Cold Spring Harbor Symposia on Quantitative Biology

A Topological Treatment of Recombination and Topoisomerases

N.R. Cozzarelli, M.A. Krasnow, S.P. Gerrard, et al.

Cold Spring Harb Symp Quant Biol 1984 49: 383-400
Access the most recent version at doi:10.1101/SQB.1984.049.01.045

A Topological Treatment of Recombination and Topoisomerases

N.R. COZZARELLI,* M.A. KRASNOW,† S.P. GERRARD,* AND J.H. WHITE‡

*Department of Molecular Biology, University of California, Berkeley, California 94720; †Department of Biochemistry, University of Chicago, Chicago, Illinois 60637; ‡Department of Mathematics, University of California, Los Angeles, California 90024

A feature of several site-specific recombination systems currently under study is that they can reduce the supercoiling of the substrate and produce knots or catenanes in addition to rearranging DNA sequences. The topological changes can be analyzed to reveal specific information about the mechanism of recombination.

The change in the topology of circular DNA during supercoiling or relaxation by a topoisomerase is well-defined as a change in the linking number (Lk) of the DNA. No such index of topological structure exists to measure the topological changes made during knotting or catenation by either topoisomerases or recombination enzymes.

We describe an index called linkage (Lg) that can be applied uniformly to knots, catenanes, supercoils, and double-helical twist. It is based on counting the crossings, or nodes, in a plane projection of a DNA molecule. For unknotted single rings of DNA, Lg = Lk. For knots and catenanes, Lg also includes the number of nodes specific to knot and catenane structure that we label Kn and Ca, respectively. We also count the nodes introduced by an enzyme and label them Me for mechanism.

Simple equations then relate the linkage value of the substrate and product to the mechanism of the enzyme. We first describe reactions of topoisomerases in terms of node language. We next treat quantitatively the products of site-specific recombination reactions and develop specific conclusions about the mechanisms of resolvase, the cointegrate resolution enzyme encoded by the Tn*3* transposable element, and of Int, the λ-phage integration system.

The relationships we propose have not yet been established with mathematical rigor. They have, however, been verified with ribbon models for DNA and have been a valuable guide for our research on recombination and topoisomerases.

To make the treatment more accessible, we have separated the more technical aspects in numbered subsections and in the figure and table legends. Literature references have been minimized.

NODES AND LINKING NUMBER

Nodes are the fundamental unit by which topological properties of DNA can be measured. If DNA is represented on a flat surface, then where two segments cross they form a node. The helical twist of duplex DNA topologically links the two strands of circular DNA. The number of twists, Tw, can be determined by counting the number of single-strand crossings and dividing by 2, because there are two nodes per twist. Thus, a single-strand node is assigned a magnitude of ½. Naturally occurring duplex circular DNA generally also has writhe in the form of supercoils. In the standard planar projection, the supercoils form nodes in which one duplex segment crosses another. Although there are four single-strand nodes in such a duplex crossing, two of them are crossings of one strand over itself. Since these do not contribute to linking of the two strands, only two nodes are counted and the magnitude of a duplex DNA node is 1 (Banchoff and White 1975). The number of duplex nodes is the writhe, Wr, which also links the two strands of DNA. The topological linking number for a circular DNA is the algebraic sum of the single-strand nodes of twist and the duplex nodes of writhe (White 1969; Fuller 1971; Crick 1976), or

$$Lk = Tw + Wr. \quad (1)$$

Breaking the DNA backbone is required to change Lk.

A node is ascribed sign according to the orientation and overlay of the crossing segments. When the two segments of a node are from the same DNA molecule, the orientation of the segments can be determined by tracing the path between them. Regardless of which direction the DNA is traced, the two segments will cross only one of two distinct ways, and the node can thus be given a sign using the convention shown in Figure 1.

The sign of the single-strand nodes that make up twist is determined by orienting both strands in the same direction as the duplex DNA axis, even though they are of opposite chemical polarity. Supercoil nodes are illustrated in Figure 2, which shows a (−) and a (+) supercoil. For naturally occurring circular DNA, the twist of the helix is (+) and the writhe of the axis is (−). Such negatively supercoiled DNA has a smaller linking number than circular DNA with no writhe.

1. The linkage contributed by a node is independent of the DNA path connecting the two segments. The node in Figure 2a and the two nodes in Figure 3 that

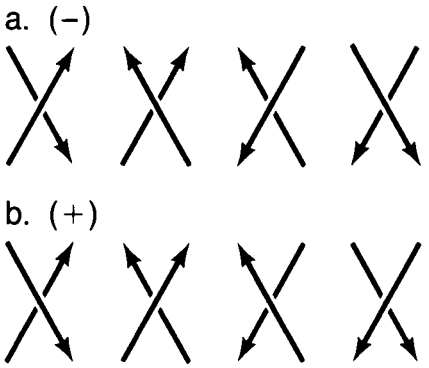

Figure 1. Sign of nodes. The arrows indicate the orientation of crossing segments of DNA where orientation is a choice of direction in which to travel along a DNA molecule. The four representations of (−) nodes in (*a*) and (+) nodes in (*b*) are equivalent and related by successive 90° rotations. We define the sign convention in terms of the rotation needed to make the segment on top congruent with the underlying segment. For angles $<180°$, clockwise and counterclockwise motion define (−) and (+) crossings, respectively. The convention can be remembered by a hand rule. With thumbs crossed over forefingers, the pointing of these fingers of the left hand describe a (−) node, whereas the crossing thumb and forefinger of the right hand trace a (+) node.

are set off by bold arrows all have a value of −1, even though there is a net positive supercoiling between the nodal segments in Figure 3a, negative supercoiling in Figure 2a, and in Figure 3b the segments come from different DNA molecules. The node is a local property dictated only by the relative orientation and overlay of the segments. Thus, the presence of nodes shows that a DNA has shape of a certain sign and complexity but not what that shape is.

2. There are two types of DNA supercoils (Bauer and Vinograd 1975). Solenoidal, or toroidal, coils (Fig. 4a) are wrapped around an external axis, as in the coursing of DNA around histones in chromatin. Plectonemic supercoils (Fig. 4b) are interwound and are interconvertible with solenoidal coils of the same sign. The supercoils in free circular DNA are plectonemic when viewed by electron microscopy and this form is thought to predominate in solution. The conceptual difficulty that (−) solenoidal coils are left-handed whereas (−) plectonemic coils are right-handed is easily explained by looking at the nodes (Fig. 4). DNA doubles back on itself when it forms plectonemic supercoils, and thus the orientation of the crossing segments is opposite to that in solenoidal supercoils. To maintain sign constant, the overlap of the crossing segments must be interchanged, which changes the hand of the helix. This example shows that handedness and sign are distinct properties.

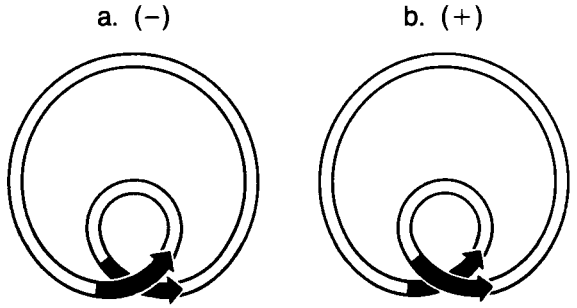

Figure 2. The sign of supercoils. The parallel lines represent the strands of closed duplex DNA that contains a (−) supercoil in (*a*) and a (+) one in (*b*). The arrows aid in determining node sign by showing orientation.

3. Unfortunately, no sign convention for nodes in DNA is free of difficulties. We recommend convention A, the one described above, but there may be circumstances where a second convention, B described below, would be superior. Whatever choice is made, one must maintain internal consistency and make sure that conclusions describe the physical situation and are not implicit in the convention.

   Using convention A to calculate the linkage due to supercoiling, one can count either the helix axis nodes or the crossings of the single-strands, designated W and C; C–C and W–W crossings are ignored. For simple rings, ignoring same-strand crossings makes physical sense because they can be removed without strand breakage, i.e., they are not required topologically. For consistency, we maintain this convention for duplex DNA knot and catenane nodes, where self nodes link just as much as nodes between complementary strands. Convention A simplifies analysis because ascribing the same orientation to the helix axis and both component strands allows single- and double-strand nodes to be described equivalently.

   The disadvantage of convention A is that it ignores the antiparallel chemical polarity of standard double helices. Convention B assigns antiparallel orientation to the W and C strands. To ensure that right-handed double helices are (+), the opposite sign convention to that shown in Figure 1 must be used. An advantage of convention B is that the nodes of a right-handed helix will be (+) for both intramolecular and intermolecular helices, whereas in convention A they are (−) for fold-back helices but (+) for standard double helices. We noted above, however, that for any convention, hand and sign need not correlate. The difficulty with convention B is that it gives the wrong sign for supercoils and knots. For example (−) supercoils (Fig. 4) would have (+) nodes by convention B, instead of the sign opposite that of the helix. It is possible to rectify the situation by adding appropriate sign change rules, but to treat all curves uniformly we use convention A.

4. The two uses of the term "supercoiling" can be distinguished by considering the special case in which the helix axis lies in a plane. Such a relaxed DNA has no writhe and its linking number, Lk°, is equal to Tw. Supercoiling can be defined as $Lk - Lk°$, the linking difference. More commonly, supercoiling is used interchangeably with writhe. The two

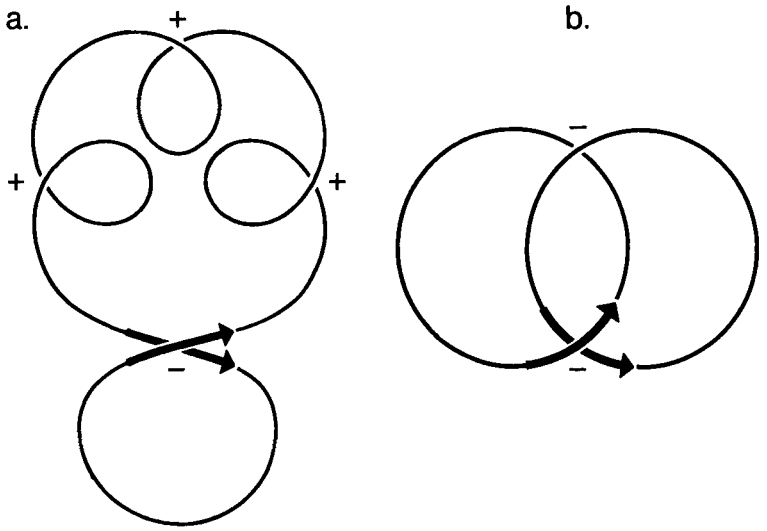

Figure 3. A node is a local property. The node shown by bold arrows is (−) in both (*a*) and (*b*), even though the net coiling between the nodal segments is (+) in (*a*) and the segments making up the node in (*b*) are from two different molecules. For simplicity, duplex DNA is shown as a single line.

uses are equivalent when Tw = Tw°. We will use Wr rigorously, but supercoiling heuristically.

5. Given a ribbon model for DNA in which the opposite edges are the W and C strands, it has been proven that for any plane projection of the model, the sum of the value ± ½ attached to each of the crossings of opposite edges, i.e., the algebraic sum of the single-strand nodes, is Lk. Lk for duplex DNA is necessarily integral because there are an even number of crossings. It is an approximation though to describe Tw and Wr by node counting because they are not necessarily integral. If the helix axis is planar, then Wr and therefore Tw will be integral. However, when one duplex segment crosses another as in supercoils (Fig. 2), the axes for both segments cannot be in the same plane. Hence, additional Tw must be introduced at the node so that one segment may rise and cross over the other. Therefore, Wr of a single supercoil may be slightly different than 1 because of the additional Tw and Equation 1. Because DNA is a long thin molecule, we make the approximation that the axis can be made essentially planar in a flattened model and thus that Tw near these supercoil nodes is negligible and both Tw and Wr are integral.

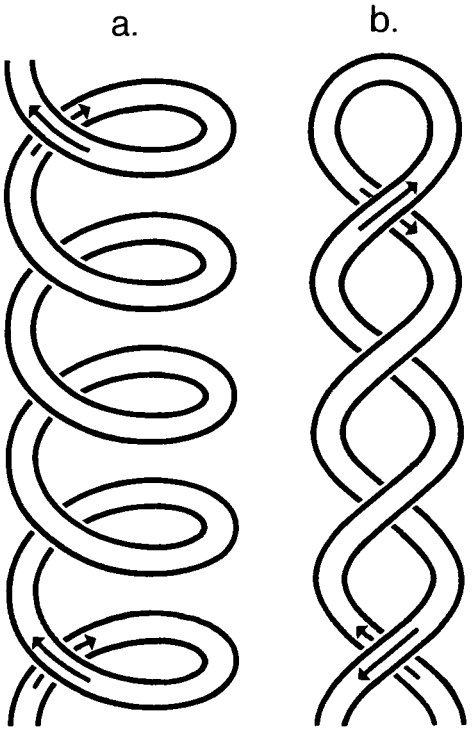

Figure 4. Sign and handedness of supercoils are distinct properties. The solenoidal supercoils in (*a*) and plectonemic supercoils in (*b*) are both (−), as can be seen from the node-defining arrows, but are left-handed and right-handed, respectively.

LINKAGE OF KNOTS

We define the intrinsic linkage of a knot, Kn, as the algebraic sum of the nodes that constitute the knot, where the sign and magnitude of the nodes are determined as above. To obtain the complete linkage value of a knotted DNA ring, Kn is added to the other forms of Wr and Tw. Conversely, Kn is the linkage value after all Wr and Tw not essential to the knot have been removed. Kn can be readily determined from a standard representation of a knot because the relative orientation of crossing intramolecular segments is apparent. For the simplest knots, trefoils, the three nodes have the same sign. For the trefoil in Figure 5a, Kn = −3, and for the mirror image knot in Figure 5b, Kn = +3.

Kn distinguishes the two possible knots with three nodes. However, there can be many knots with five or

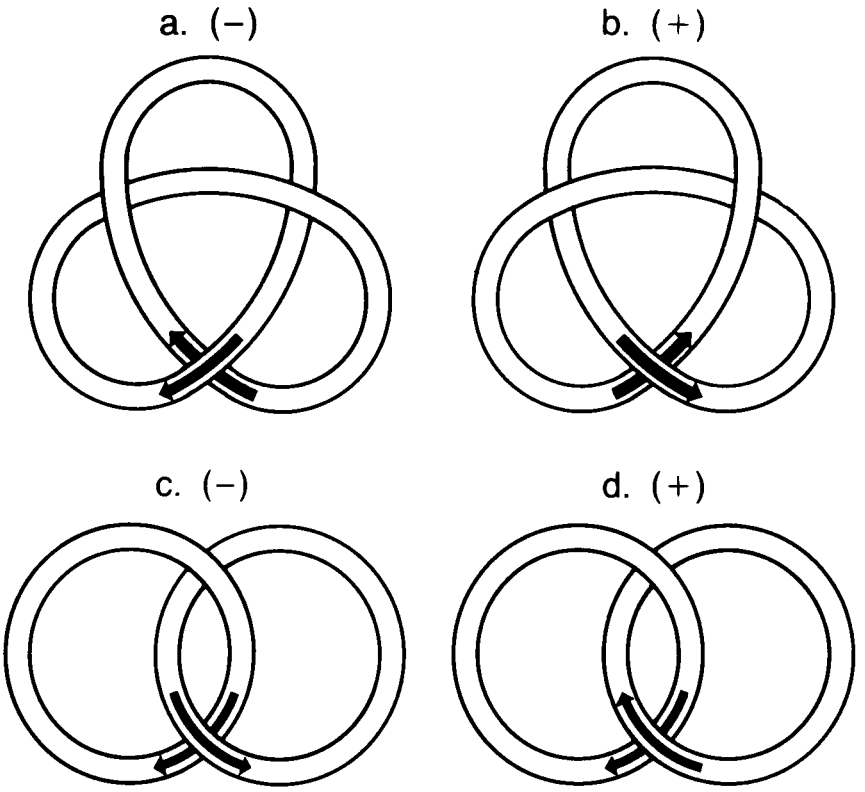

Figure 5. Sign of trefoil knots and singly interlocked catenanes. The duplex knot in (*a*) has 3 (−) nodes and the one in (*b*) has 3 (+) nodes. The duplex DNA catenane in (*c*) has 2 (−) nodes and that in (d) has 2 (+) nodes. The arrows define the nodes and the orientation.

more nodes with the same value of Kn (Rolfsen 1976). Therefore, although Kn provides useful topological information about knots, it does not usually describe a knot uniquely. We have shown elsewhere how the treatment of curves by Schubert (1956) uniquely classifies naturally occurring DNA knots and catenanes (White and Cozzarelli 1984).

Linkage indices do specify completely the members of the torus family of knots and catenanes that includes the predominant products of site-specific recombination. All the nodes are equivalent in these curves that can be drawn as star-shaped forms on the surface of a torus, a doughnut-shaped object, without crossings. The simplest examples are trefoils and singly linked catenanes (Fig.5).

LINKAGE OF CATENANES

For nodes within a single circle, the sign is easy to determine because the path of the DNA provides a way to orient the crossing segments. Nodes between two different rings, or intermolecular nodes, can be given a sign only if a direction is assigned to each ring. For two heterologous circles, such an assignment is arbitrary. However, for two homologous rings or two rings derived from a single ring by recombination, the sequence of the DNA provides a natural way to orient the two rings.

We define the intrinsic linkage of a catenane, Ca, as the algebraic sum of the intermolecular nodes. The linkage values of duplex and single-strand catenane nodes are assigned values of 1 and ½, respectively. It is easy to show that for single-strand catenanes this is consistent with the definition of Lk. In Figure 6, a double-strand circle with $Lk = Tw = +4$ is denatured, then rearranged as a single-strand torus catenane with $Ca = +4$. The winding of the DNA strands in Figure 6, a and c, is identical. For double-strand intermolecular nodes, we again count only the single-strand nodes formed by strands of opposite chemical polarity and not the two same-strand crossings, as already noted for determining the value of Lk. Thus, Lg is 1 for such nodes, whether one considers them in terms of crossing duplexes or of the constituent crossing single strands. In Figure 5 are shown two forms of a singly interlinked duplex catenane, with $Ca = -2$ for Figure 5c and $Ca = +2$ for Figure 5d. Ca uniquely specifies catenanes of the torus family but usually not other catenanes.

We have shown thus far that node counting provides a simple way of treating the linkage in knots and catenanes that provides indices which are strictly compatible with the twist, writhe, and linking number used to de-

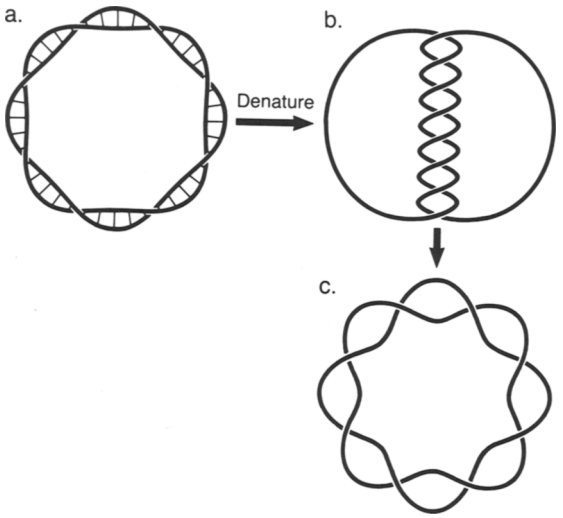

Figure 6. Denaturation of closed duplex circular DNA yields a single-strand catenane. (*a*) A closed duplex circular DNA with Lk = Tw = +4; the lines perpendicular to the helix axis represent nucleotide pairs. Denaturation forms a catenane of circular single-stranded DNA drawn with clustered interlocks in (*b*) and with interlocks evenly spaced in (*c*); since the strands are still oriented as in the helix, there are 8 (+) single-strand nodes and thus Ca = +4. Since the backbone is not broken during denaturation, the Lk of the helix must remain constant and is equal to Ca of the catenane. Topologically, closed circular duplex DNA is a catenane of single-strand rings that happen to be base-paired.

scribe circular DNA. This allows a unitary quantitative treatment of the various topological changes introduced by topoisomerases and recombination. We define linkage by the following equation:

$$Lg = Lk + Ca + Kn. \quad (2)$$

1. If we define writhe as a configuration of DNA that prevents the helix axis from being planar, then knots and catenanes, like supercoils, are forms of writhe. Therefore, the linkage of knots and catenanes is Kn or Ca plus the value of Lk that does not include knot and catenane contributions to writhe. To reduce nomenclature, we still call this restricted linking number Lk.
2. We show more generally that Lk of a closed duplex ring is equal to Ca of the equivalent single-strand catenane using the rigorous spanning surface definition of Lk. If a surface is spread across one of two interlocked curves, then the other curve will puncture the surface from above or below (Fig. 7). If the curves are oriented, Lk is just the sum of such punctures with appropriate sign attached. This is equivalent to our definition of Ca for single-strand catenanes, which this figure could also represent, and therefore Lk = Ca in this case. If we consider each curve in Figure 7 to represent the helix axis of duplex DNA, then by the same argument we prove that Lk = Ca for duplex DNA, only now each puncture has a magnitude of 2 rather than 1.

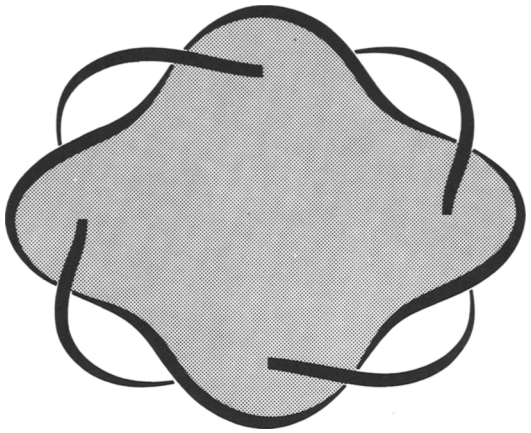

Figure 7. Use of the spanning surface definition of Lk to justify the node-counting measure of Ca. If a surface is spread across one of two interlocked curves, Lk is equal to the algebraic sum of the times the other curve punctures this surface. To attach sign in the example shown, the curves are oriented in parallel and all of the punctures have a (+) sign; therefore, Lk = +4. Since Lk is a property of any two interlocked curves, the figure can equally represent duplex circular DNA or single-strand catenanes. Since for the latter Ca = +4, the linkage assignments for Lk and Ca are compatible.

3. To show that Lk, Ca, and Kn are all measures of topological linkage of the strands of duplex DNA, the DNA is represented as a paper ribbon with the edges corresponding to the W and C strands. If the ribbon is sealed to form a simple uncoiled closed loop, then cutting down the middle of the ribbon, i.e., denaturing the DNA, separates the strands. However, if either |Lk|, |Kn|, or |Ca| > 0, then the same operation forms topologically linked single-strand rings. None of these topological indices, however, is a perfect measure of physical linkage. Knots and catenanes can have Lk, Ca, and Kn = 0 even though strand breakage is required to convert them to free rings, i.e., W and C are topologically bonded. Examples include the four-noded knot and figure-eight catenane (see Fig. 14) in which the linking nodes sum to zero. Moreover, by convention, we ignore the clear physical linkage imparted by C–C and W–W crossings in duplex catenanes but not in single-strand catenanes.

LINKAGE CHANGES BY TOPOISOMERASES

Topoisomerases are enzymes that change a topological property of DNA such as linking number or knot and catenane structure (Wang and Liu 1979; Cozzarelli 1980; Gellert 1981). The best-studied examples are from prokaryotes. By breaking one of the crossing DNA segments at the node and passing the other segment through the break, such enzymes invert the sign of a node (Brown and Cozzarelli 1979). Since they then reseal the break, they invert only one node during each cycle of activity. Two classes of topoisomerases have been distinguished (Liu et al. 1980). Type-1 topoiso-

merases invert single-strand nodes and thereby change Lk in steps of 1; type-2 enzymes act at double-strand nodes and change Lk in steps of 2. If we call this change Me, for mechanism, we can represent the reactions of topoisomerases by the equation

$$Lk^s + Me = Lk^p, \quad (3)$$

where the superscripts s and p refer to substrate and product, respectively. A more general form of this equation, using Lg, applies to knots and catenanes as well,

$$Lg^s + Me = Lg^p. \quad (4)$$

For a circular substrate that is not knotted or catenated, $Lg^s = Lk^s$, and combining Equations 2 and 4, we obtain,

$$Lk^s + Me = Lk^p + Kn^p + Ca^p. \quad (5)$$

We describe the four reactions of type-2 topoisomerases in terms of linkage changes to illustrate the utility of the treatment. The DNA linkage changes made were determined using a ribbon model for DNA (Fig. 8). Gyrase is a type-2 topoisomerase that can introduce supercoils into DNA by inverting a (+) node (Fig. 8a). Lk diminishes by 2, and from Equation 3, $Lk^s - 2 = Lk^p$. Relaxation by gyrase is not an enzymatic reversal of this process, but rather the same DNA segment passage event with the orientation of one segment reversed (Fig. 8b). Thus, relaxation inverts a (−) node and changes Lk by +2, or $Lk^s + 2 = Lk^p$. In the catenation and knotting reactions a single node is inverted, which may be (+) or (−). In the catenation reaction in Figure 8c, a (+) node is inverted and thus $Lk^s - 2 = Lk^p + Ca^p$. Since Lk is unchanged, Ca is predicted to be −2 and is so (Fig. 8c). In the knotting reaction in Figure 8d, a (−) node is inverted and thus $Lk^s + 2 = Lk^p + Kn^p$. Since $Lk^s = 0$, we get $0 + 2 = Lk^p + Kn^p$. Since $Kn^p = +3$ for the knot formed, $Lk^p = -1$, i.e., a (−) supercoil must be introduced as well, as observed. Thus, as predicted by equation 5, supercoiling, catenation, and knotting by type-2 enzymes change Lg equivalently, in steps of 2, even though the linkage is 1 for a supercoil, 2 for the simplest catenane, and 3 for the simplest knot.

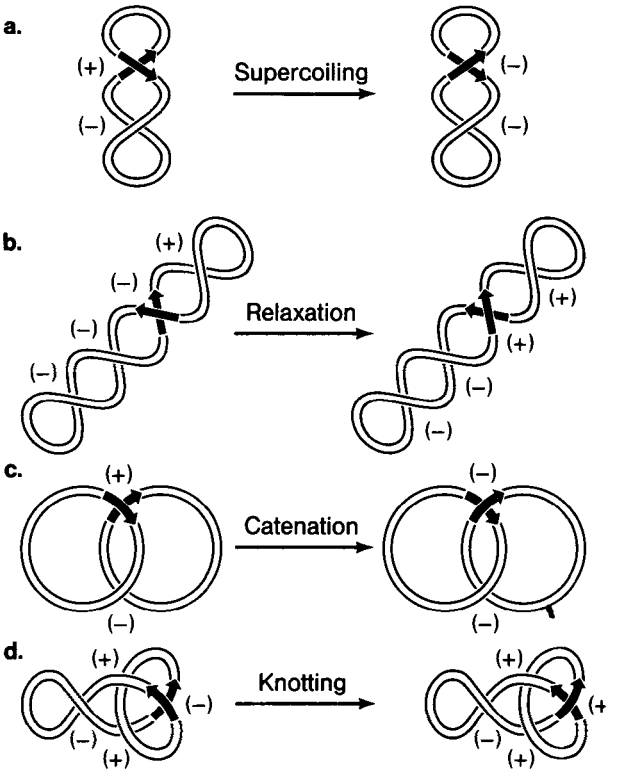

Figure 8. Reactions of type-2 topoisomerases. The duplex substrates shown are: relaxed circular DNA for supercoiling (*a*) and knotting (*d*), a ring containing two (−) supercoils for relaxation (*b*), and two relaxed rings for catenation (*c*). Each substrate contains additional nodes that cancel in the substrate but are trapped in the product. The filled arrows delineate the crossing segments of the sign inversion node. The products were determined using a ribbon model for DNA in which Tw = 0 and in each case sign inversion changed |Lg| by 2.

The Domain Concept and the Conservation of Interdomainal Nodes

In the examples thus far with simple substrates and products, Equation 5 sufficed to predict product structure. In general though, the distribution of linkage in Equation 5 into Lk^p and either Kn^p or Ca^p is not obvious. We next derive equations that specify values for Kn^p or Ca^p and so lead to a complete solution of Equation 5.

To do so, we utilize the concept of a domain introduced by Mizuuchi et al. (1980). The node where the enzyme acts divides the substrate into two domains, which are the continuous DNA lengths bounded by the crossing segment of the sign inversion node. An intradomainal node is made up of segments from the same domain; the crossing segments of an interdomainal node are contributed by different domains. Since Tw nodes are intradomainal, then for a substrate that is neither knotted nor catenated only writhe nodes can be interdomainal. This component of writhe is symbolized $^{ter}Wr^s$, where ter stands for interdomainal. Since solenoidal supercoils can form at most one interdomainal node at each domain boundary, interdomainal writhe is generally from plectonemic supercoiling. The linkage change by Me consists only of interdomainal nodes.

The value of Ca^p is the algebraic sum of all the interdomainal nodes, and we describe the conservation of interdomainal nodes by

$$^{ter}Wr^s + Me = Ca^p. \quad (6)$$

For knots, we must have an additional term, $^{ter}Wr^p$, the interdomainal writhe of the product, because in cir-

cumstances detailed below, interdomainal nodes do not become part of a knot. Thus, the conservation equation for knotting is

$$^{ter}Wr^s + Me = Kn^p + {}^{ter}Wr^p. \quad (7)$$

We illustrate these relationships by considering the production of complex knots. In Figure 9 is shown a generalized substrate for a knotting reaction by a type-2 topoisomerase. One domain is gray and the other clear, and a variable $^{ter}Wr^s$ is depicted by bracketing n interdomainal nodes. In this example, intradomainal supercoils and twist are irrelevant to knotting and not shown. If $n = 0$, then the product will not be knotted because a knot must have at least three nodes. In that case, $Kn^p = 0$ and by Equation 7, $^{ter}Wr^p = Me$. The second instance when $^{ter}Wr^p \neq 0$ is when a single node cannot be trapped in the knot and $^{ter}Wr^p = \pm 1$. This occurs because only an odd number of nodes can be trapped if Me and $^{ter}Wr^s$ are of the same sign, and only an even number if they are of opposite sign. After the enzyme releases the DNA $^{ter}Wr^p$ becomes part of Lk^p.

We determined the structure of the knots formed from the substrate shown in Figure 9 using a ribbon model for DNA. The knots formed were all of a family called twist knots (Rolfsen 1976); we call them stevedore knots to avoid confusion with Tw. Knowing the family, Kn then precisely describes the knots formed. The products for a range of $^{ter}Wr^s$ and for $Me = +2$ and $Me = -2$ are shown in Table 1. In each case, the equations and the properties of $^{ter}Wr^p$ given above correctly predicted the stevedore knot and the supercoils formed. Thus if Me is known and $^{ter}Wr^s$ can be determined, we can solve Equations 5 and 7 and deduce the structure and superhelicity of the knot formed by a topoisomerase.

1. We have been unable to develop a set of topological indices that apply generally to interrupted DNA as well as to intact molecules. The difficulty is well illustrated by topoisomerase reactions. Type-1 topoisomerases, for which $|Me| = 1$, efficiently knot and catenate nicked or gapped DNA by passing a DNA segment through the intact strand opposite a break (Dean et al. 1983). Either $|Ca|$ is 2 for the resulting interrupted singly interlocked catenane and Equation 5 does not apply, or $|Ca|$ is 1 and therefore Ca is not a state function because its value depends on whether the catenane is formed by a type-1 or type-2 enzyme. Because neither alternative is acceptable, Lg is well-defined only for unbroken circular DNA.

However, there are limited ways in which interrupted DNA is defined topologically. It can be useful to ascribe values of Kn and Ca to interrupted DNA because, unlike Lk, these indices are defined in terms of helix axis nodes and, therefore, are unchanged by single-strand interruptions. Indeed, they are most easily measured experimentally by nicking the DNA to avoid the complication of supercoils. Interrupted DNA can even be supercoiled if the structure is stabilized by an interaction with proteins, as in nucleosomes, or if the W and C strands are not free to rotate about each other because noncovalent bonds hold the broken ends together. There is no necessity for covalent bond continuity in the mathematical description of DNA, but is does simplify the process.

LINKAGE CHANGES DURING RECOMBINATION

Although we will focus on site-specific intramolecular recombination because much more is known about its topology, the principles apply equally to other forms of genetic exchange. Recombination within a plasmid containing directly repeated sites, a resolution reaction, results in the division of the parental ring into two daughter rings. In addition to redistributing supercoils, a quantitative change in linkage may occur as strands cross over and rotate about each other. By a procedure similar to that used for topoisomerases, we can derive equations that establish the relationship of the product structures to the arrangement of sites at synapsis and the change in linkage from the mechanism of strand exchange.

Linkage Terms for Recombination

Any mechanism for strand exchange can be assigned an intrinsic linkage change, Me, by considering it as a node-forming process. In the example illustrated in Figure 10a, we have aligned the two recombination sites in parallel. After a double-strand break is made in both sites and one recombination join is made, the

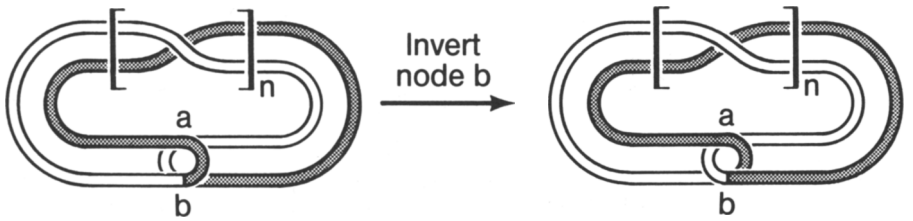

Figure 9. A mechanism for formation of complex knots by a type-2 topoisomerase. The duplex DNA substrate is shown with a variable number (n) of (−) plectonemic interdomainal supercoils set off by brackets from the rest of the DNA. Sign inversion by a type-2 topoisomerase could be at either node a or node b; the latter is illustrated. The sign inversion node divides the circle into two domains that are distinguished by the shading of one domain. The knots formed will increase in complexity in proportion to the number of interdomainal nodes.

Table 1. Conservation of Linkage during Knotting by Type-2 Topoisomerases

$^{ter}Wr^s$	Me	Node inverted[a]	Sign of node inverted	Kn	Rolfsen notation[b]	$^{ter}Wr^p$
+3	+2	a	−	+5	5_2	0
+2	+2	b	−	+3	3_1	+1
+1	+2	a	−	+3	3_1	0
0	+2	b	−	0	0_1[c]	+2
−1	+2	a	−	0	0_1[c]	+1
−2	+2	b	−	0	4_1	0
−3	+2	a	−	0	4_1	−1
−4	+2	b	−	−2	6_1	0
−5	+2	a	−	−2	6_1	−1
+3	−2	b	+	0	4_1	+1
+2	−2	a	+	0	4_1	0
+1	−2	b	+	0	0_1[c]	−1
0	−2	a	+	0	0_1[c]	−2
−1	−2	b	+	−3	3_1	0
−2	−2	a	+	−3	3_1	−1
−3	−2	b	+	−5	5_2	0
−4	−2	a	+	−5	5_2	−1
−5	−2	b	+	−7	7_2	0

A ribbon model for DNA containing the indicated $^{ter}Wr^s$ was knotted by a double-strand passage event according to the scheme in Fig. 9. The properties of the knot formed and $^{ter}Wr^p$ are tabulated. In some instances, node a or b is incorporated into a knot even though it is apparently intradomainal. To avoid the complication that this introduces, it suffices to exclude both nodes a and b from the value of $^{ter}Wr^s$.

[a]The letter designation for the node inverted is shown in Fig. 9.

[b]A standard knot nomenclature contained in Rolfsen's (1976) book on knots and links. The first number is the number of nodes and the subscript indexes the examples with that number of nodes. All of these knots are of the twist, or stevedore, class.

[c]An unknotted circle; not listed in Rolfsen table.

DNA segment axes are still flat in a plane and no change in linkage has occurred. The second join must either cross over or under the first one, thus creating a new double-strand node. If the second join is made over the first, as shown, it will form a (+) node and Me = +1. A second join under the first would result in Me = −1.

In Figure 10b, we illustrate topology of the mechanism proposed for phage λ integrative recombination (Nash et al. 1981). The sites are aligned in parallel but with the strands of opposite chemical polarity (called W and C) across from each other. Like strands of each site, e.g., W strands, are broken and one portion is rotated about the C strand in the direction that dissi-

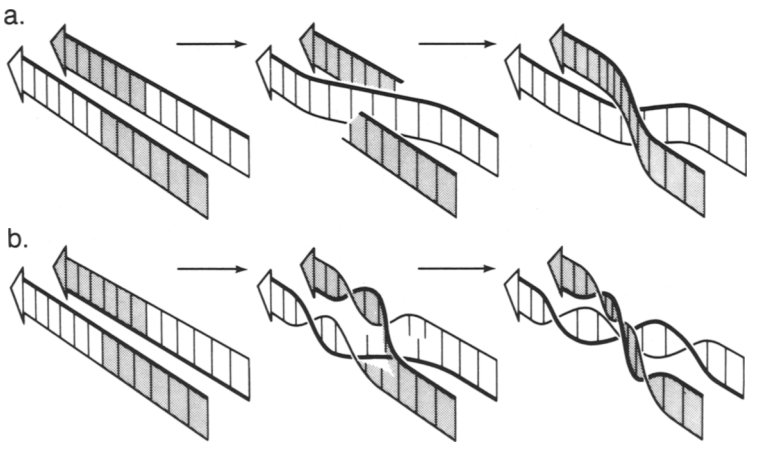

Figure 10. Two topological models for resolution. A ribbon representation of the sites of recombination in a circular DNA. The W and C strands are distinguished as bold and thin lines and the two recombination domains as clear and shaded. The arrows show the direction of the directly repeated sites. The vertical lines for base pairs and the untwisting at the junctions are schematic representations. (*a*) A simple resolution model in which Me = +1. After double-strand breakage and the first crossed reunion, the DNA is still planar and no new node is made. The second join creates a (+) interdomainal node, thus $^{ter}Me = +1$ and $^{tra}Me = 0$. (*b*) The topology of the Nash model for resolution. Because the W and C strands of the two sites are now directly across from each other, breakage and reunion require a rotation about the helix axis. The result of the first single-strand exchange is a Holliday structure. The Me for the total reaction is +4, with $Me^{tra} = +3$ and $Me^{ter} = +1$.

pates (−) supercoils until it meets and joins the remainder of the strand of the other site. Four (+) single-strand nodes are formed between the W and C strands in this Holliday intermediate, and therefore the linkage change is +2. The central W-W node is ignored, according to convention. Repeating this process for the C strands produces another change in linkage of +2 and thus Me = +4 for this mechanism.

Since only interdomainal nodes, whether trapped in the substrate or introduced by the enzyme, contribute generally to Ca^p and Kn^p, it is necessary to determine which nodes formed during recombination are interdomainal. The two domains are bounded by the two recombination sites. Although Me for topoisomerases is entirely interdomainal, we must break down Me for recombination into an interdomainal component, ^{ter}Me, and an intradomainal component, ^{tra}Me, where $Me = {}^{ter}Me + {}^{tra}Me$. In both examples illustrated in Figure 10, the axes of the recombination sites are not crossed at synapsis and cross just once after recombination, so that for both mechanisms, $^{ter}Me = +1$. The additional nodes introduced in the example in Figure 10b are intradomainal and only twist the two individual helices by a total of 3, and therefore $^{tra}Me = +3$. Although the two mechanisms are quite different, the fact that ^{ter}Me is the same for both pathways causes the values of Ca^p and Kn^p to be the same.

As with topoisomerases, the second feature of a recombination event that contributes to Ca^p or Kn^p is the value of $^{ter}Wr^s$, the interdomainal nodes in the substrate. Some plectonemic supercoils and random flops may be trapped as interdomainal nodes at the time of recombination. In addition, wrapping of the DNA around the enzyme, such as in the nucleosomelike structure suggested for Int protein bound to its sites, could stabilize interdomainal nodes (Better et al. 1982). Alignment of the sites in a specific orientation on the enzyme sometimes requires an additional interdomainal node. The interdomainal writhe, $^{ter}Wr^s$, is made up of those nodes that are crossings of the two recombination domains, whether trapped supercoils and flops, enzyme wraps, or a consequence of site alignment. An enzyme can contribute to product structure by an effect on $^{ter}Wr^s$, as well as on ^{ter}Me (see Spengler et al., this volume).

Equations for Recombination

We begin with Equation 5, the conservation of linkage equation for a substrate that is neither catenated nor knotted ($Lk^s + Me = Lk^p + Ca^p + Kn^p$). Recombination of directly repeated sites produces two daughter rings which are not knotted, but may be catenated, and the equation for resolution is

$$Lk^s + Me = Lk^p + Ca^p. \tag{8}$$

This can be broken down into intradomainal and interdomainal components, producing

$$^{tra}Lk^s + {}^{tra}Me = Lk^p \tag{9}$$

and

$$^{ter}Wr^s + {}^{ter}Me^{res} = Ca^p \tag{10}$$

where res stands for resolution and tra stands for intradomainal. Equation 9 describes the redistribution of parental intradomainal supercoils and supercoils introduced by the enzyme into daughter ring supercoils. Equation 10 describes the conversion into catenane links of parental interdomainal supercoils and links introduced by the enzyme itself. Equations 9 and 10 hold for all mechanisms in which nodes do not cross domainal boundaries during the reaction.

Recombination of two inverted sites, or inversion, produces a single full-length circle in which one domain has been inverted. The inversion of one domain reverses the orientation of the crossing segments of interdomainal nodes, but not their overlay. Therefore, inversion reverses the sign of all interdomainal nodes. To preserve the sign convention for both substrate and product, we reverse the sign of $^{ter}Wr^s$ to reflect the sign these nodes will have after recombination. When inversion occurs by the same mechanism as resolution, the sign change also causes $^{ter}Me^{inv}$ to equal $-{}^{ter}Me^{res}$, where inv stands for inversion. Finally, we again use $^{ter}Wr^p$ to represent those interdomainal nodes that cannot be trapped in a knot. Then we can write, respectively, the equations for intradomainal and interdomainal node conservation during inversion:

$$^{tra}Lk^s + {}^{tra}Me = {}^{tra}Lk^p \tag{11}$$

and

$$-{}^{ter}Wr^s + {}^{ter}Me^{inv} = Kn^p + {}^{ter}Wr^p \tag{12}$$

The term $^{ter}Wr^p$ has nonzero values in the same instances as described above for topoisomerases and combines with $^{tra}Lk^p$ as the linking of the product once the enzyme releases.

In Figure 11, we illustrate a simple recombination event between directly repeated sites that is an example of the general mechanism in Figure 10a. To align the sites in parallel, the DNA is bent such that there is a (+) node near the sites and a compensating (−) node further away. At the time of synapsis and site cleavage, the (+) node is interdomainal, so $^{ter}Wr^s = +1$. The (−) node is intradomainal, and since there is no other Tw or Wr in this model substrate, $^{tra}Lk^s = -1$. Since the second recombination join is made over the first, creating a (+) node, $^{ter}Me = +1$; $^{tra}Me = 0$. From Equation 9 we find that $Lk^p = -1 + 0 = -1$ and there is one (−) supercoil among the two product rings. From Equation 10 we find that $Ca^p = +1 + 1 = +2$ and, indeed, the product rings are joined by one (+) double-strand link. Equation 5 is satisfied because the total change in linkage, Lg, is $Lk^p + Ca^p - Lk^s = +1$, which equals the value of Me. Thus, even a simple

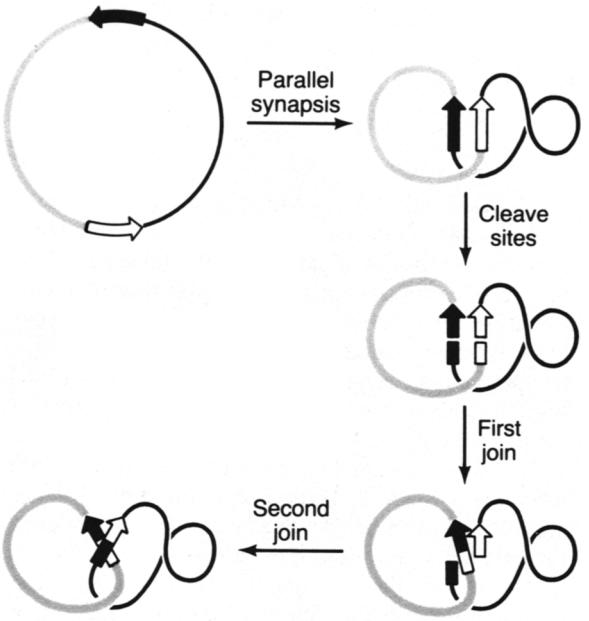

Figure 11. A simple Me = +1 mechanism for resolution. This is an example of the mechanism drawn in general topological form in Fig. 10a. The two recombination sites are shown as filled and open arrows and the domains between the sites are gray and black lines representing duplex DNA. In the first step, the two sites are aligned in parallel by a 90° rotation of each site in opposite directions. A (+) interdomainal node results that is compensated by a (−) intradomainal supercoil. Both sites are cleaved homologously; the second recombination join is above the first, yielding a Ca = +2 catenane. Because the (−) supercoil put in at synapsis remains, Me = ^{ter}Me = +1.

mechanism with Me = +1 acting on a relaxed circle can produce catenated rings containing a supercoil.

An example of recombination with a supercoiled substrate is illustrated in Figure 12. Negative plectonemic supercoils in the substrate are trapped as interdomainal nodes at the time of recombination, so that in this case $^{ter}Wr^s = -6$. Me = 0 for the reaction shown and Equation 10 predicts the correct product: two catenated rings with Ca = −6.

Recombination Basics

We were surprised to find that simple qualitative relationships connect the recombination parameters we have developed. These are illustrated using the generalized substrates for resolution and inversion in Figure 13. The two sites are shown as one open and one filled arrow, aligned in parallel, and broken in anticipation of exchange. Because one round of recombination always produces two hybrid sites, inverted repeats necessarily lead to an inversion product, whereas direct repeats will always lead to a resolution product. Figure 13 shows the qualitative relationship between the orientation of the two sites in the primary sequence and the values of $^{ter}Wr^s$ and ^{ter}Me.

In Figure 13b, the directly repeated sites are synapsed in parallel with one interdomainal node. In Figure 13a, inverse sites are synapsed in parallel with no intervening nodes. The DNA crossings in this figure are equal to $^{ter}Wr^s$. If sites are synapsed in parallel, then for directly repeated sites, $^{ter}Wr^s$ must be odd, and for inverted repeats, $^{ter}Wr^s$ must be even. The crossing of the dotted lines in Figure 13 shows the number of interdomainal nodes put in by the enzyme, or ^{ter}Me; for sites synapsed in parallel, ^{ter}Me must be odd for both direct and inverted sites.

The basic relationships for both parallel and antiparallel alignment are summarized in Table 2. There are three terms to consider: the alignment of synapsed sites and whether $^{ter}Wr^s$ and ^{ter}Me are odd or even. There are just two ways of assorting these three terms to get resolution and two different combinations that lead to inversion. Therefore, once one of these three parameters is determined experimentally for a reaction, the other two are fixed. Furthermore, since synapsed site alignment will in all likelihood be dictated by the structure of the enzyme·DNA complex, it will be the same for resolution and inversion. Thus, ^{ter}Me will be either odd for both reactions or even for both, and $^{ter}Wr^s$ will be even for one and odd for the other.

Test of the Equations

We tested the recombination equations using ribbon models to determine the linkage of products of resolution and inversion. Lk^s was varied between +5 and −7 and the results are tabulated for the simple mechanism illustrated in Figure 10a (Table 3) and the Nash mechanism depicted in Figure 10b (Table 4). In all cases, product Lk, Ca, and Kn obtained with models were correctly predicted by the equations. Because the knots and catenanes were exclusively of the torus family, the computed values of Ca and Kn specify their structure uniquely.

As predicted by the conservation of interdomainal nodes, the twists put in during recombination, ^{tra}Me, affect Lk^p but not Ca^p or Kn^p. This can be seen by comparing Tables 3 and 4, since the two mechanisms

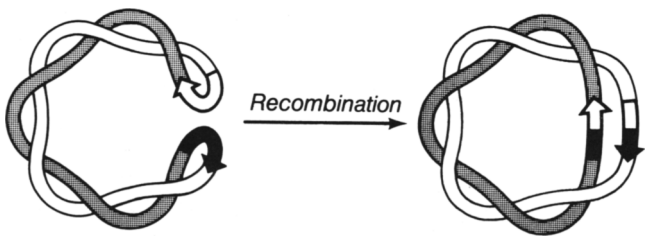

Figure 12. Resolution of a supercoiled substrate by an Me = 0 mechanism. The substrate contains 5 plectonemic (−) supercoils and two directly repeated recombination sites shown as open and filled arrows. The two domains that are separated by recombination are shaded and clear, and all the supercoils are interdomainal. Direct crossover between the antiparallel sites creates no new nodes (Me = 0) but generates a catenane with Ca = −6 which preserves the right-handed helical intertwining and the (−) sign of the substrate supercoils.

Table 2. Basic Relationships of Recombination Parameters

Sequence organization of sites	Alignment of synapsed sites	$^{ter}Wr^{s}$	^{ter}Me	Result
Direct	parallel	odd	odd	resolution
Direct	antiparallel	even	even	resolution
Inverse	parallel	even	odd	inversion
Inverse	antiparallel	odd	even	inversion

have the same ^{ter}Me but different values for ^{tra}Me. Nor does the mechanism of site alignment affect Ca^{p} and Kn^{p}. There are two basic operations in site alignment: reversal of site orientation by introduction of an odd number of supercoils between the sites and a 180° rotation about the helix axis to position strands of appropriate polarity. For example, if a ribbon model for DNA has two directly repeated sites, then if Lk^{s} is even, addition of an odd number of supercoils between the sites will make them parallel; but now strands of the same polarity at each site are opposite each other. If one site, designated site 1, is rotated 180° about the helix axis, then both strands of same polarity are across from each other. Designating the product ring that has the tip of the site 1 arrow as ring 1, and the other as ring 2, we tabulate the products when the orientation of the sites and the direction of the twist are varied (Table 5). Again, the value of Ca^{p} is unaffected; only the supercoiling of the product and the distribution of the supercoils between rings 1 and 2 are affected. Therefore, only $^{ter}Wr^{s}$ and ^{ter}Me influence the catenane and knot structure of the product, regardless of the precise nature of strand synapsis and exchange. The exact correspondence between the predictions and the results using DNA models validate the linkage assignments and the equations.

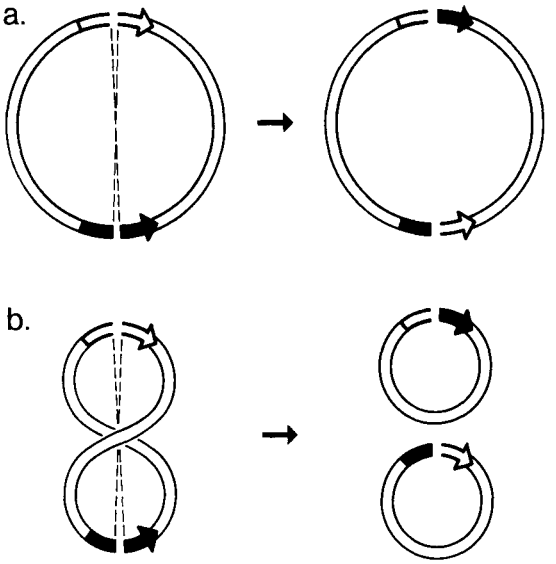

Figure 13. Generalized substrates for inversion (*a*) and resolution (*b*). The two recombination sites in circular duplex DNA substrates are shown as filled and open arrows that are broken at the crossover sites and aligned in parallel in anticipation of recombination. The dotted lines indicate the path of crossover. The number and sign of the crossings of the dotted lines and of the DNA equal ^{ter}Me and $^{ter}Wr^{s}$, respectively. The figure shows why inversion of sites aligned in parallel requires an odd ^{ter}Me but an even $^{ter}Wr^{s}$, whereas for the comparable resolution both of these parameters must be odd.

1. The value of $^{ter}Wr^{p}$ is always zero for a single simple round of resolution. This follows because $^{ter}Wr^{s}$ and ^{ter}Me always sum to an even number for a round of resolution with an unknotted substrate (Table 2), and the resultant products are torus catenanes that can only have an even value of Ca. Therefore $^{ter}Wr^{p}$

Table 3. Recombination Products Found According to a Simple Mechanism

	Resolution				Inversion					
Lk^{s}	$^{ter}Wr^{s}$	$^{tra}Lk^{s}$	Lk^{p}	Ca^{p}	$^{ter}Wr^{s}$	$^{tra}Lk^{s}$	Kn^{p}	Lk^{p}	$^{tra}Lk^{p}$	$^{ter}Wr^{p}$
5	5	0	0	6	6	−1	−7	−1	−1	0
4	5	−1	−1	6	4	0	−5	0	0	0
3	3	0	0	4	4	−1	−5	−1	−1	0
2	3	−1	−1	4	2	0	−3	0	0	0
1	1	0	0	2	2	−1	−3	−1	−1	0
0	1	−1	−1	2	0	0	0	−1	0	−1
−1	−1	0	0	0	0	−1	0	−2	−1	−1
−2	−1	−1	−1	0	−2	0	0	1	0	1
−3	−3	0	0	−2	−2	−1	0	0	−1	1
−4	−3	−1	−1	−2	−4	0	3	0	0	0
−5	−5	0	0	−4	−4	−1	3	−1	−1	0
−6	−5	−1	−1	−4	−6	0	5	0	0	0
−7	−7	0	0	−6	−6	−1	5	−1	−1	0

A ribbon model for an intact circular DNA substrate with no initial twist and containing the indicated Lk^{s} and $^{ter}Wr^{s}$ was recombined as illustrated in Fig. 11. All supercoils were initially interdomainal but a (+) $^{ter}Wr^{s}$ node was introduced as indicated to align the sites in parallel, resulting in a $^{tra}Lk^{s}$ of −1. $^{tra}Me = 0$, $^{ter}Me^{res} = +1$, and $^{ter}Me^{inv} = -1$. Torus catenanes resulted from resolution of substrates with directly repeated sites, and torus knots arose from inversion of inverted repeats. The measured values of Ca, Kn, and Lk^{p} are tabulated as are $^{tra}Lk^{p}$ and $^{ter}Wr^{p}$. $^{ter}Wr^{p}$ was determined by inspection of the products of recombination immediately after strand exchange and before the products were arranged in standard configuration. $^{tra}Lk^{p}$ was determined as the difference between Lk^{p} and $^{ter}Wr^{p}$.

Table 4. Recombination Products Found According to Nash Scheme

	Resolution				Inversion					
Lk^s	$^{ter}Wr^s$	$^{tra}Lk^s$	Lk^p	Ca^p	$^{ter}Wr^s$	$^{tra}Lk^s$	Kn^p	Lk^p	$^{tra}Lk^p$	$^{ter}Wr^p$
5	5	0	3	6	6	−1	−7	2	2	0
4	5	−1	2	6	4	0	−5	3	3	0
3	3	0	3	4	4	−1	−5	2	2	0
2	3	−1	2	4	2	0	−3	3	3	0
1	1	0	3	2	2	−1	−3	2	2	0
0	1	−1	2	2	0	0	0	2	3	−1
−1	−1	0	3	0	0	−1	0	1	2	−1
−2	−1	−1	2	0	−2	0	0	4	3	1
−3	−3	0	3	−2	−2	−1	0	3	2	1
−4	−3	−1	2	−2	−4	0	3	3	3	0
−5	−5	0	3	−4	−4	−1	3	2	2	0
−6	−5	−1	2	−4	−6	0	5	3	3	0
−7	−7	0	3	−6	−6	−1	5	2	2	0

Recombination of a ribbon model for DNA was carried out by the Nash scheme illustrated in Fig. 10b for resolution. $^{tra}Me = +3$, $^{ter}Me^{res} = +1$, and $^{ter}Me^{inv} = -1$. The alignment and linkage determinations are as described in Table 3.

Table 5. Effect of Variation in Manner of Synapsis on Linkage of Recombination Products

Lk^s	Interdominal node added	$^{ter}Wr^s$	$^{tra}Lk^s$	Hand of half-twist of site 1	Lk^{p_1}	Lk^{p_2}	Ca^p
3	0	3	0	none	0	0	+4
2	+1	3	−1	right	−1	0	+4
1	0	1	0	none	0	0	+2
0	+1	1	−1	right	−1	0	+2
−1	0	−1	0	none	0	0	0
−2	+1	−1	−1	right	−1	0	0
−3	0	−3	0	none	0	0	−2
3	0	3	0	none	0	0	+4
2	+1	3	−1	left	0	−1	+4
1	0	1	0	none	0	0	+2
0	+1	1	−1	left	0	−1	+2
−1	0	−1	0	none	0	0	0
−2	+1	−1	−1	left	0	−1	0
−3	0	−3	0	none	0	0	−2
3	0	3	0	none	0	0	+4
2	−1	1	+1	right	0	+1	+2
1	0	1	0	none	0	0	+2
0	−1	−1	+1	right	0	+1	0
−1	0	−1	0	none	0	0	0
−2	−1	−3	+1	right	0	+1	−2
−3	0	−3	0	none	0	0	−2
3	0	3	0	none	0	0	+4
2	−1	1	+1	left	+1	0	+2
1	0	1	0	none	0	0	+2
0	−1	−1	+1	left	+1	0	0
−1	0	−1	0	none	0	0	0
−2	−1	−3	+1	left	+1	0	−2
−3	0	−3	0	none	0	0	−2

The Lk of each of the product rings and their interlocking were determined with ribbon models for a DNA substrate with two recombination sites designated 1 and 2. For Lk^s of odd value, sites were aligned parallel with all supercoils as interdomainal nodes and recombined with an $Me = ^{ter}Me = +1$. For Lk^s of even value, sites were first aligned antiparallel; the indicated node was added to make the sites parallel, and site 1 was twisted to the right or left as indicated to bring strands of like polarity together. A right-hand half-twist is defined as a 180° clockwise rotation of a site pointing away from you, or the usual right-hand rule. Lk^{p_1} and Lk^{p_2} are the writhe of the product circles with the arrowheads of sites 1 and 2, respectively.

can be ignored for resolution, because the only even value for product interdomainal nodes not consistent with a catenane is zero.

2. For convenience in testing the equations with ribbon models for DNA, the substrates were initially assigned neither intradomainal supercoils nor twist and only a small number of interdomainal nodes. It is illuminating to calculate the linkage changes, using the Nash model, for a more realistic substrate. A standard substrate for Int is BP86, which is a 9.4-kb-long plasmid containing two recombination sites in direct repeat, separated by 1.5 kb (Mizuuchi et al. 1980). Assuming one helical twist per 10.5 bp, Tw = +895 and Wr = −54 if the supercoil density is 0.06. If all these supercoils are plectonemic coils and the sites collide randomly, $^{ter}Wr^{s}$ would be about −15. The experimental result is that Ca of the product is distributed about −10 (Spengler et al., this volume). As throughout, we assume that the product will have the same twist, but now Wr = −41. The four supercoils lost because Me = +4 provide an ample source of free energy to drive the reaction, and the conversion of 10 supercoils to catenane links may provide still more.
3. A useful ribbon model for DNA can be constructed from 1.75-in-wide strips of 0.01-in-thick plastic such as Mylar. The ends of 4-ft lengths are bent back and riveted or glued so that two ordinary clothes snaps, in opposite orientation, can be attached to each end. Colored 0.5-in plastic tapes are applied along each edge up to the snap to indicate W and C strands. Two ribbons are snapped together to form a model intramolecular recombination substrate. Arrows in third and fourth colors are placed across the joins to indicate the polarity of the recombination sites.

Application to Resolvase Recombination

The recombination activity of resolvase encoded by the Tn*3* family of transposons is limited to resolution (Reed 1981; Heffron 1983). Resolvase produces a singly linked catenane regardless of the substrate supercoil density. To do so, both ^{ter}Me and $^{ter}Wr^{s}$ must be fixed (Krasnow and Cozzarelli 1983). Using the linkage equations and an interpretation of the minor products of resolvase recombination, we can deduce the value for ^{ter}Me. Although the major product is a singly interlinked catenane, there is a gradual appearance of first a four-noded knot and then the five-noded figure-eight catenane (Krasnow et al. 1983a; A. Stasiak, unpubl.). Further analysis suggests the eventual formation of even smaller amounts of a six-noded knot. One very accomodating model for these unusual results is that the minor products are generated by successive rounds of recombination by resolvase that remains on the same substrate molecule, or iterative recombination (Fig. 14).

If the four-noded knot results from a second recombination event on a singly linked catenane and the figure-eight catenane from recombination on the four-noded knot, as we suggest, then ^{ter}Me must be either +1 or −1, since only one more node is introduced in each round. Because the entrapped supercoil node of the figure-eight catenane is uniquely (+) (Krasnow et al. 1983b), the ^{ter}Me component of Me for resolvase must be +1. From Table 2 we deduce that $^{ter}Wr^{s}$ at the time of recombination is odd and that the sites are aligned in parallel.

The structures found using a DNA ribbon model for three processive rounds of recombination with ^{ter}Me = +1 and variable $^{ter}Wr^{s}$ are listed in Table 6. Only when $^{ter}Wr^{s} = -3$ are all three observed products accounted for. A prediction of this model is that Ca for the singly linked catenanes will be −2, not +2, and this has recently been shown (S.A. Wasserman and N.R. Cozzarelli, in prep.). This result supports both the iterative recombination model and our conclusions that the sites are synapsed in parallel with $^{ter}Wr^{s} = -3$ and $^{ter}Me = +1$.

1. In a topological treatment of recombination, there is no distinction between a mechanism that introduces a (+) node and one that removes an existing (−) node, since the Lg change is +1 in each case. Similarly, we make no statement about the physical organization of synapsed sites on resolvase beyond that the sites are topologically equivalent to a parallel alignment. In addition to parallel and antiparallel alignment, we must consider sites that cross precisely at the point of recombination that are not strictly defined as either intradomainal or interdomainal. Using a DNA model, one can show that recombination of crossed sites requires that ^{ter}Me be odd even if they are nearly antiparallel. Parallel sites also require an odd ^{ter}Me (Table 2) and thus there is no topological difference between the same Me acting on crossed and parallel sites.

   For both crossed and parallel sites, $^{ter}Wr^{s}$ must be odd for direct sites and even for inverted sites (Table 2). However, modeling shows there can be only an even number of interdomainal nodes in crossed site substrates, not counting the crossed site node, because the two domains form two distinct loops. Therefore, a crossed site node is topologically equivalent to an interdomainal node with directly repeated sites, but an intradomainal node with inverted sites. This can be justified by treating the crossed site node as the limiting case and moving the node a short distance away from both domainal boundaries. For both direct and inverted sites, the node can be moved such that it is either intradomainal or interdomainal, but in each case, only one movement is compatible with converting the sites to the topologically equivalent parallel sites. For directly repeated sites, it becomes interdomainal, and for inverted sites, intradomainal. Thus, the effect of interdomainal writhe is also topologically equivalent for parallel and crossed sites.
2. A single recombination event that does not involve topoisomeraselike strand-passage events produces

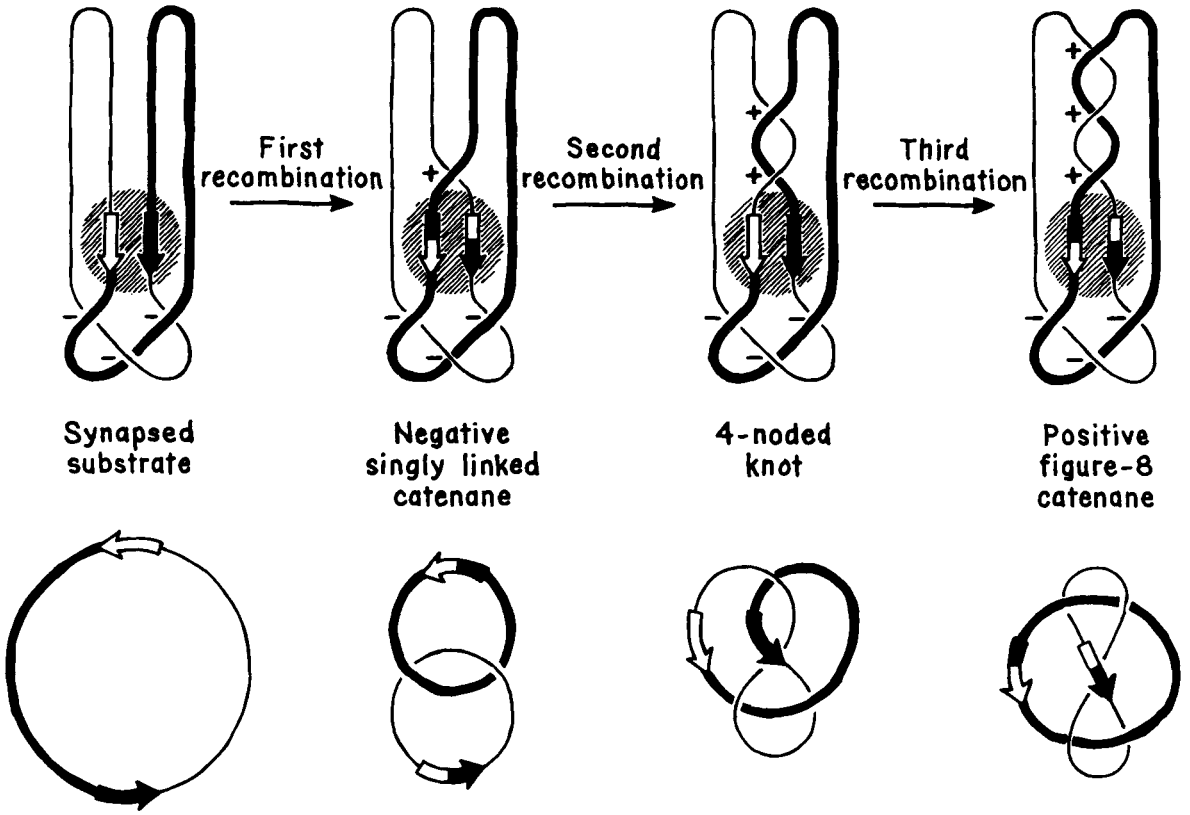

Figure 14. Iteration of the resolvase mechanism. The two directly repeated resolvase sites are shown as filled and open arrows, and the domains are distinguished by bold and standard lines that represent duplex DNA. The row of molecules above are shown bound to resolvase (crosshatched circle) during successive rounds of recombination. The row of structures below are deproteinized and in standard form. Since $^{ter}Wr^{s} = -3$ and $^{ter}Me = +1$, the first round gives a (−) singly linked catenane. Ordinarily, the DNA is released at this point, but we postulate that occasionally it remains bound by the enzyme, processively recombining the DNA further to a four-noded knot and a (+) figure-eight catenane.

only torus knots and catenanes that are specified uniquely by Ca and Kn. The multiple events of iterative recombination can produce knots and catenanes that are not of this class. It is often possible to predict their structure and to calculate $^{ter}Wr^{p}$ if the following rules of thumb (presented for knots) are used along with the linkage equations and Rolfsen's (1976) compendium of knots and links. Rule 1: If interdomainal nodes can form a knot, they will; Rule 2: $|^{ter}Me| \geq |Kn^{p} - {}^{ter}Wr^{s}|$; Rule 3:

Table 6. Predictions of Iterative Recombination

	Product						
	round 1		round 2			round 3	
Substrate $^{ter}Wr^{s}$	Ca^{p}	Rolfsen notation	Kn^{p}	$^{ter}Wr^{p}$	Rolfsen notation	Ca^{p}	Rolfsen notation
5	6	6_1^2	7	0	7_1	8	8_1^2
3	4	4_1^2	5	0	5_1	6	6_1^2
1	2	2_1^2	3	0	3_1	4	4_1^2
−1	0	0_1^2	0	1	0_1[a]	2	2_1^2
−3	−2	2_1^2	0	−1	4_1	0	5_1^2
−5	−4	4_1^2	−2	−1	6_1	−2	7_1^2
−7	−6	6_1^2	−4	−1	8_1	−4	9_1^2

Products of iterative recombination. Recombination using a ribbon model for DNA was carried out as described in Table 3 except that a second and third round of recombination were performed without an intervening release of the DNA. Therefore, the $^{ter}Wr^{p}$ node, if present, was not released into writhe and no new node was introduced to align the sites. The Rolfsen notations for the four-noded knot and the figure-eight catenane are 4_1 and 5_1^2, respectively.

[a]An unknotted circle; not listed in Rolfsen table.

$^{ter}Wr^{p} = 0, +1$, or -1 when a knot is formed. All the structures listed in Table 6 were correctly predicted in this manner.

Application to the Int System

The λ Int system (Nash 1981) contrasts with resolvase in that substrate supercoiling is passively incorporated into product structure, and the complexity of the catenanes and knots produced increases with substrate supercoiling (Pollock and Nash 1980, 1983). The application of the linkage equations to Int is detailed elsewhere in this volume (Spengler et al., this volume), and we shall just summarize some major conclusions. The product knots and catenanes are exclusively of the torus class and have a right-handed helical interwinding because the substrate interdomainal nodes are (−) plectonemic supercoils (Fig. 4). The knot nodes are exclusively (+), as expected from the inversion of sequence that accompanies their formation. Thus, interdomainal negative supercoils are converted to catenane or knot nodes precisely as predicted by the equations.

The Int system works readily with direct and inverse sites. Because the sign of interdomainal nodes is reversed during inversion, the sign of ^{ter}Me for inversion is the opposite of that for resolution. Therefore, the overall Me for the same mechanism operating on direct and inverted sites is the same only if there is no interdomainal component. Since ^{tra}Me is the same for both types of recombination, it follows that $Me^{res} - Me^{inv} = (^{tra}Me^{res} + {}^{ter}Me^{res}) - (^{tra}Me^{res} - {}^{ter}Me^{res})$, which simplifies to

$$Me^{res} - Me^{inv} = 2\ {}^{ter}Me^{res}. \quad (13)$$

Thus there is a fixed relationship between the Me values for resolution and inversion that is determined by the interdomainal component of Me. Using the Nash model as an example, the values of Me for resolution (+4) and for inversion (+2) differ by +2, and $^{ter}Me^{res}$ is +1.

We can describe the results of Nash and Pollock (1983) in terms of our topological formulation. They determined a value for Me^{inv} for Int of +2 by assuming that it is equal to the change in supercoiling during inversion. To simplify the analysis of topoisomers, they looked only at unknotted products and assumed that the change in Lk (ΔLk) for these equals Me^{inv}. From Equations 11 and 12, we derive instead that when $Kn^{p} = 0$, $Me^{inv} = \Delta Lk\ 2^{ter}\ Wr^{s}$. Therefore, $\Delta Lk = Me^{inv}$ only when there is no net interdomainal writhe in the substrate, and there is no necessity for this. For example, if $Me = {}^{ter}Me = 0$ and $^{ter}Wr^{s} = -1$, then $^{ter}Wr^{p} = +1$ and ΔLk equals +2, even though Me = 0. Although a more difficult experiment, the ambiguity would be removed if knotted products were analzyed because then critical terms could be measured and Me^{inv} calculated directly from Equations 11 and 12. However, the value of Me^{inv} of +2 is reasonable and equals that predicted by the Nash model for recombination. In this model, $^{ter}Me^{res} = +1$, the same as the value deduced for resolvase.

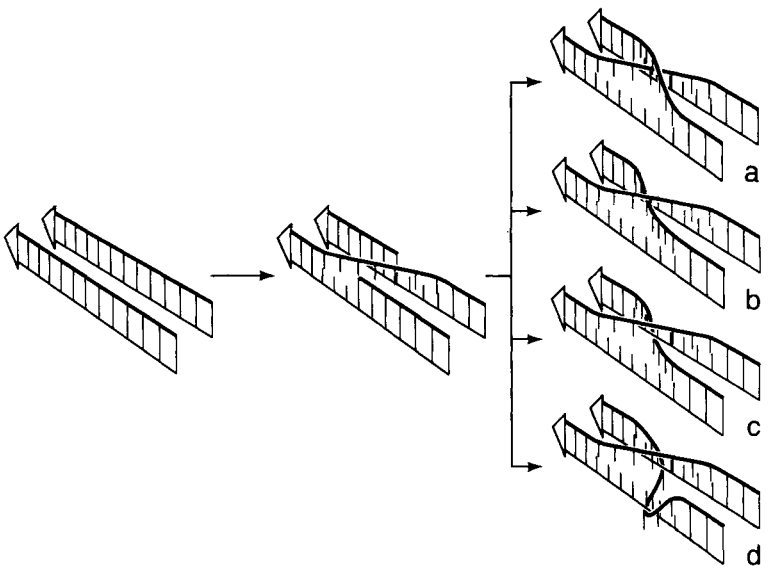

Figure 15. Linkage changes during formation of Holliday intermediates. Only the two directly repeated recombination sites of circular duplex DNA are shown. The bold line is the W strand and the narrow line is the C strand of the sites, whose direction is shown by the arrow. The vertical lines for base pairs and the denaturation at the crossover points are schematic representations to facilitate viewing the topological changes. In the first step, both sites are broken at the same place in the W strands and the first cross-join is made. In the second step, the second cross-join is made either over (*a*) or under (*b*) the joined W strand. It is made under the intact duplex in *c* and under all three strands in *d*. The Me is 0, 0, −1, and −2 for these four alternatives, respectively.

Holliday Structures

The Holliday structure is an important intermediate in general recombination and is a part of the Nash mechanism for Int recombination. A Holliday structure is formed when one reciprocal single-strand exchange occurs before the second set of strands is broken. There is no unique linkage change associated with Holliday structure formations because Me depends on the motion of the DNA strands during the exchange.

Four mechanisms for Holliday structure formation are illustrated as strand passage events in Figure 15 to emphasize the topological consequences. After the W strands of each site are broken and there is one cross-religation, there are four simple ways to form a node by religating the remaining strand. In paths a and b, this occurs above and below the other religated strand; because these are both W strands, the node formed is not counted and Me = 0. However, if the second religation is under one (path c) or both (path d) of the C strands, then W–C nodes are formed and the Me for path c is -1 and for path d, -2. The formation of the Holliday intermediate in the Nash model (Fig. 10b) is the mirror image to path d and Me is $+2$. The linkage change of the overall reaction is the sum of Me for each exchange, and not necessarily twice the Me of the first. Branch migration can occur before maturation of the Holliday structure, but because the DNA is not broken in the process, it does not affect Me. However, it can significantly influence the shape of the products by moving the domain boundaries.

Separation of the two single-strand exchange points by more than one helical turn, as shown in Figure 16, results in the formation of a hemicatenane, in which the two sections of parental helix between the crossover points topologically interlink the daughter rings. For this recombinant, Ca is composed of one double-strand node and six single-strand nodes and equals -4. The linkage changes in the formation and resolution of the Holliday intermediate are not equal in this case because of the change in $^{ter}Wr^{s}$ after the first exchange. Separation of such a hemicatenane could require a topoisomerase.

1. Holliday intermediates formed with inversion substrates have particularly interesting topological properties. Inversion of the sequence between the recombination sites reverses the strand connections of the sequence with the flanking regions so that its C and W strands are connected, respectively, to flanking W and C strands. The result is that the W and C strands are joined together to form a single continuous backbone (Figure 17). The linkage in a limiting case in which the two recombination sites are adjacent is a single-strand stevedore knot that maintains the circular conformation (Fig. 17a). In the more realistic case with sites farther apart (Fig. 17b), a more complicated single-strand knot is formed with two topologically segregated interwound domains. Nonetheless, the intermediates can be matured into standard recombinants. Branch migration in knotted intermediates of inversion is energetically limited because it segregates knot in-

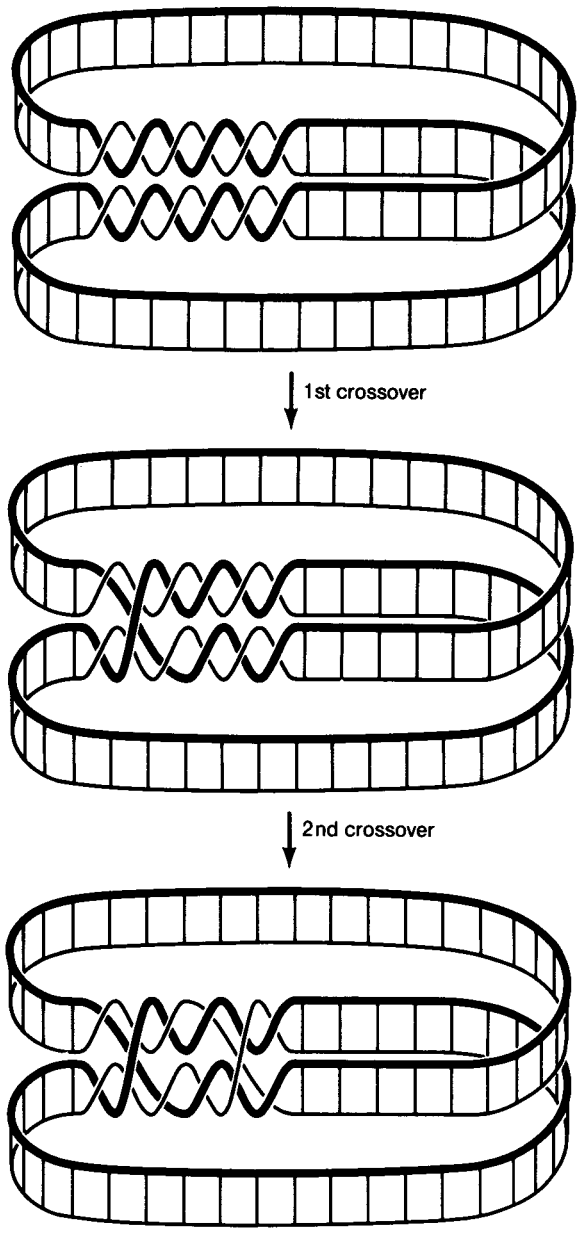

Figure 16. Formation of hemicatenanes by separation of single-strand crossover points. A ribbon representation for circular duplex DNA is shown with the W and C strands distinguished by bold and standard lines. The double helix is shown in the region of crossover to emphasize the nodes, but elsewhere base-pairing is schematized by vertical lines. The formation of the Holliday intermediate is by the first crossover that links the bold lines. Because the second crossover is separated from the first, a hemicatenane is generated. It is linked by the helical turns between the first and second join and by the Me of -1. Six single-strand intermolecular nodes and one duplex intermolecular node are formed that sum to -3 and -1, respectively, so that Ca = -4. In this example, the $^{ter}Wr^{s}$ term is contributed by the twisting of the W and C strands of different domains around each other rather than the writhe of the duplex DNA axis.

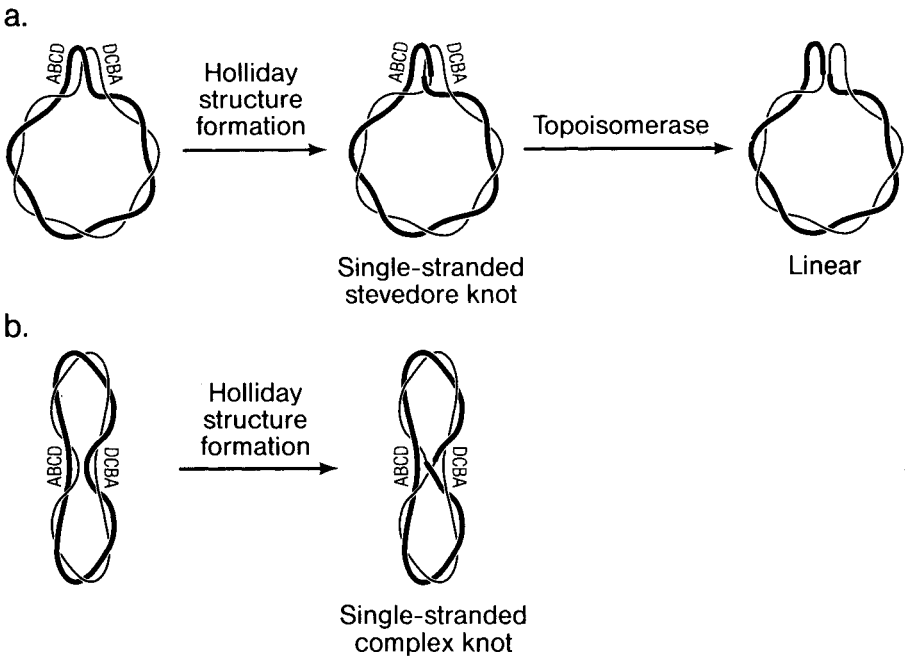

Figure 17. Unusual topological consequences of Holliday structure formation during inversion. The substrate contains two recombination sites shown as ABCD. The W and C strands are distinguished by bold and standard lines. In *a* the sites are less than one helical turn apart and in *b* they are more distant. Because of the inverse orientation of the sites, breakage at exactly the same point in the two sequences interrupts the W strand of one site but the C strand of the other. Reunion then links W to C in a continuous strand that is a molecular Möbius strip. Concomitant with this is single-strand knot formation; a stevedore-class knot in *a* and a complex knot in *b*. Both of these knots can be matured to recombinants by the usual resolution of a Holliday structure, but if the stevedore knot were untied by a topoisomerase a linear DNA with hairpin ends is formed. The stevedore knot could, in turn, be generated from the complex knot by branch migration and untangling by a topoisomerase.

terwindings into a domain that continually diminishes in size during migration.

We mention two instances where this analysis may be relevant to processes other than recombination. A topoisomerase could lift the restriction on branch migration in the inverted intermediate and allow migration to proceed until the stevedore knot is generated and then untied to give a linear duplex with hairpin termini (Fig. 17). A mechanism involving the combined action of a topoisomerase and a recombination enzyme could resolve the fused linear dimer chromosomes joined head-to-head that are the proposed product of replication of some hairpin termini of linear chromosomes (Blackburn and Szostak 1984). Second, the well-known genetic instability of long palindromes (Collins 1981) could result from an analogous process because migration of inversion Holliday intermediates could lead to chromosome breaks. This would explain the stability of inverted repeats separated by nonhomologous regions and the stability of some palindromes in certain recombination mutants (Leach and Stahl 1983).

CONCLUSIONS

We have presented a system for treating quantitatively the changes in linkage that occur when DNA is knotted or catenated by topoisomerases and by recombination enzymes. This expands on the concept of linking number, which describes the helical twist and the supercoiling of intact circular DNA. Our definition of the linkage value, and the terms that make it up, are entirely consistent with those used for linking number; all of these indices can be defined and understood in terms of nodes.

We have defined the parameters Ca and Kn to describe the interlinking of catenanes and knots. Using the concept of interdomainal and intradomainal nodes, we have sorted out those nodes that determine the value of Ca and Kn and those that affect supercoiling. The nodes introduced by the action of the enzyme itself, called Me, are also broken down this way.

The result is a simple algebra validated using ribbon models that describes precisely all the topological changes of topoisomerase reactions and of recombination of direct and inverted sites. The table of basic relationships correlates interdomainal writhe and Me with site alignment during recombination. Use of these definitions, equations, and this table leads to a clear way to describe the forms of supercoils and the complexity of knots and catenanes, a unified view of the diverse reactions of topoisomerases, specific conclusions about the topology of recombination by Tn*3* resolvase, a concise analysis of the possibilities for recombination by Int, and useful insights into the Holliday intermediate.

ACKNOWLEDGMENTS

We thank Frank Dean for helpful discussions, Ellis Myers for imaginative skillful drawings, and Elisabeth Lindheim for efficient preparation of the manuscript. The work was supported by National Institutes of Health grants GM 31655 and GM31657. M.A.K. was supported by National Institutes of Health grant GM 07281.

REFERENCES

BANCHOFF, T.F. and J.H. WHITE. 1975. The behavior of the total twist and the self linking number of a closed space curve under inversion. *Math. Scand.* **36:** 254.

BAUER, W.R. and J. VINOGRAD. 1975. Circular DNA. In *Basic principles of nucleic acid chemistry* (ed. P.O.P. Ts'o), vol. 2, p. 265. Academic Press, New York.

BETTER, M., C. LU, R.C. WILLIAMS, and H. ECHOLS. 1982. Site-specific DNA condensation and pairing mediated by the Int protein of bacteriophage lambda. *Proc. Natl. Acad. Sci.* **79:** 5837.

BLACKBURN, E.H. and J. SZOSTAK. 1984. The molecular structure of centromeres and telomeres. *Annu. Rev. Biochem.* **53:** 163.

BROWN, P.O. and N.R. COZZARELLI. 1979. A sign inversion mechanism for enzymatic supercoiling of DNA. *Science* **206:** 1081.

COLLINS, J. 1981. Instability of palindromic DNA in *E. coli*. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 409.

COZZARELLI, N.R. 1980. DNA topoisomerases. *Cell* **22:** 327.

CRICK, F.H.C. 1976. Linking number and nucleosomes. *Proc. Natl. Acad. Sci.* **73:** 2639.

DEAN, F., M.A. KRASNOW, R. OTTER, M.M. MATZUK, S.J. SPENGLER, and N.R. COZZARELLI. 1983. *Escherichia coli* type-1 topoisomerases: Identification, mechanism and role in recombination. *Cold Spring Harbor Symp. Quant. Biol.* **47:** 769.

FULLER, F.B. 1971. The writhing number of a space curve. *Proc. Natl. Acad. Sci.* **68:** 815.

GELLERT, M. 1981. DNA topoisomerases. *Annu. Rev. Biochem.* **40:** 879.

HEFFRON, F. 1983. Tn3 and its relatives. In *Mobile genetic elements* (ed. J.A. Shapiro), p. 223. Academic Press, New York.

KRASNOW, M.A. and N.R. COZZARELLI. 1983. Site-specific relaxation and recombination by the Tn3 resolvase: Recognition of the DNA path between oriented *res* sites. *Cell* **32:** 1313.

KRASNOW, M.A., M.M. MATZUK, J.M. DUNGAN, H.W. BENJAMIN, and N.R. COZZARELLI. 1983a. Site-specific recombination by Tn3 resolvase: Models for pairing of recombination sites. In *Mechanisms of DNA replication and recombination* (ed. N.R. Cozzarelli), vol. 10, p. 637. A.R. Liss, New York.

KRASNOW, M.A., A. STASIAK, S.J. SPENGLER, F. DEAN, T. KOLLER, and N.R. COZZARELLI. 1983b. Determination of the absolute handedness of knots and catenanes of DNA. *Nature* **304:** 559.

LEACH, D.R.S. and F.W. STAHL. 1983. Viability of λ phages carrying a perfect palindrome in the absence of recombination nucleases. *Nature* **305:** 448.

LIU, L.F., C.-C. LIU, and B.M. ALBERTS. 1980. Type II DNA topoisomerases: Enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break. *Cell* **19:** 697.

MIZUUCHI, K., M. GELLERT, R.A. WEISBERG, and H.A. NASH. 1980. Catenation and supercoiling in the products of bacteriophage λ integrative recombination *in vitro*. *J. Mol. Biol.* **141:** 485.

NASH, H.A. 1981. Integration and excision of bacteriophage λ: The mechanism of conservative site specific recombination. *Annu. Rev. Genet.* **15:** 143.

NASH, H.A. and T.J. POLLOCK. 1983. Site-specific recombination of bacteriophage lambda. The change in topological linking number associated with exchange of DNA strands. *J. Mol. Biol.* **170:** 19.

NASH, H.A., K. MIZUUCHI, L.W. ENQUIST, and R.A. WEISBERG. 1981. Strand exchange in λ integrative recombination: Genetics, biochemistry, and models. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 417.

POLLOCK, T.J. and H.A. NASH. 1980. Catenation of the products of integrative recombination: Comparison of supertwisted and nonsupertwisted substrates. In *Mechanistic studies of DNA replication and genetic recombination* (ed. B. Alberts), vol. 19, p. 953. Academic Press, New York.

———. 1983. Knotting of DNA caused by a genetic rearrangement. Evidence for a nucleosome-like structure in site-specific recombination of bacteriophage lambda. *J. Mol. Biol.* **170:** 1.

REED, R.R. 1981. Transposon-mediated site-specific recombination: A defined *in vitro* system. *Cell* **25:** 721.

ROLFSEN, D. 1976. *Knots and links*. Publish or Perish, Berkeley, California.

SCHUBERT, H. 1956. Two bridge knots. *Math. Z.* **65:** 33.

WANG, J.C. and L.F. LIU. 1979. DNA topoisomerases: Enzymes that catalyze the concerted breaking and rejoining of DNA backbone bonds. In *Molecular genetics* (ed. J.H. Taylor), vol. 3, p. 65. Academic Press, New York.

WHITE, J.H. 1969. Self linking and the Gauss integral in higher dimension. *Am. J. Math.* **91:** 693.

WHITE, J.H. and N.R. COZZARELLI. 1984. A simple topological method for describing stereoisomers of DNA catenanes and knots. *Proc. Natl. Acad. Sci.* **81:** 3322.