S2 5' GAT CCT AGG CCT CAC GTA TT and 3' TGG TAC GAG CTC TAA TGC TC.

### Branch migration assays

Assays were performed as described (Panyutin *et al.*, 1995). Briefly, a <sup>32</sup>P-labeled S1 duplex substrate (1.5 to 4 ng/μl) was mixed on ice with a tenfold molar excess of the S2 unlabeled substrate (15 to 40 ng/μl) in a total volume of 2.5 μl containing 10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA and either 100 mM NaCl or 10 mM MgCl<sub>2</sub>. After incubation at the appropriate temperature, the reaction was quenched by the addition of ethidium bromide and Mg<sup>2+</sup>, and electrophoresed on native polyacrylamide gels as described (Panyutin *et al.*, 1995). Branch migration assays were quantified using a Fuji Phosphorimager.

### Simulation of branch migration

The simulation program was written, compiled and executed with GNU C++ version 2.7.2 on the Intel-Pentium-Linux 2.0.30 platform and C version 7.0 (Silicon Graphics) on the SGI-MIPS-irix 5.3 platform. The random number generator function was adapted from Schneier (1992). The variation between platforms was within the range of the variation from multiple trials on the same platforms (<2%). On all runs, the length of the random walk was set to 60 bp and, unless otherwise specified, the mismatch resided at position 21 to correspond to experimental conditions. The program counts the number of steps at each trial required for runoff as well as the number of times that the junction resides at a specified position for each trial.

### Acknowledgments

We are grateful to Igor Panyutin and Vlad Malkov for helpful discussions, and to them and R. Daniel Camerini-Otero for comments on the manuscript. We thank George Poy for oligonucleotide synthesis, Scott McIntosh for advice on computer programming and Linda Robinson for her assistance.

### References

- Aboul-ela, F., Koh, D. & Tinoco, I., Jr (1985). Base-base mismatches. Thermodynamics of double helix formation for dCA<sub>3</sub>XA<sub>3</sub>G + dCT<sub>3</sub>YT<sub>3</sub>G (X,Y = A,C,G,T). *Nucl. Acids Res.* **13**, 4811–4824.
- Adams, D. E. & West, S. C. (1995). Relaxing and unwinding on Holliday: DNA helicase-mediated branch migration. *Mutat. Res.* **337**, 149–159.
- Duckett, D. R. & Lilley, D. M. J. (1991). Effects of base mismatches on the structure of the four-way DNA junction. *J. Mol. Biol.* **221**, 147–161.
- Fujitani, Y. & Kobayashi, I. (1995). Random-walk model of homologous recombination. *Phys. Rev. sect. E*, **52**, 6607–6622.
- Gaffney, B. L. & Jones, R. A. (1989). Thermodynamic comparison of the base pairs formed by the carcinogenic lesion O<sup>6</sup>-methylguanine with reference both to Watson–Crick pairs and to mismatched pairs. *Biochemistry*, **28**, 5881–5889.
- Grigoriev, M. & Hsieh, P. (1997). A histone octamer blocks branch migration of a Holliday junction. *Mol. Cell. Biol.* **17**, 7139–7150.
- Hahn, T.-R., West, S. & Howard-Flanders, P. (1988). RecA-mediated strand exchange reactions between duplex DNA molecules containing damaged bases, deletions and insertions. *J. Biol. Chem.* **263**, 7431–7436.
- Hsieh, P. & Panyutin, I. G. (1995). DNA branch migration. In *Nucleic Acids and Molecular Biology* (Eckstein, F., Lilley, D. M. J., eds), vol. 9 edit., pp. 42–65, Springer-Verlag, Berlin.
- Iype, L. E., Wood, E. A., Inman, R. B. & Cox, M. M. (1994). RuvA and RuvB proteins facilitate the bypass of heterologous DNA insertions during RecA protein-mediated DNA strand exchange. *J. Biol. Chem.* **269**, 24967–24978.
- Kim, J.-I., Cox, M. M. & Inman, R. B. (1992). On the role of ATP hydrolysis in RecA protein-mediated DNA strand exchange. I. Bypassing a short heterologous insert in one DNA substrate. *J. Biol. Chem.* **267**, 16438–16443.
- Kirby, A. W., Gaskin, M. N., Antezana, M. A., Goodman, S. J., Myers, E. & Bruist, M. F. (1997). Triple-helical DNA as a reversible block of the branch point in a partially symmetrical DNA four-arm junction. *J. Mol. Biol.* **271**, 349–361.
- Kolodner, R. (1996). Biochemistry and genetics of eukaryotic mismatch repair. *Genes Dev.* **10**, 1433–1442.
- 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. & Low, K. B. (1996). Homologous recombination. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., ed.), pp. 2236–2255, ASM Press, Washington, DC.
- Lloyd, R. G. & Sharples, G. J. (1993). Processing of recombination intermediates by the RecG and RuvAB proteins of *Escherichia coli*. *Nucl. Acids Res.* **21**, 1719–1725.
- Malkov, V., Sastry, L. & Camerini-Otero, R. D. (1997). RecA protein assisted selection reveals a low fidelity of recognition in a duplex DNA by an oligonucleotide. *J. Mol. Biol.* **271**, 168–177.
- Matic, I., Taddei, F. & Radman, M. (1996). Genetic barriers among bacteria. *Trends Microbiol.* **4**, 69–73.
- Modrich, P. & Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* **65**, 101–133.
- Morel, P., Stasiak, A., Ehrlich, S. D. & Cassuto, E. (1994). Effect of length and location of heterologous sequences on RecA-mediated strand exchange. *J. Biol. Chem.* **269**, 19830–19835.
- Muller, B., Burdett, I. & West, S. C. (1992). Unusual stability of recombination intermediates made by *Escherichia coli* RecA protein. *J. Bacteriol.* **178**, 1237–1241.
- Mulrooney, S. B., Fishel, R. A., Hejna, J. A. & Warner, R. C. (1996). Preparation of Figure 8 and cruciform DNAs and their use in studies of the kinetics of branch migration. *J. Biol. Chem.* **271**, 9648–9659.
- Otokiti, E. O. & Sheardy, R. D. (1997). Effect of base pair A/C and G/T mismatches on the thermal stabilities of DNA oligomers that form B-Z junctions. *Biochemistry*, **36**, 11419–11427.
- Panyutin, I. G. & Hsieh, P. (1993). Formation of a single base mismatch impedes spontaneous DNA branch migration. *J. Mol. Biol.* **230**, 413–424.

- 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.
- Parsons, C. A., Stasiak, A., Bennett, R. J. & West, S. C. (1995a). Structure of a multisubunit complex that promotes DNA branch migration. *Nature*, **374**, 375–378.
- Parsons, C. A., Stasiak, A. & West, S. C. (1995b). The *E. coli* RuvAB proteins branch migrate Holliday junctions through heterologous DNA sequences in a reaction facilitated by SSB. *EMBO J.* **14**, 5736–5744.
- Robinson, B. H. & Seeman, N. C. (1987). Simulation of double-stranded branch point migration. *Biophys. J.* **51**, 611–626.
- Rosselli, W. & Stasiak, A. (1991). The ATPase activity of RecA is needed to push the DNA strand exchange through heterologous regions. *EMBO J.* **10**, 4391–4396.
- Schneier, B. (1992). Pseudo-random sequence generator for 32-bit microprocessors. *Dr. Dobbs' J.* **17**, 34–40.
- Seeman, N. C. & Kallenbach, N. R. (1994). DNA branched junctions. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 53–86.
- Shinagawa, H. & Iwasaki, H. (1996). Processing the Holliday junction in homologous recombination. *Trends Biochem. Sci.* **21**, 107–111.
- Tibanyenda, N., De Bruin, S. H., Haasnoot, C. A. G., Van Der, Marel G. A., Van Boom, J. H. & Hilbers, C. W. (1984). The effect of single base-pair mismatches on the duplex stability of d(TATTAATATCAAGTTG)-d(CAACTTGATATTAATA). *Eur. J. Biochem.* **139**, 19–37.
- Werntges, H., Steger, G., Riesner, D. & Fritz, H.-J. (1986). Mismatches in DNA double strands: Thermodynamic parameters and their correlation to repair efficiencies. *Nucl. Acids Res.* **14**, 3773–3790.
- West, S. C. (1996). DNA helicases: New breeds of translocating motors and molecular pumps. *Cell*, **86**, 177–180.

*Edited by M. Gottesman*

*(Received 22 December 1997; received in revised form 25 February 1998; accepted 2 March 1998)*