

# Introduction to DNA Topology

Isabel K. Darcy, Stephen D. Levene, and Robert G. Scharein

**Abstract** In this expository chapter we give an elementary introduction to DNA and to proteins that can knot and link circular DNA, with a special focus on recombination. We also describe the Ernst and Sumners tangle model of the action of proteins on circular DNA.

## 1 Introduction to DNA

In 1953, Rosalind Franklin and her student R. G. Gosling published their crystal structure of DNA upon which our current models of DNA are based. As discussed in her 1953 paper [36], “DNA is a helical structure” with “two co-axial molecules.” The “co-axial molecules,” shown as blue ribbons in Fig. 1a, refer to a chain of phosphate groups connected via sugar groups. The sugar groups are shown in red in Fig. 2. Each sugar group is connected to a base: A, T, G, or C. These bases spell out our genetic information, i.e., they form our DNA sequence. Using Franklin’s data without her knowledge, Watson and Crick came to similar conclusions regarding the structure of DNA and published their model in the same issue of Nature [90].

---

I.K. Darcy (✉)

Department of Mathematics and Program in Applied Mathematical and Computational Sciences,  
University of Iowa, 14 MLH, Iowa City, IA 52242, USA

e-mail: [idarcymath@gmail.com](mailto:idarcymath@gmail.com)

S.D. Levene

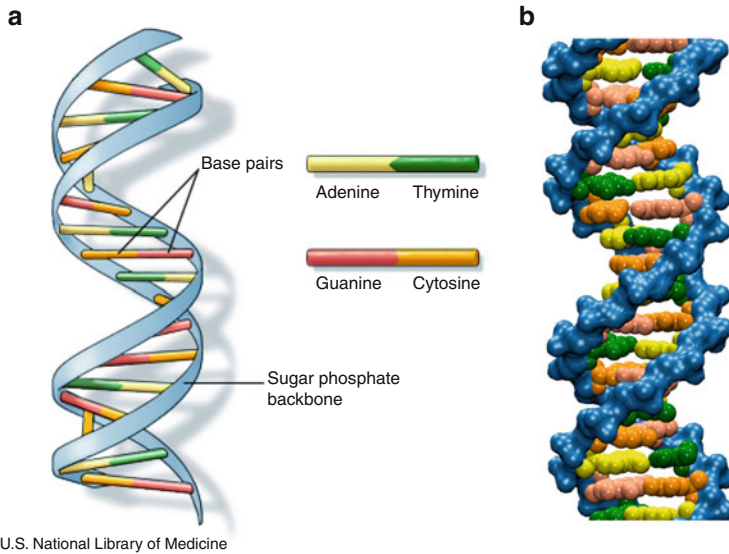
Departments of Bioengineering, Molecular and Cell Biology, and Physics, University of Texas  
at Dallas, 800 West Campbell Road, Richardson, TX 75080, USA

e-mail: [Stephen.Levене@utdallas.edu](mailto:Stephen.Levене@utdallas.edu)

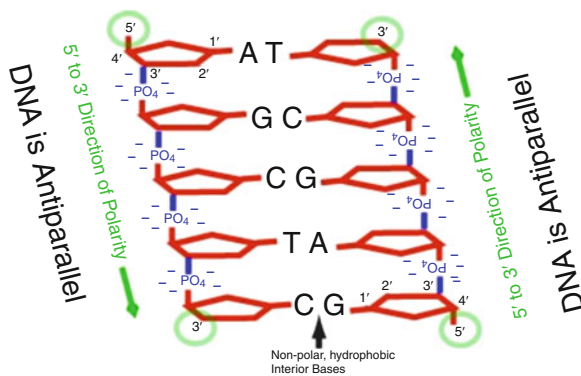
R.G. Scharein

Hypnagogic Software, Vancouver, BC, Canada

e-mail: [rob@knotplot.com](mailto:rob@knotplot.com)

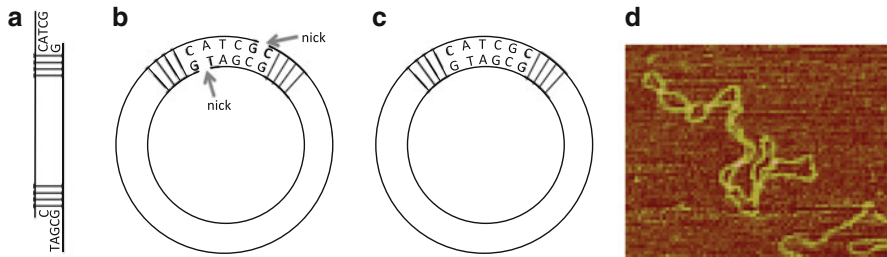


**Fig. 1** (a) Structure of DNA. Per Franklin and Gosling, the “period is  $34 \text{ \AA}$ ” and “one repeating unit contains ten nucleotides on each of two ... co-axial molecules.” They conclude “The phosphate groups lie on the outside of the structural unit, on a helix of diameter about  $20 \text{ \AA}$ ” and “the sugar and base groups must accordingly be turned inwards towards the helical axis.” [36]. The four bases are called adenine (A), thymine (T), cytosine (C), and guanine (G). Figure courtesy of the National Library of Medicine (NLM). (b) Computer model of a DNA molecule



**Fig. 2** Chemical structure of DNA. The sugars are shown in red. Note that the pairing of backbone strands is antiparallel (Figure from [50])

In most cases, the base A pairs with the base T while the base G pairs with the base C [90]. Thus, knowing the sequence of one strand of double-stranded DNA (dsDNA) means knowing the sequence of both strands. However, the two strands making up dsDNA are read in opposite directions. Rosalind Franklin noted that the two sugar–phosphate backbones are antiparallel [35]. A sugar residue contains five



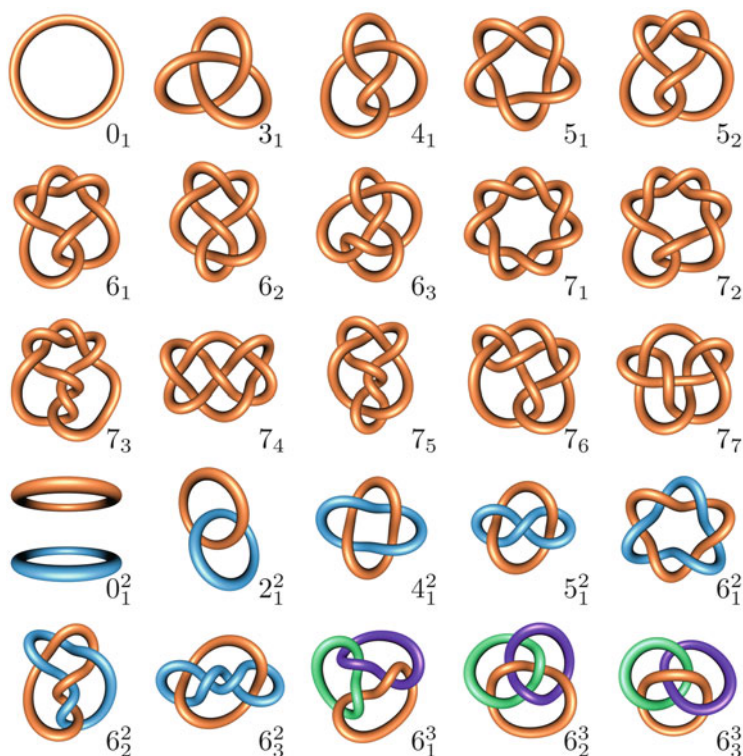
**Fig. 3** (a) Linear DNA with sticky ends. (b) Nicked circular DNA. (c) Closed circular DNA. (d) Atomic-force microscopy image of supercoiled DNA (Unpublished data courtesy of Dr. Alexandre Vetcher)

carbons that are numbered from  $1'$  to  $5'$ , as shown in Fig. 2 for the sugar in the upper left corner, which is connected to the base A. The  $5'$  carbon of this sugar is circled in green. Its  $3'$  carbon is connected via a phosphate bond to the  $5'$  carbon of the next sugar (which is connected to the base G). Hence the chemistry of this connection can be used to assign an orientation to a DNA strand. Sequences are read from  $5'$  to  $3'$ . Thus the strand on the left is read from top to bottom, and thus its sequence is AGCTC. The direction of the strand on the right goes from bottom to top. Hence its sequence is read GAGCT. Thus both AGCTC and GAGCT refer to exactly the same double-stranded DNA sequence.

In the laboratory, molecular biologists often work with linear DNA that has sticky ends. The end of the DNA is *sticky* if the portion at the end is single-stranded, as shown in Fig. 3a. This means that there are unpaired bases. If linear DNA contains two sticky ends that have complementary sequences, then if the DNA is sufficiently long, the linear DNA will circularize to form nicked circular DNA as shown in Fig. 3b. The DNA is called *nicked* because the phosphate backbone is not closed. A protein called ligase is needed to create a phosphodiester bond to close the nicks to form *closed circular DNA* (Fig. 3c). This closed DNA can be modeled by an annulus. Since the phosphate backbones are antiparallel, DNA cannot form a Möbius band (under normal circumstances). DNA has a preferred twist of about 10.5 base pairs per turn [36, 68, 85]. Closed circular DNA is called *relaxed* if it is as close as possible to its preferred twist. In nature, DNA is usually underwound, and hence it supercoils negatively (Fig. 3d). Since the DNA is underwound, the two strands are easier to pull apart for replication or transcription. For an elementary introduction to DNA, see [12]. For more on DNA topology, see [4].

## 2 DNA Knots and Topoisomerase

There are many beautiful knot tables in the literature and online. For an excellent introduction to knot theory, see [1]. Knots were first tabulated by Tait in the late 1800s [81]. The knot/link table shown in Fig. 4 was created by KnotPlot [72] based

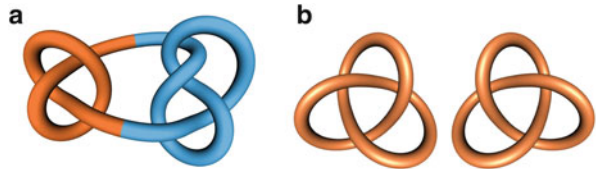


**Fig. 4** Knot/link table containing prime knots up to seven crossings and two- and three-component prime links up to six crossing. Twist knots include  $3_1$  (the trefoil knot),  $4_1$  (the Fig. 8 knot),  $5_2$ ,  $6_1$ , and  $7_2$ . Torus knots include  $3_1$ ,  $5_1$ , and  $7_1$ . The unknot,  $0_1$ , can also be considered to be a twist knot and a torus knot

on data provided by Dale Rolfsen. The knot  $n_k$  refers to the  $k$ th knot in the list of knots containing  $n$  crossings in their minimal crossing diagram. The superscript in the link table refers to the number of components. Mathematicians sometimes use the term “link” to include knots (and in rare cases, a link may be referred to as a knot). Molecular biologists normally use the term *catenane* when referring to links with at least two components. Most tables, including the one in Fig. 4, only contain *prime knots*. These are knots that cannot be subdivided into two or more simpler nontrivial knots. Knots that are not prime are called *composite*. Composite knots correspond to the operation of tying two separate knots sequentially in a piece of rope and closing the ends, as shown in Fig. 5a, where the individual prime knots are colored differently.

A knot is called *chiral* if it cannot be smoothly deformed into its own mirror image. The mirror image of a knot  $K$  is denoted  $K^*$ . The simplest example of a chiral knot is  $3_1$ ; it is shown together with its mirror image,  $3_1^*$ , in Fig. 5b. Most

**Fig. 5** (a) The composite knot  $3_1\#4_1$ . (b) Chiral pair of trefoil knots,  $3_1$  (in *gold*) and  $3_1^*$  (in *blue*)



knot tables like the one in Fig. 4 list only one enantiomer of a chiral pair. Knots that are not chiral are *achiral*. The simplest three achiral knots are  $0_1$ ,  $4_1$ , and  $6_3$ .

One of the most beautiful knot tables is the one created by topoisomerase I. This protein acting on nicked circular DNA, was able to create all different types of knots up to six crossings, 10 of the 16 possible seven-crossing knots, and a few eight- and nine- crossing knots [30]. There are two main types of topoisomerase. Type I topoisomerase will break and reconnect one strand of DNA, while type II topoisomerases will break and reconnect both strands of dsDNA. Thus type I topoisomerases can knot circular single-stranded DNA (ssDNA) as well as nicked DNA. Type II topoisomerases can knot dsDNA [89]. The normal function of topoisomerases is to keep DNA unknotted, unlinked, and properly supercoiled. For more on topoisomerase, see [87, 88].

### 3 Recombinases

The genome of any organism must possess two key characteristics. It must be stable enough to pass accurate information through inheritance, yet remain sufficiently dynamic to respond to selective environmental pressures. These requirements create a tension between genome integrity and flexibility. In all organisms, recombination systems are the principal mechanism that regulates genome stability. Chromosomal breakages and mutations stemming from problems in DNA replication or environmental stress can be repaired by recombination. Most organisms have multiple recombination pathways by which damage can be repaired, underscoring the importance of this process.

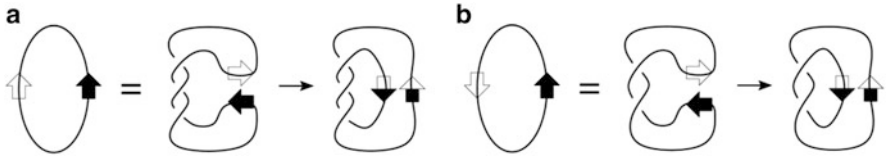
We focus here on several examples of site-specific recombination mechanisms. These processes involve interactions among specialized DNA-sequence elements that also contain specific binding sites for recombination proteins. The requirements for sequence specificity and specialized proteins distinguish site-specific recombination from general or homologous recombination, which can occur with arbitrary DNA sequences that share very high levels of sequence identity (see [58, 92] for reviews). Site-specific-recombination target sequences form the point of genetic exchange and usually are present in few copies in the genome. Often these sites are present in pairs; in the case of the bacteriophage-lambda integration site, only one copy is present in the *E. coli* genome. This extraordinary degree of specificity

leads to precisely defined genetic rearrangements. In the examples considered here, the rearrangements that occur are essentially uniquely defined.

Another important attribute of a site-specific recombination locus is the polarity of the recombination site. These loci are frequently nonpalindromic and therefore have an intrinsic polarity (Fig. 6). Recombination normally occurs only when a pair of recombination sites has been juxtaposed in a particular spatial alignment, thereby imparting both positional and orientational specificity to these systems [37,73]. This specificity has important biological consequences; moreover, the site-orientation specificity leads to the formation of specific DNA topologies in the recombination products. If two recombination sites are oriented in opposite polarities on a circular DNA molecule as shown in Fig. 6a, then the sites are said to be *inversely repeated*. Recombination on inversely repeated sites is called an *inversion* because it results in the inversion of one of the DNA segments between the two recombination sites with respect to the other DNA segment. In Fig. 6b, the two recombination loci are oriented in the same direction and are thus called *directly repeated*. Recombination on directly repeated sites results in the *deletion* (also called *excision*) of a DNA segment, changing the number of components of the substrate. The reverse reaction is called *integration*.

When supercoiled DNA substrates are used in reconstituted in-vitro (i.e., in the test tube) recombination reactions, it is possible to examine the topological changes that take place during recombination. For intramolecular recombination reactions, supercoiled plasmid substrates bearing inversely oriented sites generate knotted recombination products, whereas supercoiled substrates containing directly repeated sites generate topologically linked circles called catenanes (Fig. 6). The knots and catenanes that are formed during recombination are never random; instead, recombination generates a highly restricted subset of all the possible knotted or catenated structures that can be formed. For example, all of the knots with up to 13 crossings are known – there are over 12,000 topologically distinct knots. Integrative recombination on a circular substrate with inverted sites yields only seven of the possible knots containing up to 13 crossings, each containing an odd number of crossings. Among all possible recombination mechanisms that can lead to the formation of a knotted DNA product, the formation of this particular set of observed products can be ascribed uniquely to a particular mechanism. The topological specificity of site-specific recombination systems has been exploited to great effect in unraveling the mechanisms of many site-specific recombinases (see, e.g., [10, 21, 23, 34, 41, 79, 80, 82–84, 95]).

The complementarity of DNA strands normally plays a very limited role in site-specific recombination, more as a feature of specific recombinase–DNA interactions than a necessity for homologous pairing or strand exchange. Unlike other modes of recombination, site-specific recombination is conservative in that no DNA is gained or lost during the recombination reaction. This aspect of site-specific recombination applies both at the level of genetic information (recombination products are merely permutations of the original parental DNA) and at the level of actual DNA nucleotides (no DNA synthesis or nucleolytic degradation is involved).



**Fig. 6** (a) Recombination on inversely repeated sites on a circular DNA molecule can result in knotted DNA. (b) Recombination on directly repeated sites changes the number of components of the DNA substrate

All site-specific recombination systems that have been investigated to date fall into two superfamilies: the lambda-integrase and resolvase/invertase families. Particular examples from the lambda-integrase family are discussed below. The products of the lambda-integrase recombination reaction depend on the orientation and disposition of recombination sites; this variability permits systems such as lambda-integrative recombination to carry out excisive as well as integrative recombination in a highly regulated fashion. The two superfamilies are also distinct in terms of the intermediate structure of the DNA segments undergoing recombination; whereas lambda-integrase-type mechanisms proceed through a four-stranded DNA intermediate called a Holliday junction, the resolvase/invertase mechanisms do not. Site-specific recombination systems participate in a wide range of biological processes in both prokaryotes and eukaryotes: viral integration, antigenic variation, gene duplication and copy-number control, and the integration of antibiotic resistance cassettes. For a very nice review of site-specific recombination, including resolvase/invertase recombination, see [42].

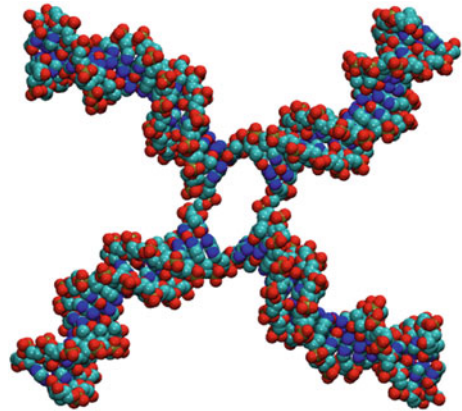
### 3.1 Holliday Junctions

In 1964, Robin Holliday proposed that recombination could be mediated by a hypothetical DNA structure consisting of four polynucleotide strands associated by a single-stranded crossover (Fig. 7) [48]. This structure, later to be named the Holliday junction, has played a central conceptual role in models of recombination. A large body of evidence has accumulated in the intervening decades that substantiates the role of these junctions in both general recombination mechanisms [46, 93, 94] and those belonging to the lambda-integrase superfamily of site-specific recombinases [3, 21].

Although the existence of this intermediate structure is no longer questioned, the details of Holliday junction geometry remain controversial. Since the early 1990s, a wide range of biochemical and biophysical tools have been used to characterize the conformation of these recombination intermediates, both as complexes with recombination proteins [6, 15, 39, 43, 49, 84] and as free DNA molecules [13, 14, 16, 18–20]. Many studies of protein-free junctions were focused on the structure and dynamics of immobile four-way DNA junctions, in which four synthetic DNA



**Fig. 7** Molecular model of a Holliday-like four-way DNA junction. The DNA sequences of the four strands in this structure lack the symmetry of a true Holliday junction, thereby inhibiting migration of the junction's branch point



strands designed with specific patterns of homology have been annealed together (Fig. 7) [51]. The limited homology among the DNA strands fixes the branch point of the junction; thus, such structures lack the ability to undergo branch migration, an essential isomerization step in general recombination. The extent to which the behavior of such immobile analogs actually mimics that of mobile junctions is an interesting issue that has remained largely unaddressed. However, it is clear that even immobile four-way junctions are conformationally quite flexible, a feature that is likely to be, if anything, more pronounced in fully mobile junctions [77]. Several groups have succeeded in obtaining high-resolution X-ray structures of four-way junctions [31, 61, 62]. These high-resolution structures exemplify many features that are consistent with those of immobile junctions based on studies in solution.

### 3.2 $\lambda$ -Int: Integration and Excision of Phage Genomes

The  $\lambda$ -integrase ( $\lambda$ -Int) system is vital to the lysogenic stage of the life cycle of bacteriophage  $\lambda$  and is one of the most intensively studied site-specific recombination systems. A notable feature of this system is the nonsymmetrical nature of the integrative and excisive recombination reactions: although the strand exchange activities are identical for both integration and excision of the phage- $\lambda$  genome, each reaction has distinct requirements for specific DNA sequences at the recombining loci and the subsets of protein cofactors involved in recombination.

Integration of phage  $\lambda$  occurs at a unique 25-bp site, termed *attB*, on the 4.6-Mbp *E. coli* chromosome. The catalytic activity for strand exchange resides in the  $\lambda$ -encoded integrase protein (Int), which functions in concert with a number of DNA-binding accessory proteins: the integration host factor (IHF) and factor for inversion stimulation (Fis) proteins of *E. coli*, and the  $\lambda$ -excisionase (Xis), which is phage-encoded. In contrast to the *attB* site, which by itself has negligible affinity for



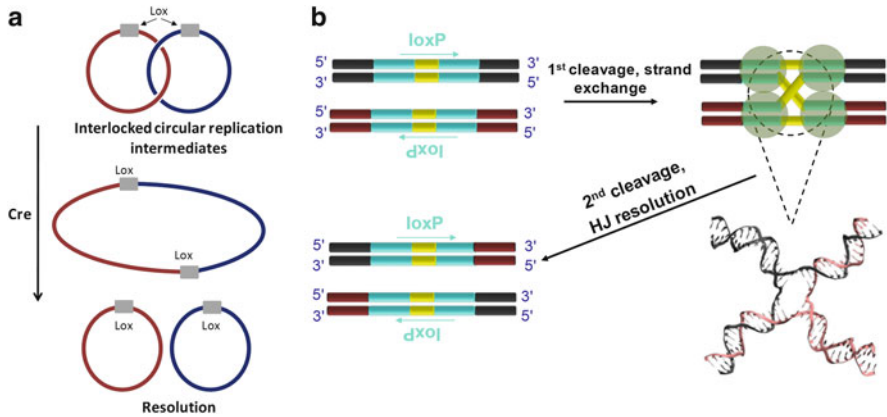
the recombination proteins, the recombination locus on the phage genome, *attP*, is about 250 bp in size and has multiple binding sites for Int and the accessory factors. Integrative recombination most likely involves assembly of Int and IHF proteins to form an organized nucleoprotein structure called the intosome [5], which subsequently captures a protein-free *attB* site during synapsis [67]. The products of the integrative recombination reaction are a functionally distinct pair of new recombination sites called *attL* and *attR* that are no longer competent to participate in subsequent rounds of integrative recombination. Instead, these sites are substrates for excisive recombination, a reaction that requires Fis and Xis in addition to Int and IHF. By coupling recombination to the intracellular levels of specific protein factors, tight regulation of the phage- $\lambda$  life cycle can be achieved in vivo. The topology of  $\lambda$ -Int recombination is discussed in [21].

### 3.3 *Cre and XerC/D: Excision and Resolution of DNA Dimers*

The genome of bacteriophage P1 is a 90-kbp circular molecule; as with all circular genomes, daughter molecules must be decatenated after replication [91]. This process is facilitated by a protein called Cre recombinase, a phage-encoded member of the  $\lambda$ -Int superfamily. The Cre mechanism acts on specific sites, denoted *loxP* in a multistep reaction scheme that involves fusion followed by resolution (Fig. 8a). A common feature of the  $\lambda$ -Int superfamily is phosphoryl transfer based on a catalytic tyrosine residue. The enzymatic reaction progresses in two distinct stages (Fig. 8b): an initial round of strand cleavage followed by DNA strand exchange to form a stable recombinase-bound Holliday junction. The junction is resolved by a second set of tyrosine-catalyzed cleavage and strand-exchange steps that lead to recombinant products.

The wild-type *loxP* target site for Cre is a 34-bp DNA sequence that consists of two 13-bp inverted repeats flanking an asymmetric 8-bp core region [38]. The core sequence confers an overall directionality on the *loxP* site. Recombination of directly repeated *loxP* sites leads to the exclusive formation of deletion products, whereas recombination of inversely repeated *loxP* sites results in an inversion of the intervening DNA sequence with respect to the parental substrate [84].

Normal replication of the *E. coli* chromosome yields intermediate forms consisting of multiply linked circular DNA molecules. The linked intermediates are resolved to unlinked monomers by the action of type II topoisomerases, most notably topo IV [78]. However, homologous recombination during replication generates concatenated dimers at a significant frequency; such structures cannot be resolved by topoisomerases. These dimers are instead resolved by the XerC/D system, which also belongs to the  $\lambda$ -Int superfamily. The activity of Xer is tightly coupled to that of FtsK, a molecular machine that controls the transport of DNA across the intercellular septum during cell division [66]. Cells lacking functional XerC or XerD genes cannot properly segregate daughter chromosomes. These



**Fig. 8** (a) Unlinking by Cre recombinase: Cre activity initially generates circular concatamers from linked circular substrates. This fused intermediate is subsequently unlinked via an excision reaction that is also mediated by Cre. (b) Mechanism of Cre acting on a pair of *loxP* target sequences. The *loxP* site consists of two inversely repeated, 13-bp Cre-binding sequence elements (cyan) flanking an 8-bp spacer region (yellow). Recombination takes place via ordered and reversible strand cleavage, exchange, and resolution reactions. The central intermediate is a Holliday junction, shown in an open, square planar conformation similar to that in Fig. 7

cells develop an anomalous filamentous-growth phenotype, in which cells elongate without dividing [59].

The target site for Xer recombination on the *E. coli* chromosome is a 28-bp sequence called *dif*. This sequence is located opposite the chromosomal replication origin and consists of a pair of 11-bp inverted repeats that flank a central 6-bp spacer region. Unlike the *loxP* site of Cre, the *dif* repeats are targeted by the Xer C and D subunits. In other respects, however, the similarities between Xer and Cre are more striking than their differences. Like the Cre–*loxP* mechanism, Xer recombination proceeds via a Holliday-junction intermediate [7].

Plasmids in *E. coli* can also become dimerized during replication. Resolution of these concatameric forms occurs via Xer activity at plasmid sequences such as *cer* and *psi*. These sequences contain the 28-bp core sequence from the chromosomal *dif* element in addition to flanking sequences that bind the accessory proteins PepA and either ArgR (in *cer*) [74] or ArcA (in *psi*) [8]. These proteins are required for recombination and play a role in organizing the active synaptic complex of proteins and DNA sequences needed for site pairing and strand exchange. This also ensures that recombination occurs exclusively via an intermolecular pathway involving directly repeated target sites. A similar mode of synaptic-complex organization occurs in gamma/delta resolvase recombination (based on a serine recombinase) and accounts for exclusivity of deletion in that system [71].

## 4 The Tangle Model for Protein Action

In this section, we start by giving some mathematical background on tangles and then describe how tangles are applied to study protein action. An  $N$ -string tangle is a collection of  $N$  disjoint arcs properly embedded in a three-dimensional ball (3-ball) that have their endpoints fixed on the 2-sphere boundary of the 3-ball. Examples of 2-string tangles are shown in Fig. 9. A tangle is *rational* if it can be formed from a zero crossing tangle by moving the endpoints in an arbitrary fashion with the constraint that they remain on the boundary of the 2-sphere and the arcs are confined to stay within the 3-ball. It is common practice and sufficient to consider only  $180^\circ$  rotations about the horizontal and vertical axes. Conway introduced rational tangles in a paper that was concerned with enumerating prime knots and links [17]. He discovered a curious relationship between rational tangles and the set of extended rational numbers<sup>1</sup> (hence the name *rational tangle*). He showed that two rational tangles are equivalent (in the sense that one may be converted into the other, keeping the boundary of the 3-ball fixed) if and only if the continued fractions corresponding to the individual tangles are equal to the same extended rational number. The details of this proof are beyond the scope of this chapter, however, a few examples will be illustrative. First of all, let us consider how to associate a tangle with a continued fraction. We denote a tangle by a list of integers  $(c_1, c_2, \dots, c_n)$ ,  $n$  odd, and create this tangle by starting with the 0 tangle and rotating about the horizontal axis by  $c_1 \times 180^\circ$ , followed by a rotation about the vertical axis by  $c_2 \times 180^\circ$ , alternating the axes in this way until we reach the end of the list. Since  $n$  is odd, we always end in horizontal twists. For the tangle  $(c_1, c_2, \dots, c_n)$ , we assign the rational number

$$r = c_n + \frac{1}{c_{n-1} + \frac{1}{\dots + \frac{1}{c_1}}}.$$

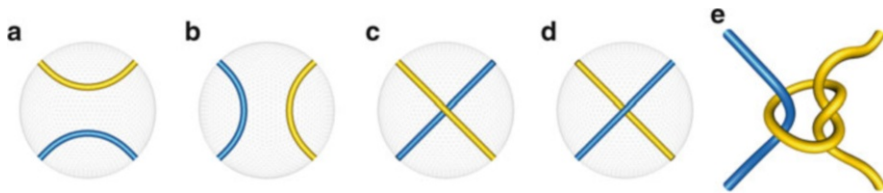
Figure 10 shows several rational tangles and one nonrational tangle. We can see that tangle  $(2, 3, 4)$  and tangle  $(-1, -1, -4, 1, 3)$  have the same rational number  $30/7$ , and therefore must be equivalent tangles from Conway's theorem.

Tangles that cannot be formed from the operations described above are known as *nonrational tangles*. Informally, these are tangles whose construction would require one of the endpoints of an arc to leave the 2-sphere, and pass through the 3-ball and around one of the arcs inside. An example is shown in Fig. 10e. Nonrational tangles are not uncommon in DNA; however, their analysis is considerably more complicated than the rational case, and we will not discuss them further in this chapter.

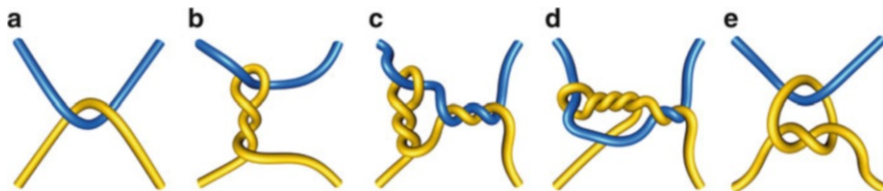
Ernst and Sumners [34] were the first to apply tangles to DNA biology. In their model, a protein complex binding  $N$  segments of DNA is represented by

---

<sup>1</sup>The rational numbers plus infinity.



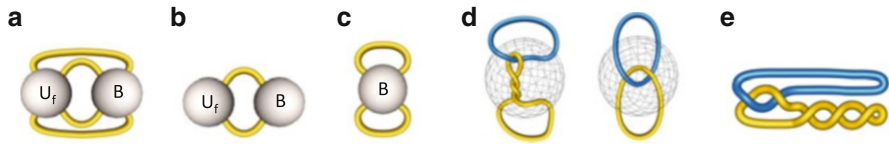
**Fig. 9** Various tangles. (a) The 0 tangle. (b) The infinity tangle. (c) the  $+1$  tangle. (d) The  $-1$  tangle. (e) The  $\frac{1}{2} + -\frac{1}{3}$  tangle



**Fig. 10** Four rational tangles with their associated rational numbers and one nonrational tangle. (a) The  $(2)$  tangle. (b) The  $(2, 3, 0)$  tangle  $= \frac{2}{7}$  tangle. (c) The  $(2, 3, 4)$  tangle  $= \frac{30}{7}$  tangle. (d) The  $(-1, -1, -4, 1, 3)$  tangle  $= \frac{30}{7}$  tangle. Note that the signs of the integers in  $(-1, -1, -4, 1, 3)$  determine the handedness of the crossings. (This tangle can be created in KnotPlot by using the command: `tangle 1z1z4z13o.`) (e) A nonrational tangle

a tangle ball, and the DNA itself by the disjoint arcs. Of course, this is a highly simplified model of the binding of proteins with DNA. A sphere is a very rough approximation to a protein complex, and the DNA is likely to exist in more complicated conformations than the arcs seen in the above illustrations. For example, the DNA likely winds around the tangle ball rather than being embedded within the 3-ball.

If at least two DNA segments are bound in a protein–DNA complex, then this complex is referred to as a synaptic complex. The protein complex together with the segments of DNA bound by protein is called a synaptosome. In many cases it is possible to prove that a tangle modeling a synaptosome is rational (e.g., [22, 23, 32–34, 47]), but there are also several biological reasons why rational tangles are the most likely models of synaptosomes. Although an upper bound for the number of DNA crossings that can be bound in a synaptosome has not yet been determined, it is believed that synaptosomes cannot be overly complicated. The simplest nonrational two-string tangles are the five-crossing tangles shown in Figs. 9e and 10e. Moreover, as protein complexes bind supercoiled DNA (Fig. 3d), rational tangles are likely models of synaptosomes, since such tangles are formed by adding twists. Lastly, a tangle is rational if and only if one can push the strings to lie on the boundary of the 3-ball so that the strings do not cross themselves on the 3-ball. Thus if DNA wraps around a protein complex without crossing itself and if the protein complex can be modeled by a topological sphere, the tangle modeling the synaptosome must be rational. Since DNA is negatively charged, it is unlikely to cross itself on the



**Fig. 11** (a)  $N(U_f + B)$ . (b) Tangle addition,  $U_f + B$ . (c) Numerator closure of a tangle  $N(B)$ . (d)  $N(\frac{2}{7}) = N(2)$ . (e) Same link as in (d), but shown in 2-bridge = 4-plat form

boundary of a protein. There are, however, protein complexes that can be modeled by higher-genus objects such as tori; but in all known cases, the protein–DNA complex can still be modeled by a spherical tangle.

Let  $K$  represent knotted circular DNA. For those who like to think of this circular DNA as living in  $S^3$ , then the tangle model of Ernst and Sumners [34] divides  $S^3$  into two tangles (Fig. 11a). One tangle,  $B$ , models the synaptosome (i.e., the protein complex together with the portion of DNA bound by protein), whereas the unbound DNA is in the complementary tangle,  $U_f$ . For simplicity, this is written as the tangle equation  $N(U_f + B)$ . Tangle addition,  $U_f + B$ , corresponds to the operation shown in Fig. 11b. The numerator closure operation is shown in Fig. 11c. The numerator closure of a rational tangle is a rational knot/link (Fig. 11d). By rotating the vertical crossings in a rational tangle so that they appear horizontal, it is easy to see that rational knots/links are equivalent to 2-bridge knots/links (also known as called 4-plats); see the example in Fig. 11e. If  $ac \geq 0$ , then two rational knots  $N(a/b) = N(c/d)$  are equivalent if  $a = c$  and  $bd^{-1} = 1 \pmod a$ .

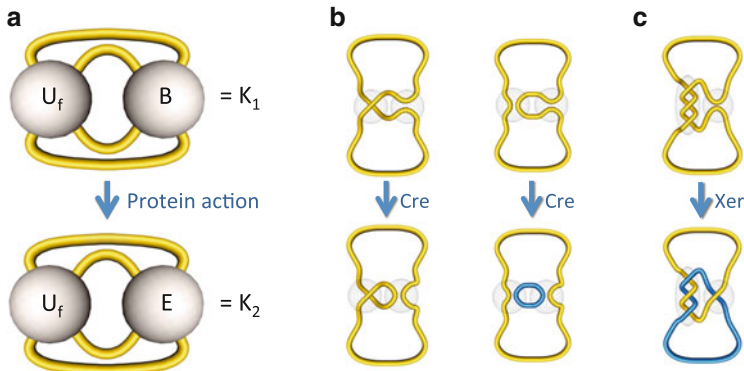
Let the tangle  $B$  represent the synaptosome before protein action, and let the tangle  $E$  represent it after protein action. Recall that the tangle  $U_f$  represents the DNA not bound by protein. A protein action that changes the knot  $K_1$  into the knot  $K_2$  is represented by the system of two tangle equations (Fig. 12):

$$N(U_f + B) = K_1, \tag{1}$$

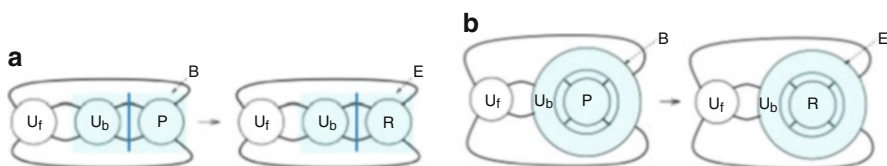
$$N(U_f + E) = K_2. \tag{2}$$

The starting conformation of the DNA,  $K_1$ , is called the *substrate*, while  $K_2$  is called the *product*. Different recombinases have different topological mechanisms. A tangle model of Cre recombination, is shown in Fig. 12b, and tangle model of Xer acting on *psi* sites, is shown in Fig. 12c. For a tangle model of Xer acting on Ftsk, see [76]. The software packages TangleSolve [70], TopoICE-R [25], and TopoICE-X [27] can be used to solve certain types of tangle equations and visualize the solutions.

In the original tangle model, the tangle  $B$  is divided into a sum of two tangles  $B = U_b + P$ , where the tangle  $P$  represents the local action of the protein, and the tangle  $U_b$  represents protein-bound DNA whose conformation is unchanged by protein action. Often proteins act on very short segments of DNA, and thus one can often assume that  $P$  is a zero-crossing tangle. Protein action is represented by



**Fig. 12** (a) Protein action is represented by the tangle equations  $N(U_f + B) = K_1, N(U_f + E) = K_2$ . (b) A tangle model for Cre recombination. (c) Tangle model for Xer acting on *psi* sites



**Fig. 13** (a) The Ernst and Sumners model of protein action is represented by the tangle equations  $N(U_f + U_b + P) = K_1, N(U_f + U_b + R) = K_2$ . (b) A more general tangle model

replacing the tangle  $P$  by the tangle  $R$ , as shown in Fig. 13 and modeled by the equations  $N(U_f + U_b + P) = K_1, N(U_f + U_b + R) = K_2$ . However, this model assumes that the local protein action can be separated from the remaining protein-bound DNA by a disk that intersects the strings of tangle  $B$  in exactly two points. However, this eliminates potential biologically relevant models. A more general tangle model is shown in Fig. 13b [23]. Note that the tangle  $B$  is transformed into the tangle  $E$  via replacing the subtangle  $P$  with the subtangle  $R$ . For more on rational subtangle replacement, see [9].

## 5 Concluding Remarks

We have only touched the surface regarding the topological modeling of protein–DNA complexes. In addition to modeling the action of proteins that can knot circular DNA, tangles can also be applied to probe the structure of multiple DNA segments in any stable protein–DNA complex regardless of the action (or inaction) of the protein via an experimental technique called difference topology [2, 26, 28, 40, 45, 52–57, 63–65]. This technique uses a recombinase or topoisomerase to trap crossings bound by the protein under study. This requires knowledge regarding

how the recombinase or topoisomerase acts. However, there are still unsolved problems regarding how these proteins act. With respect to recombination, we can only determine all solutions to the tangle equations modeling these reactions in special cases, even when the substrate and product are both rational knots/links (e.g. [10,24,29,47]). Although we can solve all tangle equations modeling topoisomerase action when the substrate and product are rational knots/links [24], there are still questions regarding preferred pathways [11,44,60,69,75,86].

**Acknowledgements** This research was supported in part by a grant from the joint Division of Mathematical Sciences (DMS) and National Institute of General Medical Sciences (NIGMS) Initiative to Support Research in the Area of Mathematical Biology (grant number NSF 0800285 to I.K.D. and NSF DMS-0800929 to S.D.L.). We wish to thank Massa Shoura for providing Figs. 1b, 8, and 9.

## References

1. C. Adams, *The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots* (American Mathematical Society, Providence, RI, 2004)
2. A. Akopian, S. Gourlay, H. James, S.D. Colloms, Communication between accessory factors and the Cre recombinase at hybrid psi-loxP sites. *J. Mol. Biol.* **357**(5), 1394–1408 (2006)
3. P. Argos, A. Landy, K. Abremski, J.B. Egan, E. Haggard-Ljungquist, R.H. Hoess, M.L. Kahn, B. Kalionis, S.V.L. Narayana, L.S. Pierson, N. Sternberg, J.M. Leong, The integrase family of site-specific recombinases: regional similarities and global diversity. *EMBO J.* **5**(2), 433–440 (1986)
4. A.D. Bates, A. Maxwell, *DNA Topology*. In Focus (IRL Press at Oxford University Press, Oxford/New York, 1993)
5. M. Better, C. Lu, R.C. Williams, H. Echols, Site-specific DNA condensation and pairing mediated by the int protein of bacteriophage lambda. *Proc. Natl. Acad. Sci. U S A* **79**(19), 5837–5841 (1982)
6. T. Biswas, H. Aihara, M. Radman-Livaja, D. Filman, A. Landy, T. Ellenberger, A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature* **435**(7045), 1059–1066 (2005)
7. G.W. Blakely, A.O. Davidson, D.J. Sherratt, Sequential strand exchange by XerC and XerD during site-specific recombination at dif. *J. Biol. Chem.* **275**(14), 9930–9936 (2000)
8. M. Bregu, D.J. Sherratt, S.D. Colloms, Accessory factors determine the order of strand exchange in Xer recombination at psi. *EMBO J.* **21**(14), 3888–3897 (2002)
9. D. Buck, K. Baker, The classification of rational subangle replacements between rational tangles. *Algebra. Geome Topol.* **13**, 1413–1463 (2013)
10. D. Buck, M. Mauricio, Connect sum of lens spaces surgeries: application to Hin recombination. *Math. Proc. Camb. Philos. Soc.* **150**(3), 505–525 (2011)
11. G.R. Buck, E.L. Zechiedrich, DNA disentangling by type-2 topoisomerases. *J. Mol. Biol.* **340**(5), 933–939 (2004)
12. C. Calladine, H. Drew, B. Luisi, A. Travers, *Understanding DNA: The Molecule and How it Works* (Elsevier, San Diego, CA, 2004)
13. S.M. Chen, F. Heffron, W.J. Chazin, Two-dimensional  $^1\text{H}$  NMR studies of 32-base-pair synthetic immobile Holliday junctions: complete assignments of the labile protons and identification of the base-pairing scheme. *Biochemistry* **32**(1), 319–326 (1993)
14. S.M. Chen, W.J. Chazin, Two-dimensional  $^1\text{H}$  NMR studies of immobile holliday junctions: nonlabile proton assignments and identification of crossover isomers. *Biochemistry* **33**(38), 11453–11459 (1994)



15. Y. Chen, U. Narendra, E.L. Iype, M.M. Cox, A.P. Rice, Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. *Mol. Cell* **6**(4), 885–897 (2000)
16. R.M. Clegg, A.I. Murchie, D.M. Lilley, The solution structure of the four-way DNA junction at low-salt conditions: a fluorescence resonance energy transfer analysis. *Biophys. J.* **66**(1), 99–109 (1994)
17. J.H. Conway, An enumeration of knots and links, and some of their algebraic properties, in *Computational Problems in Abstract Algebra (Proceedings of a Conference Held at Oxford, 1967)* (Pergamon, Oxford, 1970), pp. 329–358
18. J.P. Cooper, P.J. Hagerman, Gel electrophoretic analysis of the geometry of a DNA four-way junction. *J. Mol. Biol.* **198**(4), 711–719 (1987)
19. J.P. Cooper, P.J. Hagerman, Geometry of a branched DNA structure in solution. *Proc. Natl. Acad. Sci. U S A* **86**(19), 7336–7340 (1989)
20. J.P. Cooper, P.J. Hagerman, Analysis of fluorescence energy transfer in duplex and branched DNA molecules. *Biochemistry* **29**(39), 9261–9268 (1990)
21. N.J. Crisona, R.L. Weinberg, B.J. Peter, D.W. Summers, N.R. Cozzarelli, The topological mechanism of phage lambda integrase. *J. Mol. Biol.* **289**(4), 747–775 (1999)
22. M. Culler, C.M. Gordon, J. Luecke, P.B. Shalen, Dehn surgery on knots. *Bull. Am. Math. Soc. (N.S.)* **13**(1), 43–45 (1985)
23. I.K. Darcy, Biological distances on DNA knots and links: applications to Xer recombination. *J. Knot Theory Ramif.* **10**(2), 269–294 (2001). *Knots in Hellas '98*, Vol. 2 (Delphi)
24. I.K. Darcy, Solving unoriented tangle equations involving 4-plats. *J. Knot Theory Ramif.* **14**(8), 993–1005 (2005)
25. I.K. Darcy, R.G. Scharein, TopoICE-R: 3D visualization modeling the topology of DNA recombination. *Bioinformatics* **22**(14), 1790–1791 (2006)
26. I.K. Darcy, A. Bhutra, J. Chang, N. Druivenga, C. McKinney, R.K. Medikonduri, S. Mills, J. Navarra Madsen, A. Ponnusamy, J. Sweet, T. Thompson, Coloring the Mu transpososome. *BMC Bioinform.* **7**, 435 (2006)
27. I.K. Darcy, R.G. Scharein, A. Stasiak, 3D visualization software to analyze topological outcomes of topoisomerase reactions. *Nucleic Acids Res.* **36**(11), 3515–3521 (2008)
28. I.K. Darcy, J. Luecke, M. Vazquez, Tangle analysis of difference topology experiments: applications to a Mu protein-DNA complex. *Algebra. Geome Topol.* **9**(4), 2247–2309 (2009)
29. I.K. Darcy, K. Ishihara, R.K. Medikonduri, K. Shimokawa, Rational tangle surgery and Xer recombination on catenanes. *Algebra. Geome Topol.* **12**(2), 1183–1210 (2012)
30. F.B. Dean, A. Stasiak, T. Koller, N.R. Cozzarelli, Duplex DNA knots produced by *Escherichia coli* topoisomerase I. Structure and requirements for formation. *J. Biol. Chem.* **260**(8), 4975–4983 (1985)
31. B.F. Eichman, M. Ortiz-Lombardia, J. Aymami, M. Coll, P.S. Ho, The inherent properties of DNA four-way junctions: comparing the crystal structures of Holliday junctions. *J. Mol. Biol.* **320**(5), 1037–1051 (2002)
32. C. Ernst, Tangle equations. *J. Knot Theory Ramif.* **5**(2), 145–159 (1996)
33. C. Ernst, Tangle equations. II. *J. Knot Theory Ramif.* **6**(1), 1–11 (1997)
34. C. Ernst, D.W. Summers, A calculus for rational tangles: applications to DNA recombination. *Math. Proc. Camb. Philos. Soc.* **108**(3), 489–515 (1990)
35. R.E. Franklin, R.G. Gosling, Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* **172**(4369), 156–157 (1953)
36. R.E. Franklin, R.G. Gosling, Molecular configuration in sodium thymonucleate. *Nature* **171**(4356), 740–741 (1953)
37. M. Gellert, H. Nash, Communication between segments of DNA during site-specific recombination. *Nature* **325**(6103), 401–404 (1987)
38. K. Ghosh, G.D. Van Duyne, Cre-loxP biochemistry. *Methods* **28**(3), 374–383 (2002)
39. D.N. Gopaul, F. Guo, G.D. Van Duyne, Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination. *EMBO J.* **17**(14), 4175–4187 (1998)

40. S.C. Gourelay, S.D. Colloms, Control of Cre recombination by regulatory elements from Xer recombination systems. *Mol. Microbiol.* **52**(1), 53–65 (2004)
41. I. Grainge, M. Bregu, M. Vazquez, V. Sivanathan, S.C. Ip, D.J. Sherratt, Unlinking chromosome catenanes in vivo by site-specific recombination. *EMBO J.* **26**(19), 4228–4238 (2007)
42. N.D. Grindley, K.L. Whiteson, P.A. Rice, Mechanisms of site-specific recombination. *Ann. Rev. Biochem.* **75**, 567–605 (2006)
43. F. Guo, D.N. Gopaul, G.D. van Duyne, Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* **389**(6646), 40–46 (1997)
44. A.H. Hardin, S.K. Sarkar, Y. Seol, G.F. Liou, N. Osheroff, K.C. Neuman, Direct measurement of DNA bending by type IIA topoisomerases: implications for non-equilibrium topology simplification. *Nucleic Acids Res.* **39**(13), 5729–5743 (2011)
45. R.M. Harshey, M. Jayaram, The mu transpososome through a topological lens. *Crit. Rev. Biochem. Mol. Biol.* **41**(6), 387–405 (2006)
46. W.D. Heyer, Biochemistry of eukaryotic homologous recombination, in *Molecular Genetics of Recombination*, ed. by A. Aguilera, R. Rothstein (Springer, Heidelberg, 2007), pp. 95–133
47. M. Hirasawa, K. Shimokawa, Dehn surgeries on strongly invertible knots which yield lens spaces. *Proc. Am. Math. Soc.* **128**(11), 3445–3451 (2000)
48. R. Holliday, A mechanism for gene conversion in fungi. *Genet. Res. Camb.* **5**, 282–304 (1964)
49. K.E. Huffman, S.D. Levene, DNA-sequence asymmetry directs the alignment of recombination sites in the Flp synaptic complex. *J. Mol. Biol.* **286**, 1–13 (1999)
50. A. Hughes, Primary DNA molecular structure (2005). Connexions web site, Available at: <http://cnx.org/content/m11411/1.8/>
51. N.R. Kallenbach, R.I. Ma, N.C. Seeman, An immobile nucleic acid junction constructed from oligonucleotides. *Nature* **305**, 829–831 (1983)
52. E. Kilbride, M.R. Boocock, W.M. Stark, Topological selectivity of a hybrid site-specific recombination system with elements from Tn3 res/resolvase and bacteriophage P1 loxP/Cre. *J. Mol. Biol.* **289**(5), 1219–1230 (1999)
53. E.A. Kilbride, M.E. Burke, M.R. Boocock, W.M. Stark, Determinants of product topology in a hybrid Cre-Tn3 resolvase site-specific recombination system. *J. Mol. Biol.* **355**(2), 185–195 (2006)
54. S. Kim, A 4-string tangle analysis of DNA-protein complexes based on difference topology. ProQuest LLC, Ann Arbor, 2009. Ph.D. thesis, The University of Iowa
55. S. Kim, A generalized 4-string solution tangle of DNA-protein complexes. *J. Korean Soc. Ind. Appl. Math.* **15**(3), 161–175 (2011)
56. S. Kim, I.K. Darcy, Topological analysis of DNA-protein complexes, in *Mathematics of DNA Structure, Function and Interactions*. IMA Volumes in Mathematics and its Applications, vol. 150 (Springer, New York, 2009), pp. 177–194
57. K. Kimura, V.V. Rybenkov, N.J. Crisona, T. Hirano, N.R. Cozzarelli, 13S condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation. *Cell* **98**(2), 239–248 (1999)
58. S.C. Kowalczykowski, D.A. Dixon, A.K. Eggleston, S.D. Lauder, W.M. Rehrauer, Biochemistry of homologous recombination in Escherichia coli. *Microbiol. Rev.* **58**(3), 401–465 (1994)
59. C. Lesterlin, C. Pages, N. Dubarry, S. Dasgupta, F. Cornet, Asymmetry of chromosome Replichores renders the DNA translocase activity of FtsK essential for cell division and cell shape maintenance in Escherichia coli. *PLoS Genet.* **4**(12), e1000288 (2008)
60. Z. Liu, L. Zechiedrich, H.S. Chan, Action at hooked or twisted hooked DNA juxtapositions rationalizes unlinking preference of type-2 topoisomerases. *J. Mol. Biol.* **400**(5), 963–982 (2010)
61. P.K. Mandal, S. Venkadesh, N. Gautham, Structure of d(cgggtaccgc)4 as a four-way Holliday junction. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **67**(Pt 12), 1506–1510 (2011)
62. M. Ortiz-Lombardia, A. Gonzalez, R. Eritja, J. Aymami, F. Azorin, M. Coll, Crystal structure of a DNA holliday junction. *Nat. Struct. Biol.* **6**(10), 913–917 (1999)
63. S. Pathania, M. Jayaram, R. Harshey, Path of DNA within the Mu transpososome: transposase interaction bridging two Mu ends and the enhancer trap five DNA supercoils. *Cell* **109**, 425–436 (2002)

64. Z.M. Petrushenko, C.H. Lai, R. Rai, V.V. Rybenkov, DNA reshaping by MukB. Right-handed knotting, left-handed supercoiling. *J. Biol. Chem.* **281**(8), 4606–4615 (2006)
65. C. Price, A biological application for the oriented skein relation. ProQuest LLC, Ann Arbor, 2012. Ph.D. thesis, The University of Iowa
66. G.D. Recchia, M. Aroyo, D. Wolf, G. Blakely, D. Sherratt, FtsK-dependent and -independent pathways of Xer site-specific recombination. *EMBO J.* **18**(20), 5724–5734 (1999)
67. E. Richet, P. Abcarian, H.A. Nash, Synapsis of attachment sites during lambda integrative recombination involves capture of a naked DNA by a protein-DNA complex. *Cell* **52**(1), 9–17 (1988)
68. E.D. Ross, R.B. Den, P.R. Hardwidge, L.J. Marter, Improved quantitation of DNA curvature using ligation ladders. *Nucleic Acids Res.* **27**(21), 4135–4142 (1999)
69. V.V. Rybenkov, C. Ullsperger, A.V. Vologodskii, N.R. Cozzarelli, Simplification of DNA topology below equilibrium values by type II topoisomerases. *Science* **277**(5326), 690–693 (1997)
70. Y. Saka, M. Vazquez, Tanglesolve: topological analysis of site-specific recombination. *Bioinformatics* **18**(7), 1011–1012 (2002)
71. G.J. Sarkis, L.L. Murley, A.E. Leschziner, M.R. Boocock, W.M. Stark, N.D. Grindley, A model for the gamma delta resolvase synaptic complex. *Mol. Cell* **8**(3), 623–631 (2001)
72. R.G. Scharein, Interactive topological drawing. ProQuest LLC, Ann Arbor, 1998. Ph.D. thesis, The University of British Columbia
73. A.M. Segall, H.A. Nash, Architectural flexibility in lambda site-specific recombination: three alternate conformations channel the attL site into three distinct pathways. *Genes Cells* **1**(5), 453–463 (1996)
74. H. Senechal, J. Delesques, G. Szatmari, Escherichia coli ArgR mutants defective in cer/Xer recombination, but not in DNA binding. *FEMS Microbiol. Lett.* **305**(2), 162–169 (2010)
75. Y. Seol, A.H. Hardin, M.P. Strub, G. Charvin, K.C. Neuman, Comparison of DNA decatenation by Escherichia coli topoisomerase IV and topoisomerase III: implications for non-equilibrium topology simplification. *Nucleic Acids Res.* **41**, 4640–4649 (2013)
76. K. Shimokawa, K. Ishihara, M. Vazquez, Tangle analysis of DNA unlinking by the Xer/FtsK system. *Bussei Kenkyu* **92**, 89–92 (2009)
77. L.S. Shlyakhtenko, V.N. Potaman, R.R. Sinden, Y.L. Lyubchenko, Structure and dynamics of supercoil-stabilized DNA cruciforms. *J. Mol. Biol.* **280**(1), 61–72 (1998)
78. C. Sissi, M. Palumbo, In front of and behind the replication fork: bacterial type iia topoisomerases. *Cell Mol. Life Sci.* **67**(12), 2001–2024 (2010)
79. W.M. Stark, D.J. Sherratt, M.R. Boocock, Site-specific recombination by Tn3 resolvase: topological changes in the forward and reverse reactions. *Cell* **58**(4), 779–790 (1989)
80. D.W. Sumners, C. Ernst, S.J. Spengler, N.R. Cozzarelli, Analysis of the mechanism of DNA recombination using tangles. *Q. Rev. Biophys.* **28**(3), 253–313 (1995)
81. P. Tait, On knots. *Trans. R. Soc. Edinb.* **28**, 145–190 (1877)
82. M. Vazquez, D.W. Sumners, Tangle analysis of Gin site-specific recombination. *Math. Proc. Camb. Philos. Soc.* **136**(3), 565–582 (2004)
83. M. Vazquez, S.D. Colloms, D.W. Sumners, Tangle analysis of Xer recombination reveals only three solutions, all consistent with a single three-dimensional topological pathway. *J. Mol. Biol.* **346**(2), 493–504 (2005)
84. A.A. Vetcher, A.Y. Lushnikov, J. Navarra-Madsen, R.G. Scharein, Y.L. Lyubchenko, I.K. Darcy, S.D. Levenc, DNA topology and geometry in Flp and Cre recombination. *J. Mol. Biol.* **357**(4), 1089–1104 (2006)
85. A.A. Vetcher, A.E. McEwen, R. Abujarour, A. Hanke, S.D. Levene, Gel mobilities of linking-number topoisomers and their dependence on DNA helical repeat and elasticity. *Biophys. Chem.* **148**(1–3), 104–111 (2010)
86. A. Vologodskii, Theoretical models of DNA topology simplification by type IIA DNA topoisomerases. *Nucleic Acids Res.* **37**(10), 3125–3133 (2009)
87. J.C. Wang, Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.* **3**(6), 430–440 (2002)

88. J.C. Wang, *Untangling the Double Helix: DNA Entanglement and the Action of the DNA Topoisomerases* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2009)
89. S.A. Wasserman, N.R. Cozzarelli, Supercoiled DNA-directed knotting by T4 topoisomerase. *J. Biol. Chem.* **266**(30), 20567–20573 (1991)
90. J.D. Watson, F.H. Crick, Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**(4356), 737–738 (1953)
91. G. Witz, A. Stasiak, DNA supercoiling and its role in DNA decatenation and unknotting. *Nucleic Acids Res.* **38**(7), 2119–2133 (2010)
92. C. Wyman, R. Kanaar, Homologous recombination: down to the wire. *Curr. Biol.* **14**(15), R629–R631 (2004)
93. C. Wyman, D. Ristic, R. Kanaar, Homologous recombination-mediated double-strand break repair. *DNA Repair* **3**(8–9), 827–833 (2004)
94. K. Yamada, M. Ariyoshi, K. Morikawa, Three-dimensional structural views of branch migration and resolution in DNA homologous recombination. *Curr. Opin. Struct. Biol.* **14**(2), 130–137 (2004)
95. W. Zheng, C. Galloy, B. Hallet, M. Vazquez, The tangle model for site-specific recombination: a computer interface and the TnpI-IRS recombination system, in *Knot Theory for Scientific Objects-Proceedings of the International Workshop on Knot Theory for Scientific Objects held in Osaka (Japan), March 8–10, 2006*, vol 1 ed. by Akio Kawauchi. OCAMI Studies (Osaka Municipal universities Press, Sakai, 2007), pp. 251–271, [http://www.omup.jp/modules/tinyd1/index.php?id=25&easiestml\\_lang=en](http://www.omup.jp/modules/tinyd1/index.php?id=25&easiestml_lang=en)