SIMULATION OF DOUBLE-STRANDED BRANCH POINT MIGRATION

BRUCE H. ROBINSON* AND NADRIAN C. SEEMAN‡
*Department of Chemistry, University of Washington, Seattle, Washington 98195; ‡Department of Biological Sciences, State University of New York at Albany, Albany, New York 12222

ABSTRACT A structural and dynamic model has been developed for the branch point formed when two DNA double helices exchange strands during genetic recombination. This model, which generalizes most previous structural models, maintains the twofold symmetry inherent in the covalent and hydrogen bonded structure, yet has three degrees of freedom about virtual bonds, constituting a simplified junction. Using this structural model, a three-step dynamic model for branch point migration has been developed: (*a*) longitudinal diffusion about the virtual bonds to achieve a structure in which the helix axes are approximately parallel; (*b*) opening of the base pairs; and (*c*) rotary diffusion about the helix axis to effect a migratory event. The model, which includes the possible role of electrostatic interactions, solves problems inherent in previous treatments. We find that no significant electrostatic torques arise that promote branch point migration. The absence of a kinetic mechanism to circumvent thermodynamic barriers due to mispairing suggests that an energy source is used for those situations in living systems.

INTRODUCTION

Recombination between two homologous double helices of DNA to form two newly hybridized double strands of DNA is one of the fundamental processes in the generation of genetic diversity in biological systems. The recombinant DNA is believed to arise via the Holliday (1964) intermediate, indicated in Fig. 1 *b*. This cross-stranded structure, in which the strands from two DNA double helices have been exchanged, is shown with a hypothetical sequence. This structure is topologically identical to the "crossroads-junction" structure illustrated in Fig. 1 *a*. Both conformers of this intermediate structure have twofold rotational symmetry, as indicated by the appropriate symbols in the figure. The strands with the same shading in the figures are considered to have been paired before strand exchange was initiated. The point at which the strands interchange is termed the branch point, or, alternatively, the junction. This corresponds to the crossover point of Fig. 1 *b*, and to the middle of the intersection in Fig. 1 *a*. The interchange process is thought to occur by migration of the branch point, which involves rotation of the double helices constituting it. Either of the two migration reactions (I or II) shown in Fig. 1 may occur. Repetition of reaction I results in a return to the original pairing: strand 1 paired with strand 2 and strand 3 paired with strand 4. Conversely, repetition of reaction II yields a new hybridization: strand 1 paired with strand 4 and strand 2 paired with strand 3. If the two initial double helices (1-2 and 3-4) are exactly identical, no genetically assayable recombination event will occur, regardless of which reaction pathway is involved. However, if the two double helices are only homologous, with slight sequence differences between them, reaction II will lead to a recombinant product. In this case, resolution of the junction structure via reaction II results in two double helices that are not optimally paired, and are therefore appropriate substrates for repair enzymes.

Numerous attempts have been made to model the structural features involved in the mechanism of double-stranded branch point migration. Structural models have been proposed by Sigal and Alberts (1972), Sobell (1972), McGavin (1977), Wilson (1979), and ourselves (Seeman et al., 1979; Seeman and Robinson, 1981). The structural model proposed by Sigal and Alberts is physically analogous to the schematic diagram in Fig. 1 *b*, with both the helix axes and two parallel strands virtually unperturbed. The other two strands have altered conformations at the junction, because they are exchanged between the double helices. The sequence similarities imply twofold spatial symmetry, as noted by Sobell (1972). He has also pointed out that the backbone structures can assume fourfold symmetry, analogous to Fig. 1 *a*: The helix axes of the four double helices emanating from the branch point in that figure are related by a fourfold rotational symmetry operation, whose axis is coincident with the twofold axis shown. McGavin's (1977) and Wilson's (1979) intriguing proposals, which involve tetrahelical structures, are beyond the scope of our treatment.

The details of our structural and dynamic model are described below. The structural proposals are based on molecular model building. No experimental data yet exist at an adequate resolution to indicate the detailed structural properties and parameters of the branch point; however, electron microscopic observations (Lee et al., 1970; Broker

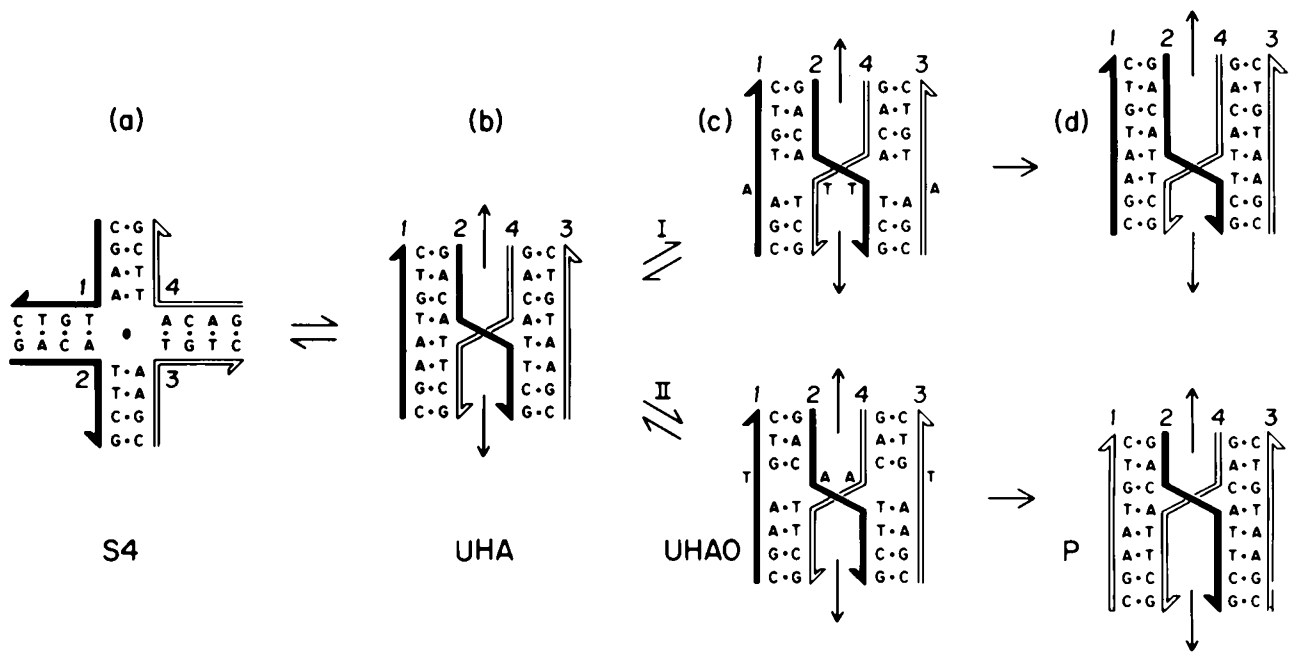


FIGURE 1 Proposed mechanism for branch point migration showing all structurally distinct intermediates. (*a*) Schematic representation of the fourfold symmetric model for the junction (S4) originally proposed by Sobell (1972). This structure is topologically identical to the Sigal–Alberts (1972) version of the Holliday (1964) structure, with unperturbed helix axes (UHA), shown schematically in *b*. Throughout the figure, the staded backbones were initially paired exclusively to each other, as were the unshaded backbones. The half-arrows indicate the 5′ → 3′ directions of the strands. The full headed arrows indicate the axis of twofold symmetry in *b–d*. The lens-shaped object in the center of *a* serves the same role there. The junction is the crossroads region of *a* and the place where the strands cross in *b–d*. Migration of the junction in either direction is indicated by reactions I and II in the transition from *b* to *c*. *c* shows the two junctions with one of the sets of base pairs that flank the junction opened (UHAO). *d* indicates the product of the reaction (P), the migrated branch point. The mechanism (see text) involves longitudinal diffusion of the helix axes from the electrostatically favored S4 structure to the UHA structure, followed by base pair opening to yield the UHAO structure, and then rotary diffusion to give the product, P, the migrated junction. The eventual end product of the repetition of reaction I is a return to the original pairing. The eventual end product of the repetition of reaction II is a newly hybridized pair of double helices.

and Lehman, 1971; Kim et al., 1972; Wolgemuth and Hsu, 1980) have confirmed the existence of the branch point and of the migratory phenomenon. It is likely that physical studies of immobile and semimobile nucleic acid junctions (Seeman, 1981, 1982; Kallenbach et al., 1983*a*,*b*) will soon generate some of the missing structural information.

It is possible to devise numerous models of branch point migration dependent upon an external energy source, either direct, such an a nucleoside triphosphate (e.g., McEntee et al., 1979), or indirect, such as DNA supercoiling (Gellert et al, 1978; Lilley, 1980; Panayotatos and Wells, 1981). The major kinetic model, independent of an external energy source, is due to Meselson (1972), in conjunction with the Sigal–Alberts unperturbed heliz axes model. He has postulated that rotary diffusion of the double helices about their axes would generate migration with this sort of structure. No comparable proposal has been forthcoming for the fourfold symmetric structures. Numerous investigators have measured the rate of branch point migration under a range of conditions (Thompson et al., 1976; Warner et al., 1978; Courey and Wang, 1983; Gellert et al., 1983; Lilley and Hallam, 1984; Warner, R. C., personal communication).

Here we present a structural and dynamic model for the migrational process. The structural model subsumes the major features of the Sigal–Alberts and Sobell models, which correspond to special cases of this more general formulation. The parameterization of the branch point junction complex results in a large number of conformers that are in dynamic equilibrium. We have calculated the electrostatic energies that characterize the different conformers, along with their associated probabilities, for the entire range of possible structures associated with our parameterization. These probabilities suggest that a longitudinal diffusion step can be added to the dynamic processes previously proposed by Meselson. We have also updated Meselson's dynamic treatment of branch point migration, by including more recent data on torsional correlation lengths in double-stranded DNA.

THE STRUCTURAL MODEL FOR THE JUNCTION

We have parameterized the junction under an inclusive formulation that retains the intrinsic twofold rotational symmetry implied by the covalent and hydrogen bonded structure of the junction. Fig. 2 shows this treatment of the junction structure. The difficulty in understanding a com-

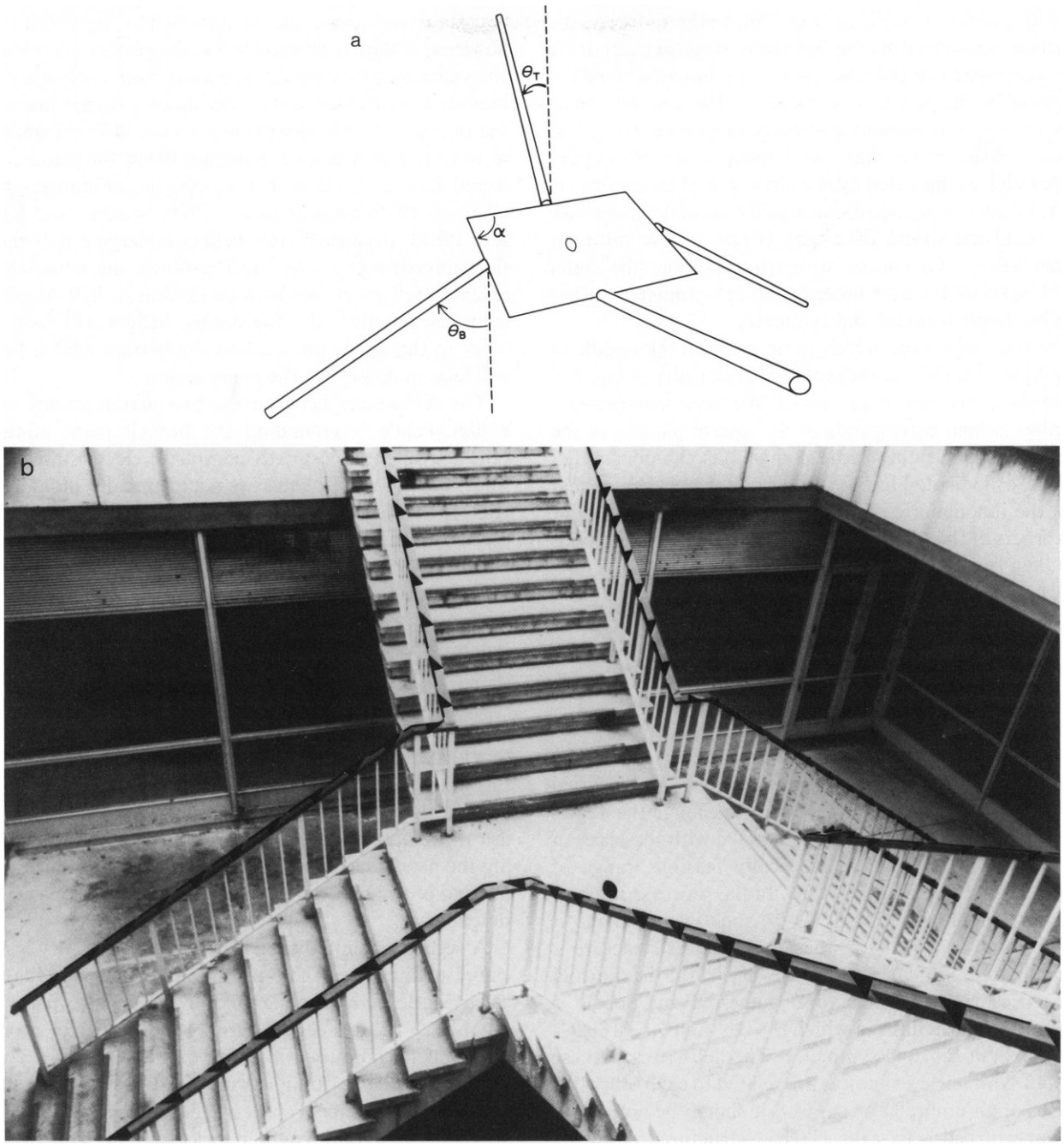


FIGURE 2 Representation and parameterization of the recombinational junction structure. (*a*) Schematic representation of the junction. The four double helices that comprise the junction are indicated by their helix axes, which are the four rod-like elements emanating from the central rhombus. This rhombus represents the junction itself, the four phosphates flanking the junction being placed near the corners of the rhombus. The twofold symmetry axis perpendicular to the plane of the rhombus is indicated by the lens-shaped symbol at the center of the rhombus. The three degrees of freedom are indicated, as well: α is the variable angle of the rhombus, θ_T is the degree of freedom for the two double helices indicated above the plane of the rhombus, and θ_B is the degree of freedom for the two double helices indicated below the plane of the rhombus. (*b*) The staircase analogy to the junction structure. In this representation, the rhombus of *a* corresponds to the plateau at the center of the staircase. The banisters correspond to the backbones of the four strands that comprise the junction, while the steps are analogous to the base pairs. The polarity of the backbones is indicated by the arrows on the banisters, showing the alternately antiparallel nature of the backbones. The twofold symmetry of the junction is again illustrated by the lens-shaped symbol at the center of the plateau. Note that two people walking down the opposite top stairways will face each other, while two people walking down the opposite bottom stairways will have their eyes oriented in opposite directions.

plete molecular model of the junction has led us to represent it both schematically (Fig. 2 *a*) and by comparison with a familiar staircase (Fig. 2 *b*). In the staircase, the banisters correspond to the backbone structures, and the steps correspond to the base pairs. The junction region is suggested by the plateau in the middle. The structure has a twofold axis of rotational symmetry perpendicular to the plateau plane. Note that the banisters are alternately antiparallel, as indicated by the directions of the arrows on them. From this figure, it can readily be understood how each backbone strand (banister) is part of two different double helices. Of course, in a true junction, the stairs would be related to each other by helical symmetry, rather than by simple translational symmetry.

The four helix axes, which go up or down the middle of each flight of stairs, are indicated schematically in Fig. 2 *a*, which shows the rest of our model. We have constructed a rhombus, which corresponds to the square plateau of the staircase. The variable parameter of this rhombus is the angle α, as indicated in Fig. 2 *a*. The four phosphates that flank the junction are positioned slightly above and below the corners of this rhombus. Their separation and positions are taken from B-DNA (Arnott and Hukins, 1972).

Besides the variability of the rhombus angle, we have constructed two virtual bonds coincident with two adjacent sides of the rhombus. The two variable angles associated with these virtual bonds, θ_T and θ_B, are indicated in Fig. 2 *a*. Note that the senses of these two angles are opposite. The maintenance of the twofold rotational symmetry axis perpendicular to the plane of the rhombus requires that the helices attached to opposite sides of the rhombus be varied in pairs. These virtual bonds allow us to change the angles θ_T and θ_B, which the double helices make with the plane of the rhombus. The unique, physically feasible, range of each of these three angles is 180°. In the staircase illustration, for example, $\alpha = 90°$, $\theta_T = 60°$, and $\theta_B = 60°$. Thus, if $\alpha = 36°$, and $\theta_T = \theta_B = 0°$, the Sigal–Alberts structure is approximated. Similarly, if $\alpha = \theta_T = \theta_B = 90°$, the fourfold symmetric backbone structure suggested by Sobell is obtained. All structures in which $\alpha = 90°$ and $\theta_T = 180° - \theta_B$ will exhibit fourfold symmetry. The helix axes in these fourfold symmetric structures are related to each other like the ribs of an umbrella or parasol. It should be noted that the unperturbed-helix-axes (UHA) structure represents one extreme structural type under this formulation, while fourfold symmetric (S4) backbone arrangements represent the other.

The twofold sequence symmetry of the recombinational intermediate is to be expected, except for an occasional mismatch. We have assumed that this will be reflected structurally in twofold rotational symmetry, which all of our calculations maintain. It should be emphasized, however, that there is no symmetry relationship between the parts of the Holliday structure above and below the branch point. This point is evident from the stairway shown in Fig. 2 *b*.

The staircase in Fig. 2 *b* illustrates an important related point: Two individuals symmetrically walking down the two staircases above the plateau would face each other; however, if they continued below the plateau, down opposite staircases, they would face away from each other. The molecular equivalent of this asymmetry is that just above the plateau, the two double helices would face each other in, say, the major groove, while just below the plateau, they would face each other in the opposite, or minor, groove (Wilson, 1979; Seeman et al., 1979; Seeman and Robinson, 1981). Because double helices undergo a half-revolution approximately every five residues, the situation will reverse itself every five base pairs; thus, a half-turn away from the junction, the top double helices will face each other in the minor groove, and the bottom double helices will face each other in the major groove.

The differences between the two juxtapositions of the double helices surrounding the branch point might be reflected in the phosphate–phosphate electrostatic repulsions, since the phosphates are not identically situated with respect to the two grooves of the double helix. For this reason, we have calculated the electrostatic energy associated with each of the conformations available to the structure, $E(\alpha, \theta_T, \theta_B)$. If one assumes that the forces that determine the conformation of the helices about the junction are primarily electrostatic, the map of these energies is readily transformed into one showing the probability that the structure will assume any of these given conformations. Since electrostatic forces are the only long range forces in this large system that are readily recognized and treated, this assumption is not without utility. Our lack of knowledge of the detailed geometry of the junction forces us to use the simplified system described above, thereby ignoring the possibly substantial energetic effects of base stacking and of single bond conformations in the vicinity of the junction.

A second potentially relevant physical quantity is associated with each of these conformations. This is the total electrostatic torque, generated by phosphate–phosphate repulsions, i.e., an electrostatic torque that can promote branch point migration in either direction. In Fig. 3 *a*, we see a sample conformation, with $\alpha = 110°$, $\theta_T = \theta_B = 0°$. A productive migrational event will occur when all the helices are rotating either counterclockwise (as indicated) or clockwise (yielding migration in the opposite direction). We term the summed total torque about the helix axes in either of these directions the "productive torque." The magnitude of the productive torque indicates its contribution to migration, while its sign determines the direction of migration. It should be noted that this concept of a productive torque remains valid, even when the double helices that constitute the junction are no longer parallel to each other; it refers to the sum of the torques about the individual helix axes, regardless of the orientations of the axes themselves.

The reaction coordinate shown in Fig. 3 *b* complements

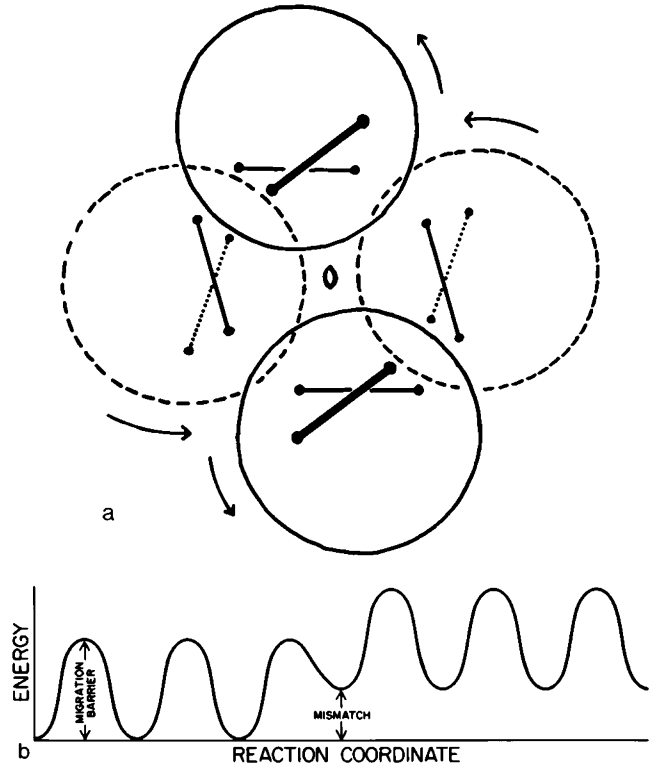


FIGURE 3 Nature of movement through the junction. (*a*) The directions of rotation to produce a productive migration. Shown is a junction with parameters $\alpha = 110°$, $\theta_T = \theta_B = 0°$. The view is perpendicular to the plane of the rhombus. The twofold symmetry is indicated by the lens-shaped symbol in the center of the figure. Two double helices emanate toward the reader from the plane of the paper; these are indicated by solid circles. Two other double helices are below the plane of the paper, emanating in the other direction; these are indicated by dashed circles. Within each circle are two pairs of dots, representing Cl' atoms. Those connected in pairs belong to the same base pair. The two pairs for each double helix nearest the junction are indicated in this fashion. Within the solid circles, the thicker line represents the base pair nearer the reader. Similarly, within the dashed circles, the solid line represents the base pair nearer the reader. In the course of migration toward the reader, the thin lines in the solid circles will move to replace the thick lines; the thick lines in the dashed circles will break and reform to replace the thin lines in the solid circles; meanwhile, the dotted lines in the dashed circles will move to replace the solid lines in the dashed circles. For these coordinate motions to occur, all four double helices must rotate in the counterclockwise directions indicated by the curved arrows. Migration in the other direction similarly entails coordinate rotation in the clockwise direction. (*b*) Schematic reaction coordinate for the migratory process. A single-step migration is separated by a single peak. The migration barrier is indicated in the first step. The result of a mismatch is also indicated: The relative barrier remains the same, but the relative minimum to which the system returns is now higher. No detailed statement about the structure of the energy function near the region of the mismatch is being made; only the qualitative nature of the phenomenon is being indicated.

the structural diagram shown in Fig. 3 *a*. Here, we have plotted the energy of the state of the system along the ordinate, while the effective reaction coordinate is indicated along the abcissa. Migration in one direction (say, counterclockwise) in Fig. 3 *a* corresponds to rightward movement along Fig. 3 *b*, while migration in the other direction (say clockwise) corresponds to movement in the leftward direction in Fig. 3 *b*. The effect of a mismatch between homologous recombinant fragments is shown. The incorporation of such an energetically unfavorable interaction would be facilitated if the effective step size were large. Therefore, we have sought large, asymmetric forces (torques in this case) to test this possibility. Our calculations indicate that electrostatic forces larger than kT do not exist for this system.

ELECTROSTATIC CALCULATIONS AND RESULTS

The quantitative details of the model used in the calculations are as follows: We have placed a single turn of B-DNA (Arnott and Hukins, 1972) on each of the sides of the rhombus, as indicated schematically in Fig. 2. In the initial conformation ($\theta_B = \theta_T = 0$), the vertices of the rhombus correspond to the projections of the phosphorus atoms onto the plateau plane; this specification serves to orient the double helices about their helix axes. The plane of the rhombus corresponds to the [Z = 0] plane, of the conventional coordinate system, thereby fixing the helices vertically. The four double helices are covalently joined into four polynucleotide backbones by the phosphate linkages associated with the corners of the rhombus. Thus, these phosphorus atoms occupy positions analogous to the corners of the four banisters of Fig. 2 *b*.

We have assumed that the monopole terms dominate the electrostatic interaction, and have placed a negative charge at the positions of the phosphorus atoms. This charge has been isotropically screened 76%, according to simple polyelectrolyte theory for monovalent cations (Manning, 1978), to yield an effective charge of −0.24. No further screening has been included due to the current lack of a sufficiently detailed structural model. The dielectric constant of water has been taken to be 78.4 (e.g., Handbook of Chemistry and Physics, 1981). Both the shielding and dielectric terms are important scaling factors for the energies and torques about the helix axes. The torques about the helix axes must be calculated using the actual positions of the phosphorus atoms, rather than positions projected onto the helix areas.

We have performed our calculations using a single turn of double helical B-DNA for each of the four double helices for all allowed conformations of the complex. However, we have used longer lengths for selected regions of conformation space, to determine the dependence of the forces on helical length. The stereochemical feasibility of all conformations has been checked. Those conformations that generated impossible steric contacts were omitted from the calculations. The only exception to this exclusion involves bad intrastrand contacts within a single residue of the virtual bonds. Our model makes no statement about the detailed geometry of the linkage there, and a small amount of conformational flexibility at those sites is to be expected. Viable space-filling and Kendrew models can be constructed for these regions that eliminate the bad contacts,

with small modifications to the local geometry. We have made no attempt to include flexibility parameters for the DNA, due to computing limitations.

Fig. 4 shows the conformational probability as a function of α. This probability is the total probability for all angles of θ_B and θ_T, and it is calculated from the energy of each state using the Boltzmann relation between probability and energy at 300°K. Since the distribution maximizes at 90°, conformations in which the rhombus approximates a square are favored.

It is convenient to approximate this distribution in an analytical form:

$$P(\alpha) = a\,[\sin(\alpha)]^{\nu'}. \quad (1)$$

For fixed θ_T or for fixed θ_B, the ν' parameters of $P(\alpha)$ were similar, which justifies the use of averaging over these angles. The points in Fig. 4 were calculated only at $\alpha = 90°$ for 18 and 50 base pair arms; ν' was then computed, thus yielding the entire curve. In this way the energies and probabilities for a single turn and for five turns of helix were extrapolated to appropriate lengths.

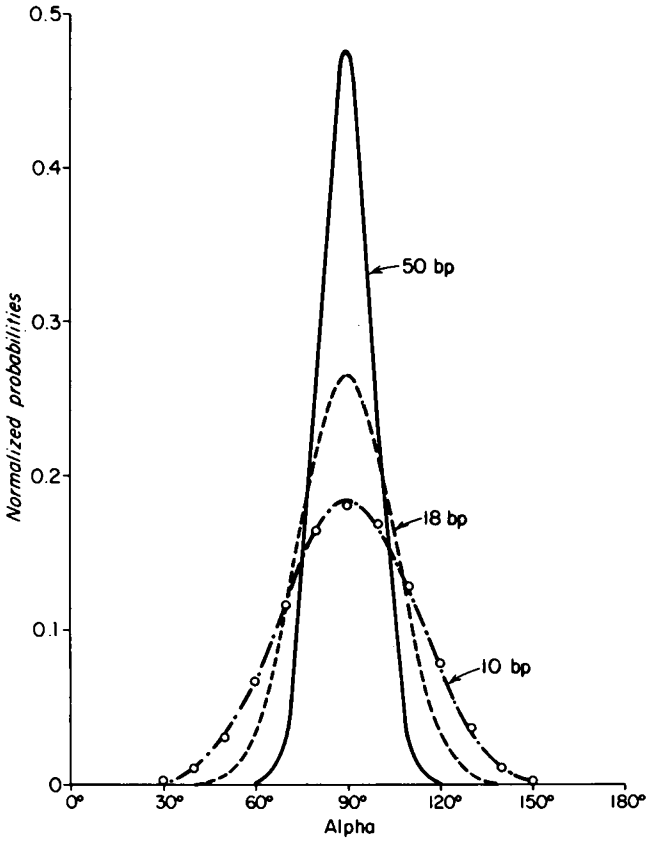


FIGURE 4 The probability distribution of α for various chain lengths. Curves are shown for lengths of double helix of 10, 18, and 50 base pairs. The points along the 10 base-pair curve are the calculated values. The curve through these points is the best fit to these points via the function indicated in the text. The curve for 50 base pairs is based on calculated values for selected points, from which an appropriate scale factor was derived. The curve for 18 base pairs is based on the function derived in the text. Note that the probability for the rhombus being a square increases with the length of the double helices emanating from the junction.

We present the most favored section in α, the section at $\alpha = 90°$, in Fig. 5 *a*. We have plotted the angles θ_T and θ_B in two dimensions, and contoured the electrostatic energy of the structure in increments of $\frac{1}{4}\,kT$. Note that the innermost contours are of the lowest energy, while the contours of highest energy are on the outside of the figure.

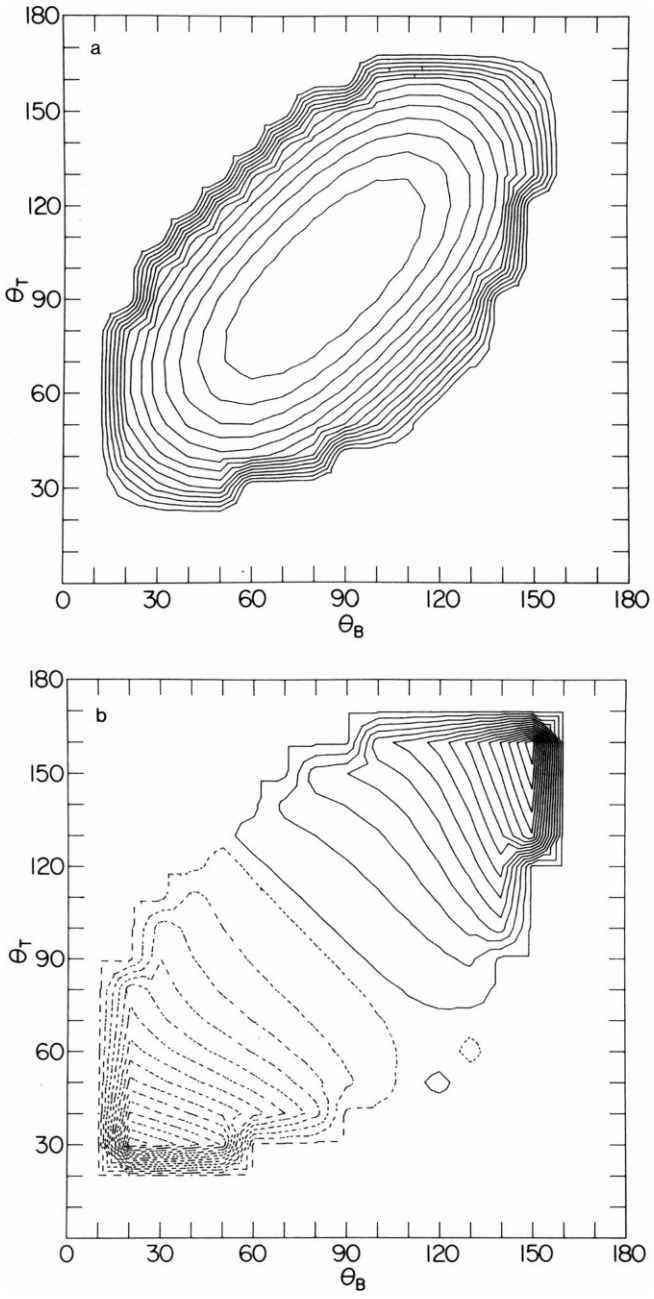


FIGURE 5 The energy and torque functions for the square function. (*a*) The energy section, contoured at intervals of $\frac{1}{4}\,kT$. The minimum is at the center of the figure, and higher energies are indicated as one crosses contour lines from that point. Note the twofold rotation axis lying in the plane of the section along the fourfold symmetry diagonal, $\theta_T + \theta_B = 180°$. (*b*) The productive torque function, contoured at intervals of 6 cal/mol-rad. The positive contours are indicated by solid lines, while negative contours are indicated by dotted lines. Note that the twofold antisymmetry rotation axis along the fourfold symmetry diagonal.

The region outside the highest contours is filled primarily with conformations that are sterically inaccessible. The diagonal corresponding to $\theta_T + \theta_B = 180°$ is an axis of twofold rotational symmetry for the three-dimensional function. This diagonal corresponds to the set of fourfold symmetric umbrella-rib-like backbone structures. The actual minimum of the energy function is very close to the position $\alpha = \theta_T = \theta_B = 90°$, which corresponds to a square planar configuration of the helix axes. The most probable configuration occurs at $\theta_T = 95°$, and $\theta_B = 80°$, and is still on the diagonal; this configuration deviates from square planarity because of differences between the major groove and minor groove phosphate structures.

The productive torques about the helix axes show symmetry about this same fourfold diagonal. Fig. 5 *b* indicates the productive torques calculated for the same section, $\alpha = 90°$. The function is contoured at intervals of 6 cal/mol-rad. Here, we can see that the diagonal becomes an axis of twofold rotational antisymmetry for the three-dimensional function. From the magnitudes of the contoured figure, it is evident that the torques in this section are very small, typically ~30–70 cal/mol-rad. The values of the torques throughout the other sections of the map do not differ from these values markedly: only rarely do values exceed 100 cal/mol-rad, and then, only for positions of extremely low probability. Thus, if the values of the dielectric constant and the shielding factor are accepted, a single turn of helix clearly does not generate sufficient torque to provide the driving force that would ameliorate the mispairing problem. The torques have been calculated for selected longer lengths, at selected conformations, and they do increase to significant values.

The structures of the energy and torque maps indicate that the symmetry diagonal is the most useful direction in which to project these functions. The probability plot in Fig. 4 indicates that little information will be lost if the functions are projected in α, as well. We have projected the results in both these directions, as shown in Fig. 6. This is a plot of the conformational probabilities versus the distance from the symmetry diagonal. In Fig. 6, we show the plot directly calculated, for a single turn of double helix emanating from each side of the junction, and an analytical curve that simulates this function.

$$P(\delta) = a\left[\cos\left(\frac{\delta}{2}\right)\right]^{\nu}, \tag{2a}$$

where

$$\delta = \theta_T + \theta_B - 180°. \tag{2b}$$

Extrapolations to longer lengths of double helix, based on calculations at selected values, are also shown. It should be noted that the abcissa, δ, corresponds roughly to deviation from fourfold symmetry. It describes the difference angle between the θ_B and θ_T ribs of the umbrella: Those values closest to the center relate to structures whose sets of helix

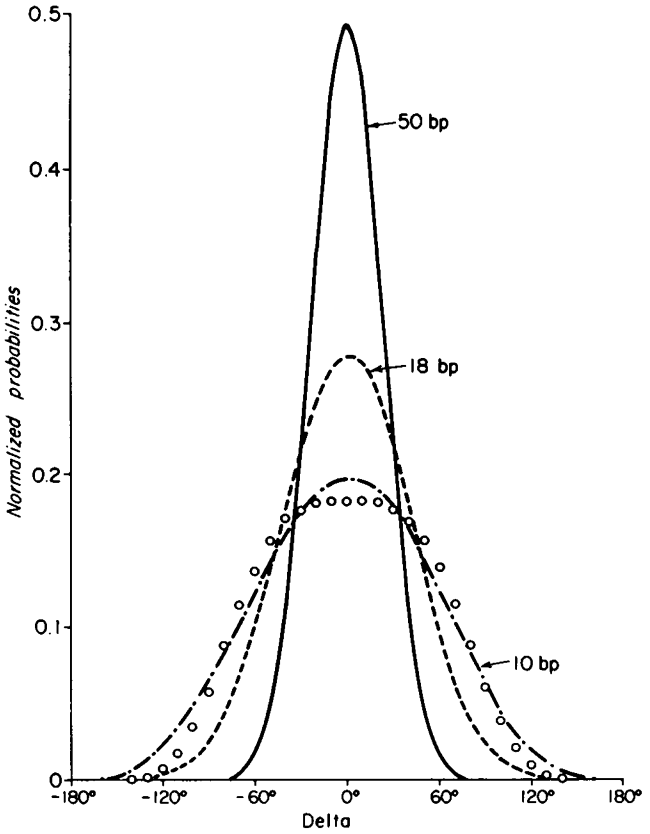


FIGURE 6 The probability distribution of δ for various chain lengths. Curves are shown for lengths of double helix of 10, 18, and 50 base pairs. The points along the 10 base-pair curve are the calculated values. The curve through these points is the best fit to these points via the function indicated in the text. The curve for 50 base pairs is based on calculated values for selected points, from which an appropriate scale factor was derived. The curve for 18 base pairs is based on the function derived in the text. Since δ is an indication of the deviation of the structure from fourfold symmetry, the probability of a fourfold symmetric structure increases as the chain length increases.

axes are oriented most similarly relative to the axis of twofold symmetry. There is an empirical relation that exists between the parameter ν (of Eq. 2a) and the number of base pairs, (n_b), in an individual arm of the structure:

$$\nu = \tfrac{1}{2}(n_b - 4). \tag{3}$$

This relationship scales ν to <2% disagreement over the range of helix lengths explored. One of the extrapolated curves corresponds to five turns of double helix, whereas the other is an extrapolation to an intermediate distance of 18 base pairs.

QUALITATIVE NATURE OF THE DYNAMIC MODEL

Meselson (1972) originally treated the problem of the dynamic mechanism of branch point migration. He performed a hydrodynamic calculation that showed that rotary diffusion is certainly fast enough to drive the process. Meselson's calculated diffusion constant was a

slowest-case analysis, which included the entire double length of bacteriophage lambda DNA (100 kb) involved in forming the recombinant complex. It is physically more meaningful to calculate the diffusion constant corresponding to one torsional correlation length of DNA, which is 580 Å (Allison and Schurr, 1979). This length will be invariant to the size of the genome being studied (neglecting end effects). Using the torsional correlation length rather than the genome length yields a step rate of several hundred nanoseconds; this is much faster than any observed rate (Courey and Wang, 1983; Warner, R. C., personal communication).

For the junction to migrate, two base pairs must break and two new ones must form. The process of base pair opening may be rate limiting at low temperatures, as suggested by early studies (Warner et al., 1978), but it is not likely to determine the kinetics at higher temperatures, because of its high enthalpy of activation (vide infra). Clearly some further process is involved in slowing down branch point migration at higher temperatures. Here, we suggest a possible structural basis for the observed kinetics at all temperatures.

The dynamic and structural results of the previous section, combined with molecular model building, suggest the involvement of flexural diffusion in the dynamic mechanism of branch point migration. From a mechanical viewpoint, rotary diffusion should be an effective driving force for structures with unperturbed helix axes (UHA, see Fig. 1 *b*), since the helix axes and a fortiori the bases are parallel. However, the double helical segments of fourfold symmetric (S4; see Fig. 1 *a*) conformations do not fulfill this condition, even though they appear to be favored on electrostatic grounds (see Fig. 6). Molecular models indicate that the migrating junction structure must closely resemble the UHA structure; otherwise, rotary diffusion forces will not be effectively transduced into migratory forces. Rather, they will be dissipated by the viscous drag of the other arms of the junction. Thus, if rotary diffusion is to drive the migratory process, an important step in the mechanism might involve the transition from S4-like structures to UHA-like structures. This could occur via diffusion of the double helices perpendicular to their axes, in directions corresponding to θ_T and θ_B.

This new model incorporates the probability of the junction structure assuming a UHA-like structure, as well as the migrational efficacy of the conformer. The electrostatic calculations described above suggest that the UHA-like structures are not readily available regions of conformation space: they are characterized by high values of $|\delta|$. Since these structures are far removed from fourfold symmetric regions of the energy map, their probability of occurrence is very low. Nevertheless, these UHA states are accessible to the system via the process of dynamic equilibrium. The inclusion of this structurally indicated low-probability state as an obligatory intermediate decreases the rate of the migration process.

Fig. 7 shows the qualitative nature of the migratory transition state potential surface for this model. In this two-dimensional diagram, the horizontal direction, phi (ϕ), corresponds to rotation about the helix axes, while the vertical direction corresponds to differential rotation about the virtual bonds perpendicular to the helix axes, i.e., along delta (δ). A rotation of 36° in ϕ will advance the branch point by a single transition in a given direction. The figure shows conformations near fourfold symmetry associated with large barriers to this advancement; those conformations far from fourfold symmetry are shown with a much lower barrier to rotation. Meselson's single step migratory mechanism is indicated by the double arrow at the bottom of the diagram; this arrow extends horizontally from side to side, over the low barrier. We propose a model in which the most probable pathway that the reaction will take is composed of (at least) two steps: first, a step in the vertical direction to assume a UHA structure, and then a second step that involves crossing a rotation barrier in the horizontal direction. At low temperatures, this barrier includes resistance not only to rotation, but also to the opening of the base pairs. Our model states that the lowest energy pathway for migration involves the angle δ being greater than a critical minimum angle, δ_c. Since the potential energy surface is symmetric about $\delta = 0$, δ may also be less

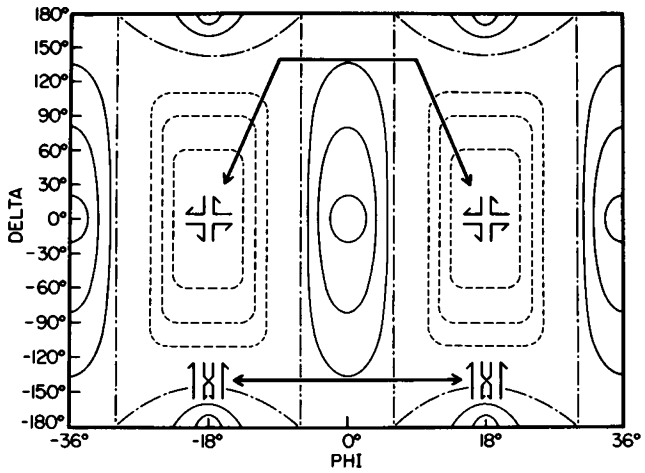
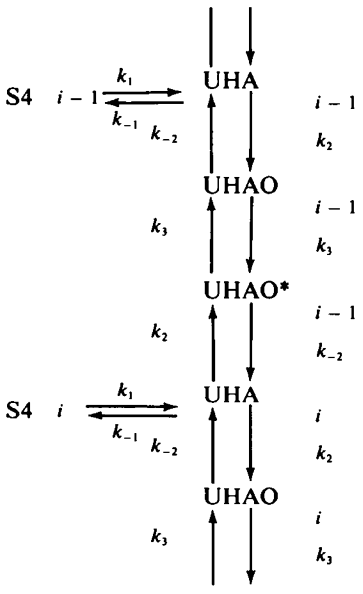
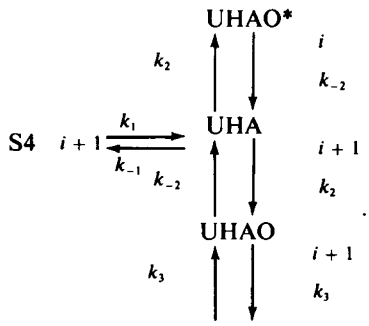


FIGURE 7 A qualitative picture of the reaction coordinate in two dimensions. Phi (ϕ) is the direction of migration, and delta (δ) is the deviation of the helix axes from fourfold symmetry. 36° steps in ϕ indicate a unit migrational event. An arbitrary zero of energy is indicated by the lines of alternating dashes and dots. Positions negative with respect to this energy are indicated by dashed contours, while positions positive with respect to this energy are indicated by solid contours. Note that the function is mirror symmetric about $\delta = 0$. This line corresponds to the most probable fourfold symmetric structures. Sigal–Alberts-like structures correspond to structures with δ near 180. Meselson's (1972) mechanism is indicated at the bottom of the figure: migration over a barrier between two Sigal–Alberts-like structures. The mechanism proposed in this paper is indicated by the thick line at the top of the figure: movement towards a Sigal–Alberts-like UHA structure (δ) (140°), and then movement over the barrier. The barrier between the two fourfold symmetric structures arises because the rotary diffusion forces cannot generate migrational forces unless the structure is UHA.

than $-\delta_c$. This lowest energy pathway is indicated by the doubly bent double arrow, at the top of the diagram in Fig. 7. Implicit in this pathway is the assumption that the barrier between the wells is higher than that for the pathway we have selected, as a result of the structural considerations described above. Although we have indicated this path as containing two distinct steps, for purposes of clarity, mixing the two steps is certainly permitted by the surface.

The three successive steps involved in the migratory event are shown in Fig. 1 in schematic molecular form. Below, we illustrate the process for several steps, including the associated rate constants. The first of these concerns moving from an S4-like structure to a UHA-like structure. The forward and backward rate constants for this step are k_1 and k_{-1} for the diffusion controlled flexing of the helices in the region of the junction. The second step involves base pair opening and closing to yield the intermediate UHAO (Fig. 1 *c*). The rates of opening and closing are k_2 and k_{-2}. The third step is the rotary diffusion process, which actually effects the migratory event, denoted by the rate constant k_3. Rotory migration must have the same rate in either direction if it is controlled by diffusion. At this point, the bases must close again, this time through the second reaction in reverse, which has rate constant k_{-2}. The migratory event and the closure are shown as a single step in Fig. 1, generating the product, P. In the more general case, schematized below, the position of the junction along a reference chain is denoted by the index i; i runs from 1 to the number of base pairs (N_b). The junction migrates from some initial position $[i]$, either ahead to the next position, $[i + 1]$, or backward to position $[i - 1]$. A step off the chain at either end irrevocably separates the strands and the junction vanishes. In the diagram below, the product P of Fig. 1 is replaced by two states, the migrated structure with the bases open, UHAO* $[i]$, and the product structure with the bases closed, UHA $[i + 1]$.





The first step of the reaction is a diffusion controlled flexing, related to flexural diffusion, D^f. The effective rate constants (k_1 and k_{-1}) are determined by the diffusion process itself (calculated below). Rather than treat the entire surface, we partition the material into two fractions. Only a fraction, F, of junctions deviate from fourfold symmetry sufficiently so that their δ is greater than δ_c, and only this fraction is able to migrate; the rate is proportional to this fraction. The first step generates a pre-equilibrated fraction of helices able to undergo migration. F is related to the equilibrium constant $K_1 = k_1/(k_{-1})$ so that $F = K_1/(1 + K_1)$. It should be emphasized that δ_c is an adjustable parameter, unfixed by any known statistical or dynamic parameters describing DNA or its environment; however, δ_c can be derived from migration rates, as noted below. δ_c is depicted in Fig. 7, where it corresponds to the value of δ at the horizontal leg of the three-step trajectory. The second and third steps are then considered independently of the first step, as a two-step mechanism. The stepping is controlled by the rotary diffusion, D^r, and k_3 is proportional to D^r. If the opening and closing rates (k_2 and k_{-2}) are very fast, then k_3 is the rate-limiting step; otherwise k_2 and k_{-2} dominate the reaction rate. K_2 ($=k_2/k_{-2}$) is the equilibrium constant for this step. The rate constants (and related equilibrium constants) for processes 1 and 3 are calculated (in the next section) using only hydrodynamic data and the results of the electrostatic calculations. k_2 and k_{-2} are phenomenological.

We have compared the results of calculations using this model (as shown in Fig. 8) with the single-step model of runoff used by Thompson et al. (1976). These calculations agree with the formulas that we now present. The three-step migration model may be used to calculate the fraction of junction (f) remaining at time t. Thompson et al. described the disappearance of junctions as runoff (RO) and defined RO = 0.5 $(1 - f)$. The rate constant (R_d) describing the disappearance of junctions, or the rate of runoff, is computed from f or RO as:

$$R_d = -\frac{d \ln f}{dt} = -\frac{d \ln (1 - 2RO)}{dt}.$$

R_d, in general, may then be computed as a function of time or of f. If the reaction were first order, then f would be exponential and R_d would be time independent. Thompson et al. related the rate of disappearance of junctions

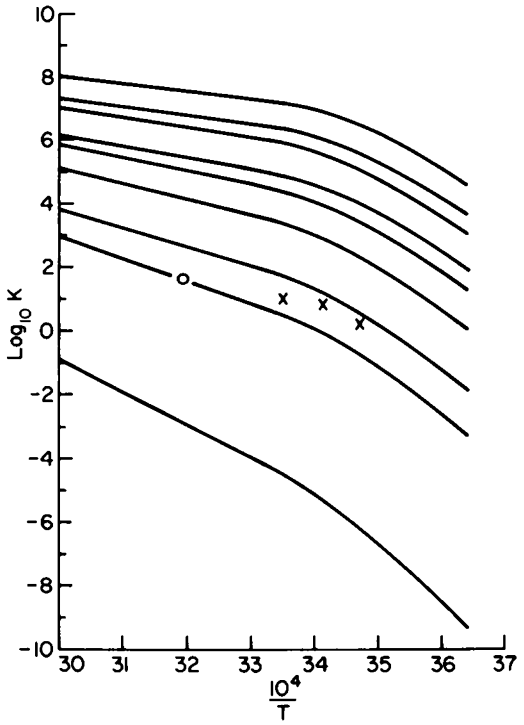


FIGURE 8 Calculated Arrhenius plots for various values of Y. The curves represent calculated Arrhenius plots that span a reasonable range of the unspecified variables ν and δ_c. From the top, the curves correspond to (ν, δ_c, Y) values of: 4,120,2.77; 4,140,4.29; 8,120,5.55; 4,160,7.00; 8,140,8.58; 16,120,11.09; 8,160,14.00; 16,140,17.16; 16,160,28.01. The points corresponding to the observations of Courey and Wang (1983) are represented by an "X", and the point due to Warner (personal communication) is represented by an "O." It can be seen that the observed points fit within the range of curves specified by the model. The curves corresponding to lower rates, and higher values of Y are below those with higher rates and lower values of Y.

(runoff) to the fundamental step rate of migration (called k_{tcw}), assuming a one-dimensional random walk. The runoff is nearly exponential in time when early times are neglected. From the model of Thompson et al. (see Appendix A), it may be seen that in this long-time region the runoff rate, called RO, is related to the junction migration rate by the relation

$$RO = C(RO)\, k_{tcw}\, (n_b)^{-2}, \qquad (4a)$$

where C is an empirical function. A plot of C as a function of RO is shown in Fig. 9. $C = 10$, for RO >50%, but in the region where the runoff does not obey first-order kinetics, C is larger, e.g., $C = 32$ at 10% runoff. We have used Eq. 4 to compare, at 50% runoff, calculations of R_d from our three-step model and from the single-step model of Thompson et al.

To compare the predictions of our dynamic model with other models, it is necessary to obtain an approximation to k_{tcw} using the mechanism described here. The results of the analysis of step one suggest that this step will pre-equilibrate the active and inactive populations. The rate constant for step 1, with this approximation, will then be

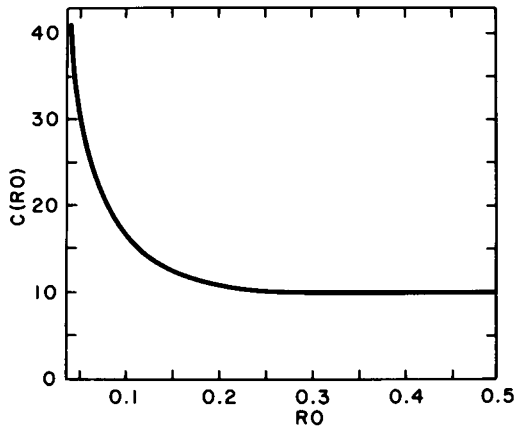


FIGURE 9 The dependence of the empirical constant, C, on the percentage of runoff. The runoff, RO, is described in Eq. 4a as RO = C(RO) k_{tcw} $(n_b)^{-2}$, where n_b is the number of base pairs, and C(RO) is an empirical function calculated from equations in the Appendix for the model used by Thompon et al. (1976). This figure shows the function C(RO).

proportional to F, the fraction of junctions able to migrate. With this assumption, the opening and closing step and the rotary diffusion step can be compared with the steady state for diffusion-controlled reactions, in which the steady-state rate constant for a reaction that contains a diffusion-controlled process with a rate constant (k_3) and an intrinsic reaction with a rate constant ($k_2 = k_{-2}$) is: $k_3\ k_2/(k_3 + k_2)$. Therefore, the overall rate constant is k_0:

$$k_0 = 0.5\, F\, k_3\, k_2/(k_3 + k_2), \qquad (4b)$$

where 0.5 is a constant of proportionality. The detailed nature of the dependence of the runoff rate on the different rate constants can only be determined by detailed calculation. Appendix A shows the equations used. In Table I, we compared k_{tcw} and k_0 for a range of rate constants for two 100 base-pair strands. The competition between processes two and three are seen to be similar to the above expression and do not depend simply on the equilibrium constant for step 2. There are some more subtle effects of k_1 and k_{-1} in the model. As shown in Table I, when the rate constants for the first process are near those of the other two steps, the runoff rate changes. The rate constants for step 1 markedly change the early time runoff rates (not shown). From Table I it is clear that k_0 is a reasonably

TABLE I

k_1	k_{-1}	k_2	k_{-2}	k_3	R_d	k_{tcw}	k_0
$1 \cdot 10^3$	$1 \cdot 10^4$	$1 \cdot 10^6$	$1 \cdot 10^6$	$1 \cdot 10^6$	26	$26 \cdot 10^3$	$23 \cdot 10^3$
$1 \cdot 10^2$	$1 \cdot 10^3$	$1 \cdot 10^6$	$1 \cdot 10^6$	$1 \cdot 10^6$	24	$24 \cdot 10^3$	$23 \cdot 10^3$
$1 \cdot 10^3$	$1 \cdot 10^4$	$1 \cdot 10^6$	$1 \cdot 10^6$	$5 \cdot 10^6$	36	$36 \cdot 10^3$	$41 \cdot 10^3$
$1 \cdot 10^2$	$1 \cdot 10^3$	$5 \cdot 10^6$	$5 \cdot 10^6$	$1 \cdot 10^6$	41	$41 \cdot 10^3$	$41 \cdot 10^3$

Computations of R_d were done for two 100 base-pair lengths of DNA, at 50% runoff. R_d was computed from equations in the Appendix. k_{tcw} is computed from R_d using Eq. 4a, suitably rearranged. k_0 is computed from Eq. 4b.

close approximation to k_{tcw}. Therefore, we use k_0 in succeeding calculations.

ESTIMATION OF RATE CONSTANTS FOR THE DYNAMIC MODEL

The fundamental step rate k_{tcw} and the temperature dependence of this rate constant were calculated based on the five independent rate constants given above. We now describe how k_1, k_{-1}, and k_3 are calculated from the hydrodynamic properties of DNA and the electrostatic repulsion energies of the helices in the region of the junction.

The rate constants for the third step and first step depend on the torsional diffusion constant about the helix axes, D^r, and the flexural diffusion constant perpendicular to the helix axes, D^f. The values for D^r and D^f are estimated from the diffusion coefficients for the rotation of two noninteracting helices, one persistence length long and freely hinged at the junction. The flexural persistence length is ~600 Å (Barkley and Zimm, 1979), while the torsional persistence length is 580 Å (Allison and Schurr, 1979). The radius of the individual helices has been taken as 12 Å. The diffusion coefficient D^f is related to the effective friction factor, f by

$$D^f = kT/f, \tag{5}$$

where k is the Boltzmann constant and T is the absolute temperature. The effective friction factor is given in terms of f_r^+ and f_t^+, which are the friction factors for pure rotation and pure translation for a cylinder moving, respectively, about and along an axis perpendicular to the helix axis. The relation is

$$f = \left[f_r^+ + \frac{L_p}{2} f_t^+ \right]. \tag{6}$$

The factor of 2 accounts for two helices in a single unit, which are hydrodynamically noninteracting to a first approximation (Schurr, J. M., personal communication). L_p is the persistence length of double helical DNA, 600 Å. The translational term f_t^+ contributes to the frictional factor, since the rotation of a helix about its end at the junction is coupled to a translation. The terms f_r^+ and f_t^+ are calculated by Broersma's (1960) equations. At 25°C, the flexural diffusion coefficient $D^f = 7{,}500$ rad^2-s^{-1}. Since this diffusion coefficient is for a full persistence length, and since direct hydrodynamic interaction has been neglected, this is the smallest value expected. The torsional diffusion coefficient D^r is 33,000 rad^2-s^{-1} (Barkley and Zimm, 1979).

Eq. 2a provides an analytic form for the equilibrium probability distribution of conformations as a function of the angle, δ, with a single adjustable parameter. The equilibrium distribution is related to the electrostatically generated restoring potential $U(\delta)$ as:

$$P(\delta) = Z^{-1} \exp\left[-\beta U(\delta)\right], \tag{7a}$$

where $\beta = 1/kT$. Since we require that $P(\delta)$ be normalized to unity, Z is the partition function. By comparison with Eq. 2a, we may, with no loss of generality, set

$$a = 1/Z \tag{7b}$$

and

$$-\beta U(\delta) = \nu \ln\left[\cos(\delta/2)\right]. \tag{7c}$$

We use the one-dimensional Smoluchowski equation (Chandrasekhar, 1943), which incorporates the restoring potential, $U(\delta)$, to describe the diffusion driven relaxation to equilibrium of the distribution of conformations, $P(\delta, t)$:

$$\frac{d\,P(\delta, t)}{dt} = D^f \frac{d}{d\delta}\left[e^{-\beta U(\delta)} \frac{d}{d\delta}\left(e^{+\beta U(\delta)} P(\delta, t) \right) \right]. \tag{8}$$

The above expression considers only flexure and does not require a sine weighting. D^f is the Einstein diffusion coefficient for flexure given by Eq. 5. This equation can be written in terms of the torque, $Q(\delta)$, which is generated by the potential $U(\delta)$.

$$Q(\delta) = -\frac{dU(\delta)}{d\delta}, \tag{9}$$

and the distribution relative to the equilibrium distribution, $\rho(\delta, t)$, which is

$$\rho(\delta, t) = P(\delta, t)/P(\delta). \tag{10}$$

Substitution of these definitions and algebraic manipulation yields an equation for $\rho(\delta, t)$:

$$\frac{d\rho}{dt}(\delta, t) = A \cdot \rho(\delta, t), \tag{11a}$$

where A is the differential operator:

$$A = D^f\left[\beta Q(\delta) + \frac{d^2}{d\delta^2} \right]. \tag{11b}$$

Eq. 11a has the formal solution:

$$\rho(\delta, t) = e^{At}\, \rho(\delta, 0). \tag{11c}$$

From the relationship of Eq. 7 and 9, we obtain the relation that

$$\beta Q(\delta) = -\nu/2 \tan(\delta/2). \tag{11d}$$

$Q(\delta)$ may be interpreted as the longitudinal torque on the helices, generated by the electrostatic potential. This should not be confused with the productive torque about the helix axes. The longitudinal torque is generated from the conformation energy profile calculated above, and is just another way of expressing the probability distribution given in Eq. 2a.

Since the description of motion is symmetric about $\delta = 0$, let us consider only those conformations with positive δ. We have postulated that conformers with δ angles close to

the UHA structure will undergo branch migration, but conformers with δ angles less than some critical cutoff angle, δ_c, will be more like S4 and hence will not undergo branch migration. Clearly, the time that the structures are in the UHA-like state is relevant to the overall rate of branch migration. Therefore, we wish to estimate the time that it takes for a conformation with the angle δ greater than the cutoff angle δ_c to move (by diffusion) to a position δ less than the cutoff angle δ_c. We estimate this time as the "mean first passage time" (Montroll and Shuler, 1958), t^*. Thus, we are interested in knowing the rate at which all conformations with $\delta > \delta_c$ become conformations with $\delta < \delta_c$. This problem is solved in the following way: We consider only those conformations with $\delta > \delta_c$ to be at equilibrium, $\rho[\delta, (t = 0)] = 1$, and require that the system evolve according to Eq. 11. In addition to a reflective boundary condition at $\delta = 180°$, we include an absorptive boundary condition at $\delta = \delta_c$.

The form of Eq. 11 is a partial differential equation in the variable δ, and $\rho(\delta, t)$ is a function of δ. However, we can make the continuous variable δ discrete, and then the differential operator, A, becomes a matrix, **A**, and the function $\rho(\delta, t)$ becomes a vector. For example, the operation

$$\frac{d^2\rho(\delta, t)}{d\delta^2}$$

becomes

$$\frac{1}{h^2}[\rho(t)_{i+1} + \rho(t)_{i-1} - 2\rho(t)_i],$$

where h is the distance between successive of δ, and $\delta = i \cdot h$. The index i marks the ith position of $\rho(\delta, t)$ in the $\boldsymbol{\rho}(t)$ vector. The matrix **A** now has the appearance of a series of kinetically coupled equations. This approximation to the differential equation is called the finite difference (see Scheid, 1968) or transition matrix approximation. In this form, it is straightforward to set both reflective and absorptive boundary conditions by adjusting the **A** matrix.

With this understanding, the rate, $R^*(t)$, at which conformers leave the system is:

$$R^*(t) = \frac{\int_{\delta-\delta_c}^{\pi} \frac{dP(\delta, t)}{dt} d\delta}{\int_{\delta-\delta_c}^{\pi} P(\delta) d\delta} = \frac{\int_{\delta-\delta_c}^{\pi} P(\delta) \frac{d\rho}{dt}(\delta, t) d\delta}{\int_{\delta-\delta_c}^{\pi} P(\delta)\rho(\delta, 0) d\delta} = \frac{\langle \mathbf{P}_0 | \dot{\boldsymbol{\rho}}(t) \rangle}{\langle \mathbf{P}_0 | \boldsymbol{\rho}(0) \rangle}. \quad (12)$$

The mean first passage time is then computed from the rate, $R^*(t)$ as

$$t^* = \int_{t=0}^{\infty} tR^*(t) dt. \quad (13)$$

As a simple example, if $P(\delta, t)$ obeyed a simple first-order decay, then the rate would be exponentially decaying and the mean first passage time would be the reciprocal of the rate constant for the decay.

If we substitute the definition of $\dot{\rho}(\delta, t)$ [$= d\rho(\delta, t)/dt$] from Eq. 11 into the expression for t^*, and evaluate the time integral, we find (after some algebraic manipulation) that:

$$t^* = \frac{\langle \mathbf{P}_0 | \mathbf{A}^{-1} | \boldsymbol{\rho}(0) \rangle}{\langle \mathbf{P}_0 | \boldsymbol{\rho}(0) \rangle}. \quad (14)$$

The above expression is used to directly calculate t^*, where $\mathbf{A}^{-1}$ is the matrix inverse to the **A** matrix, and $\mathbf{P}_0$ and $\boldsymbol{\rho}(0)$ are the vector approximations to $P(\delta)$ and $\rho(\delta, 0)$, where $\rho(\delta, 0)$ equals unity by definition (see Eq. 10).

These calculations give $k_{-1} = 1/t^*$ and the equilibrium constant K_1 is calculated from the equilibrium distribution. The rotary diffusion constant (D^r) is used to estimate the effective diffusion rate: $k_3 = D^r/\langle\phi^2\rangle$ where $\langle\phi^2\rangle = \phi_0^2/6$, and $\phi_0 = 36°$.

RESULTS OF THE DYNAMIC CALCULATIONS

The rates discussed below are k_0 at 50% runoff and are shown in Fig. 8 as a function of temperature. These have been determined from the effective rate at that point using the equations in Appendix A. The temperature dependence of the rate constant k_3 is proportional to the viscosity/temperature. The electrostatic potential U is inversely proportional to the dielectric constant of the medium. The temperature dependence of U comes from the temperature dependence of the dielectric constant. The temperature dependence of the viscosity coefficient and dielectric constant of aqueous 0.01 mM NaCl are well known (Handbook of Chemistry and Physics, 1981). The rate constant k_2 was given a value of $9 \times 10^8\ s^{-1}$ at 20°C and enthalpy of 50 kcal/mol (via an Arrhenius dependence on the temperature) to simulate a low-temperature process, which is base-pair-opening limited below 20°C. This was suggested by early data (Warner et al., 1978) and has not been refuted (Warner, R. C., personal communication) to our knowledge. The forward and backward rates (k_2 and k_{-2}) were taken to be the same (Mandal et al., 1979; Preisler et al., 1981). The value of F in Eq. 4 b was calculated directly from $P(\delta)$ in Eq. 2a, for a fixed value of δ_c and ν. Since the temperature dependence of $U(\delta)$ is known, Eq. 7c gives the temperature dependence of ν (for a fixed δ). The values of ν reported are at 20°C.

The model we have presented contains two free parameters: ν, the effective length of helix for which electrostatic interactions must be accounted, and δ_c, the flexural cutoff angle, below which the migratory event will not occur. It is instructive to examine the behavior predicted by variation of these two parameters. Fig. 8 shows the Arrhenius plots corresponding to particular reasonable values of these

parameters. The curves are all biphasic, because at lower temperatures the opening of the base pairs has been suggested to be the rate-limiting step (Warner et al., 1978). This step, which has a very high temperature dependence (Bird et al., 1970), is expected to stop being rate limiting at higher temperatures, and the rest of this discussion concerns the higher temperature domain of the kinetic process.

Inherent in the formulation of this model is the linear variation of the enthalpy and entrophy of activation with Y, where

$$Y = -\nu \ln [\cos (\delta_c/2)].$$

Inspection of Eq. 7c would suggest that a quantity of this sort might behave as a single independent parameter for ΔH and ΔS in the high temperature region. The variation of these quantities with Y are shown in Fig. 10. An important prediction of the model is that a particular value of the activation enthalpy corresponds to a given value of the activation entropy. It should be noted that the series of curves shown in Fig. 8 is also a series in increasing Y. Y may be taken to be a measure of the difficulty of achieving the effective UHA structure, and hence of undergoing the migratory event. As noted above, Y has two components, the effective extent of double helix, which is involved in electrostatic repulsive interactions, and the cutoff angle, which must be achieved before migration can occur. Longer regions of interaction result in lower rates; similarly, a higher cutoff angle will slow the rate down. It is worth noting that the electrostatic calculation, which assumed rigid helices, did not account for systematic bending of the helices, which would occur when δ becomes large, and would, of course, vanish at $\delta = 0$. Electrostatic repulsion is likely to force the helices to be somewhat distorted from linearity, so the effective length corresponding to ν would correspond to a longer actual length.

Numerous studies in migrating branch points have been conducted. Often, however, the studies are combined with cruciform formation (Lilley and Hallam, 1984; Gellert et al., 1983), and there are extra steps involved in that process. Courey and Wang (1983) have made a few measurements on the relaxation of cruciforms, and Warner and his colleagues made some early measurements on the branch point migration of bacteriophage G4 replication intermediates, which are no longer accepted (Warner, R. C., personal communication). Courney and Wang's observations along with a later value of Warner's (1984) have been plotted in Fig. 8. It is to be expected that future systematic measurements with semi-mobile junctions (Seeman, 1981, 1982) will soon fill in the gaps in these observations. It can be seen that all of the observed rates fit in a range of Y between 14 and 17. This corresponds to an activation enthalpy between 11.5 and 13.6 kcal/mol, and an activation entropy between 0.042 and 0.046 kcal/mol-deg. It should be emphasized that both the flexural and torsional diffusion contribute 5.5 kcal/mol activation enthalpy apiece, while the rest derives from the electrostatic repulsion in the model. Thus, the model we have proposed can emulate the kinetics reasonably well.

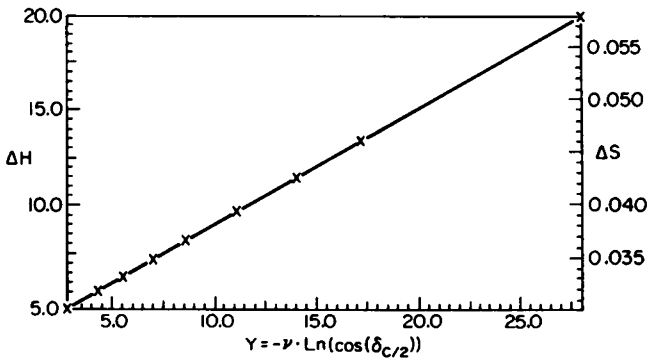


FIGURE 10 The dependence of the activation enthalpy and activation entropy on Y. Since both quantities are linear functions of Y, they are represented on the same plot. The scale for the enthalpy is on the left, in kcal/mol, and the scale for the entropy is on the right, in kcal/mol-deg. The points indicated on the plot correspond to the individual curves of Fig. 8, with their particular values of ν, δ_c, and Y.

DISCUSSION

The Mechanism of Migration

The above results indicate that the electrostatic and three-step dynamic model qualitatively emulates double-stranded branch point migration. The extrema of the branch point conformation range are of substantially different electrostatic energies, and the transitions among them appear to be critical factors in the dynamics of the high-temperature migratory process. It is interesting that the two major structures previously proposed viz the Sigal–Alberts UHA structure analogous to Fig. 1 *b*, and the Sobell S4 structure analogous to Fig. 1 *a*, are both representative of conformational classes relevant to the migratory process. Those structures analogous to Fig. 1 *a* appear to be the most probable structures from the standpoint of electrostatics; on the other hand, our structural arguments suggest that those same conformers are least likely to be involved in the migratory process. Rather, those structures furthest from fourfold symmetry, those with unperturbed (or minimally perturbed) helix axes, appear to be critical intermediates in the migratory process, corresponding to those structures with $\delta > \delta_c$.

Meselson's notion that rotary diffusion drives the actual migratory process has been the basis for this model, and is not contradicted by the existing data. However, the data indicate that rotary diffusion alone is probably not rate limiting. We have suggested that a process such as diffusion perpendicular to the helix axes contributes to the rate-limiting process. We have treated the physics of this model and have obtained satisfactory agreement with experiment. In the absence of further data, it is impossible to exclude the possibility that another process with a

different structural and physical basis could be rate limiting. Since we have ignored the effects of base stacking and of detailed junction conformational states, we must be very careful in associating a detailed structural basis with the kinetics.

Base Pair Mismatches in Homologous DNA

Our calculations bear on the kinetics of recombination of homologous DNAs, as illustrated by the relative rate constants of reaction I and reaction II shown in Fig. 1. If the completely paired intermediates shown in Fig. 1 are presumed to be obligatory, which is reasonable on stereochemical grounds, then the relative kinetics of reactions I and II (Fig 1) will be different for identical and homologous double helices: In the case of identical double helices, an equal probability is to be expected for reaction I or reaction II at each transition, while this is not necessarily true for homologous double helices. When a mismatch occurs, the system is moving (by reaction II) from a state containing two Watson–Crick base pairs (bridging arms [1-2] and [3-4]), to a state having no Watson–Crick base pairs stabilizing chains (1-4) and (2-3) at the junction point. This would generate a barrier to reaction II, thereby introducing a large bias in favor of reaction I. At the right side of Fig. 3 *b*, we have indicated the nature of the energetics when a mismatch is encountered. There is a raised region, in which the relative minimum is higher than in the rest of the diagram, but the barriers between the minima are not higher. The kinetic barriers on either side of the mismatch will be unaffected, but the equilibrium will be shifted in favor of the direction without the mismatch. The extent of this shift is a function of the step size. A mismatch will affect the relative stability of a long length of DNA much less than the stability of a very short length.

Had they turned out large enough, the productive torques might have provided the necessary driving force for branch point migration in the presence of mispairing. It is a priori conceivable that mispairing could be kinetically overcome by an averaging mechanism that jumped the system from one side of the mismatch trauma to the other, thereby increasing the step size. A large electrostatic driving force that involved a great many multiple steps would fall into this category, hurling the state of the system from the left side of Fig. 3 *b* to the far right. However, the productive electrostatic torques inherent in the isolated system are trivial on the scale needed to accomplish this task.

In a thermodynamically equivalent variation of this polystep scheme, an unraveling–reraveling mechanism of migration involving many bases could also effect this jump. The vanishing of the productive torques about the helix axes on the symmetry diagonal ($\delta = 0$) results from the cancellation of two equal and opposite forces at this highly probable set of conformations. One must ask, therefore, whether the magnitudes of these cancelling forces are indeed large enough to contribute to an unraveling mechanism for the accommodation of mispairing, as discussed above. They are not: these forces are extremely small, of the order of 0.010 or 0.020 kcal/mol-rad, at the vanishing points. Thus, we feel that polystep mechanisms, independent of external energy sources, involving unraveling are not likely to have a significant electrostatic component to them.

The results of the dynamic calculations permit us to estimate the number of steps a junction will migrate, before relaxing to a state with $\delta < \delta_c$. The total number of steps taken depends on the time during which the helices are correctly positioned, and the rate at which the migration can proceed by torsional diffusion. The root mean square average number of steps (N) taken by migrating molecules is

$$N = (k_3 t^*)^{1/2}, \tag{15}$$

where t^* is the mean time during which those conformations with $\delta > \delta_c$ will maintain themselves in this migrating state. Our calculations show that the mean lifetime of the UHA state allows an average of two to three base pairs to migrate before the flexual motion drives the structure to a nonmigrating conformation. The model suggests that multiple transitions may occur, but large polystep transitions are not expected. Thus, this model does not provide any mechanism for rapidly moving across mismatches. If recombination of strands is to be an effective mechanism for exchanging genetic information among homologous DNAs then energy transducing enzymes must be active during the process.

APPENDIX

From the kinetic scheme shown in the text the kinetic equations are:

$$\begin{aligned}
(\dot{\mathrm{UHA}})_i &= -(2k_2 + k_1)(\mathrm{UHA})_i \\
&\quad + k_2(\mathrm{UHA}_i^+ + \mathrm{UHA}_i^{+*}) + k_1(\mathrm{S4}) \\
(\dot{\mathrm{UHA}}^+)_i &= -(k_2 + \mathrm{k}_3)(\mathrm{UHA}^+)_i \\
&\quad + k_3(\mathrm{UHA}^{+*})i + 1 + \mathrm{k}_2(\mathrm{UHA})_i \\
(\dot{\mathrm{UHA}}^{+*})_i &= -(k_2 + k_3)(\mathrm{UHA}^{+*})_i \\
&\quad + k_2(\mathrm{UHA})_i + \mathrm{k}_3(\mathrm{UHA}^+)_{i-1} \\
(\dot{\mathrm{S4}})_i &= -k_i(\mathrm{S4})_1 + k_1(\mathrm{UHA})_i
\end{aligned}$$

$$\text{Let } \mathbf{v}_i = \begin{pmatrix} \mathrm{UHA}_i^{+*} \\ \mathrm{UHA}_i \\ \mathrm{S4}_i \\ \mathrm{UHA}^+ \end{pmatrix}$$

$$\alpha = \begin{pmatrix} -(k_3 + k_2) & k_2 & 0 & 0 \\ k_2 & -(2k_2 + k_1) & k_1 & k_2 \\ 0 & k_1 & -k_1 & 0 \\ 0 & k_2 & 0 & -(k_3 + k) \end{pmatrix}$$

$$\beta = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ k_3 & 0 & 0 & 0 \end{pmatrix}.$$

The total diffusion problem then is written

$$\begin{pmatrix} \dot{V}_1 \\ V_2 \\ \vdots \\ V_N \end{pmatrix} = \begin{pmatrix} \alpha & \beta & 0 & \cdot & \cdot & \cdot \\ \beta^t & \alpha & \beta & 0 & \cdot & \cdot \\ 0 & \beta^t & \alpha & \beta & 0 & \cdot \\ \cdot & \cdot & \beta^t & \alpha & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{pmatrix},$$

which has the general form

$$\dot{V} = \mathbf{A}\, V.$$

The solution is $V(t) = \exp(\mathbf{A}t)\, V(0)$.

The fraction left after time t is $F(t) = (\epsilon^t V(t)/\epsilon^t v(0))$, the total runoff is $1 - F(t)$.

$$\epsilon = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ \cdot \end{pmatrix}.$$

We solve this problem by an evolution of V over the small time interval τ:

$$V(t + \tau) = e^{\mathbf{A}\tau}\, V(t) \text{ when } \| \mathbf{A}\tau \| < 1$$

$$V(t + \tau) = (1 + \tau\, \mathbf{A})\, V(t)$$

$$\mathbf{V}_i(t + \tau) = (1 + \tau\alpha)\, \mathbf{V}_i(t) + \tau\beta\, \mathbf{V}_{i+1}(t) + \tau\beta^t\, \mathbf{V}_{i-1}(t).$$

The initial distribution of V_i at $t = 0$ is found at $\alpha|_{k_3=0}\, v = 0$ and $\epsilon^t v = 1$.

In the case of the two-jump model, the above formalism requires

$$\alpha = -2\, k_{tcw} \qquad \beta = k_{tcw}$$

and then

$$V_i(t + \tau) = (1 - 2\tau k)\, V_i + (\tau k)[V_i(t) + V_{i-1}(t)].$$

Using the criterion that $\| \tau\alpha \| = 1$ gives

$$\tau = \tfrac{1}{2} k_{tcw}$$

and then

$$V_i(t + \tau) = \tfrac{1}{2}(V_{i+1}(t) + V_{i-1}(t)),$$

which is the one-dimensional random walk model, used by Thompson et al. (1976).

We are grateful to Drs. Robert Warner and Richard Fishel for prepublication materials. We would like to thank Dr. J. M. Schurr for valuable discussions and information about torsional correlation lengths and Dr. Neville R. Kallenbach for many helpful suggestions. We would also like to thank Ms. Kathleen A. McDonough for assistance in model building. Ms. Diane B. Robinson and Mr. Ryland Loos drew the figures, and Mr. Robert Speck made the photographs. Computational facilities were kindly provided by the SUNY/Albany Computer Center and by the University of Washington Computer Center.

This work has been supported by a Basil O'Connor Starter Grant from The March of Dimes Birth Defects Foundation, Grants GM-26467, GM-29554, and ES-00117 from the National Institutes of Health (NIH), and Grant PCM-8216762 from the National Science Foundation. N. C. Seeman is a recipient of an NIH Research Career Development Award.

Received for publication 19 September 1986 and in final form 10 December 1986.

REFERENCES

Allison, S. A., and J. M. Schurr. 1979. Torsion dynamics and depolarization of fluorescence of linear macromolecules I. Theory and application to DNA. *Chem. Phys.* 41:35–46.

Arnott, S., and D. W. L. Hukins. 1972. Optimised parameters of A-DNa and B-DNA. *Biochem. Biophys. Res. Commun.* 47:1504–1509.

Barkley, M. D., and B. H. Zimm. 1979. Theory of twisting and bending of chain macromolecules; analysis of the fluorescence depolarization of DNA *J. Chem. Phys.* 70:2991–3007.

Bird, R. E., K. G. Lark, B. Curnutte, and J. E. Mansfield. 1970. Exchange of bound tritium between DNA and water at different temperatures. *Nature (Lond.).* 225:1043–1045.

Broersma, S. 1960. Rotational diffusion constant of a cylindrical particle. *J. Chem. Phys.* 32:1626–1631.

Broker, T., and I. R. Lehman. 1971. Branched DNA molecules: intermediates in T4 recombination. *J. Mol. Biol.* 60:131–149.

Chandrasekhar, S. 1943. Stochastic problems in physics and astronomy. *Rev. Mod. Phys.* 15:1–89.

Courey, A. J., and J. C. Wang. 1983. Cruciform formation in a negatively supercoiled DNA may be kinetically forbidden under physiological conditions. *Cell.* 33:817–829.

Gellert, M., K. Mizuuchi, M. H. O'Dea, H. Ohmori, and N. Tomizawa. 1978. DNA gyrase and DNA supercoiling. *Cold Spring Harbor Symp. Quant. Biol.* 43:33–40.

Gellert, M., M. H. O'Dea, and K. Mizuuchi. 1983. Slow cruciform transitions in palidromic DNA. *Proc. Natl. Acad. Sci. USA.* 80:5545–5549.

Handbook of Chemistry and Physics. 1981. 63rd ed. R. C. Weast, editor. CRC Press, Cleveland. E-57.

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

Kallenbach, N. R., R. I. Ma, A. J. Wand, G. H. Veeneman, J. H. vanBoom, and N. C. Seeman. 1983*a*. Fourth rank immobile nucleic acid junctions. *J. Biomol. Struct. & Dyn.* 1:159–168.

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

Kim, J. S., P. Shark, and N. Davidson. 1972. Electron microscope studies of heteroduplex DNA from a deletion mutant of bacteriophase 0X-174. *Proc. Natl. Acad. Sci. USA.* 69:1948–1952.

Lee, C. S., R. W. Davis, and N. Davidson. 1970. A physical study by

electron microscopy of the terminally repetitious, circularly permuted DNA from coliphage particles of *Escherichia coli* 15. *J. Mol. Biol.* 48:1–22.

Lilley, D. M. J. 1980. The inverted repeat as a recognizable structural feature in supercoiled DNA. *Proc. Natl. Acad. Sci. USA.* 77:6468–6472.

Lilley, D. M. J., and L. R. Hallam. 1984. Thermodynamics of the ColE1 cruciform. Comparisons between probing and topological experiments using single topoisomers. *J. Mol. Biol.* 180:179–200.

Mandal, C., N. R. Kallenbach, and S. W. Englander. 1979. Base pair opening and closing reactions in the double helix. *J. Mol. Biol.* 135:391–411.

Manning, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Quart. Rev. Biophys.* 11:179–246.

McEntee, K., G. M. Weinstock, and I. Lehman. 1979. Initiation of general recombination catalyzed *in vitro* by the *rec*A protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* 76:2615–2619.

McGavin, S. 1977. A model for the specific pairing of homologous double-stranded nucleic acid molecules during genetic recombination. *Heredity.* 39:15–25.

Meselson, M. 1972. Formation of hybrid DNA by rotary diffusion during genetic recombination. *J. Mol. Biol.* 71:795–798.

Montroll, E. W., and K. E. Schuler. 1958. The application of the theory of stochastic processes to chemical kinetics. *Adv. Chem. Phys.* 1:361–393.

Panayotatos, N., and R. D. Wells. 1981. Cruciform structures in supercoiled DNA. *Nature (Lond.).* 289:466–470.

Preisler, R. S., C. Mandal, S. W. Englander, N. R. Kallenbach, F. B. Howar, J. Frazier, and H. T. Miles. 1981. Equilibrium and kinetic characteristics of the low temperature open state in polynucleotide duplexes. *In* Biomolecular Stereodynamics. R. H. Sarma, editor. Adenine Press, New York. 405–415.

Seeman, N. C. 1981. Nucleic acid junctions: building blocks for genetic engineering in three dimensions. *In* Biomolecular Stereodynamics. R. H. Sarma, editor. Adenine Press, New York. 279–300.

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

Seeman, N. C., B. H. Robinson, and K. A. McDonough. 1979. Electrostatic considerations in branch point migration. Second Basis O'Connor Symposium, Key Biscayne, FL. (Abstr.)

Seeman, N. C., and B. H. Robinson. 1981. Simulation of double stranded branch point migration. *In* Biomolecular Stereodynamics. R. H. Sarma, editor. Adenine Press, New York. 279–300.

Scheid, F. 1968. Numerical Analysis. McGraw Hill, New York. 79–104.

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

Sobell, H. M. 1972. Molecular mechanism for genetic recombination. *Proc. Natl. Acad. Sci. USA.* 69:2483–2487.

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

Warner, R. C., R. Fishel, and F. Wheeler. 1978. Branch migration in recombination. *Cold Spring Harbor Symp. Quant. Biol.* 43:957–968.

Wilson, J. H. 1979. Nick-free formation of reciprocal heteroduplexes: A simple solution to the topological problem. *Proc. Natl. Acad. Sci. USA.* 76:3641–3645.

Wolgemuth, D. S., and M. T. Hsu. 1980. Visualization of genetic recombination intermediates of human adenovirus type 2 DNA from infected HeLa cells. *Nature (Lond.).* 287:168–171.