



A Mathematical Model for Enzymatic Action on DNA Knots and Links

J. C. MISRA AND S. MUKHERJEE

Department of Mathematics
Indian Institute of Technology
Kharagpur 721302, India

A. K. DAS

Department of Biotechnology
Indian Institute of Technology
Kharagpur 721302, India

(Received and accepted October 2003)

Abstract—This paper deals with some studies pertaining to nonprocessive recombinase *viz.* Topoisomerases III, IV. The mathematics of tangles is found to be very useful in studying topoisomerases and recombinases (processive and nonprocessive). It has been seen that the enzyme acts on the DNA if it is in a certain configuration. Electron micrographs of the enzyme-DNA complex show the enzyme as a blob with DNA looping out of it, but they are unable to determine the configuration of the DNA within the blob. By using knot theory and tangle calculus, the configuration of DNA within the enzyme blob as well as the enzyme action has been determined in some cases.
© 2004 Elsevier Ltd. All rights reserved.

Keywords—DNA, Knots, Links, Enzyme, Topoisomerases, Recombinases.

1. INTRODUCTION

One of the longstanding issues in molecular biology is the three-dimensional structure (shape) of proteins and deoxyribonucleic acid (DNA) in solution in the cell and the relationship between structure and function. Usually, protein and DNA structures are determined by X-ray crystallography or electron microscopy. Because of the close packing needed for crystallization and the manipulation required to prepare a specimen for electron microscopy, these methods provide little direct evidence for molecular shape in solution.

DNA can be viewed as two very long curves that are intertwined millions of times, linked to other curves, and subjected to four or five successive orders of coiling to convert it into a compact form for information storage. Duplex DNA consists of two backbone strands wound about each other in right-handed helical fashion. Each strand consists of sugar phosphate backbone with a nitrogenous base attached to each sugar. The DNA of most bacteria and viruses are circular. Although human DNA is linear, it is extremely long and tacked down to a protein scaffold

The authors are grateful to the Ministry of Human Resource and Development, New Delhi, for supporting this investigation.

at various points on the DNA. This periodic attachment endows human DNA with topological constraints similar to those for circular DNA. These topological constraints can interfere with vital metabolic cellular processes such as replication and transcription. Enzymes are required to solve these topological entanglement problems that arise through cellular metabolism and replication. In this case, topoisomerases, which are enzymes that mediate the passage of one segment of DNA through an enzyme-bridged transient break in the backbone strands of another DNA segment, are responsible for unlinking the DNA. Other enzymes called recombinases break two DNA segments and interchange the ends, resulting in an exchange of genetic information. Tangle calculus has been successfully used to study recombinases. The topological approach to enzymology is an experimental protocol in which the descriptive and analytical powers of topology and geometry are employed in an indirect effort to determine the enzyme mechanism and the structure of active enzyme-DNA complexes *in vitro*. The packing, twisting, and topological constraints all taken together mean that topological entanglement poses serious functional problems for DNA. This entanglement would interfere with, and be exacerbated by, the vital life processes of replication, transcription, and recombination. For information retrieval and cell viability, some geometric and topological features must be introduced into the DNA, and others quickly removed. Some enzymes maintain proper geometry and topology by passing one strand of DNA through another by means of a transient enzyme-bridged break in one of the DNA strands. Other enzymes break the DNA apart and recombine the ends by exchanging them, a move performed by recombinases. Recently, it has been found that topoisomerases *viz.* Topoisomerase III and IV also help in DNA recombination where the recombination is nonprocessive. The description and quantization of the three-dimensional structure of DNA and the changes in DNA structure due to the action of these enzymes requires extensive use of geometry and topology in molecular biology. This use of mathematics as an analytic tool is especially important because there is no experimental way to observe the dynamics of enzymatic action directly. The DNA knots and links of the reaction product DNA molecules are observed by gel electrophoresis and electron microscopy. By observing the changes in geometry (supercoiling) and topology (knotting and linking) in DNA caused by an enzyme, the enzyme mechanism can be described and quantized.

The topological approach to enzymology poses an interesting challenge for mathematicians as to how one can deduce enzyme mechanisms from the observed changes of DNA geometry and topology. This requires the construction of mathematical models for enzyme action and the use of these models to analyze the results of topological enzymology experiments. The entangled form of the product DNA knots and links contains information about the enzymes that made them. In addition to utility in the analysis of experimental results, the use of mathematical models forces all of the background assumptions about the biology to be carefully laid out.

In 1990, Ernst and Sumners [1] used tangle model to analyze the Tn3 resolvase site-specific recombination system. They proved mathematically that, in a processive recombination event, Tn3 resolvase binds to its unknotted, negatively supercoiled substrate (sites in direct repeat), fixes three negative supercoils, and each round of recombination introduces a positive crossing in the domain. Darcy [2] modeled the Xer recombinase using 4-plat oriented equation. But since Xer is nonprocessive, the model gave an infinite number of solutions. The solutions of the model depended upon the initial assumptions that were made.

It has been observed that the circular DNA products produced by *in vitro* enzymology experiments fall into the mathematically well-understood family of 4-plats. This family consists of knot and link configurations produced by patterns of plectonemic supercoiling of pairs of strands about each other. All "small" knots and links are members of this family—more precisely, all prime knots with crossing number less than eight and all prime (two-component) links with crossing number less than seven are 4-plats. For *in vitro* topological enzymology, we can regard the enzyme mechanism as a machine that transforms 4-plats into other 4-plats. We need a mathematical language for describing and computing these enzyme-mediated changes. In many enzyme-DNA reactions, a pair of sites that are distant on the substrate circle are juxtaposed in

space and bound to the enzyme. The enzyme then performs its topological moves, and the DNA is then released. We need a mathematical language to describe configurations of linear strings in a spatially confined region. This is accomplished by means of the mathematical concept of tangles, which were introduced into knot theory by Conway [3]. Tangle theory is knot theory inside a 3-ball with the ends of the strings firmly glued down.

It is known that the resulting protein-DNA tangle is rational, since any tangle whose strings can be continuously deformed into the boundary of the defining ball is automatically rational. There is a classification scheme for rational tangles that is based on a standard form, i.e., is a minimal alternating diagram. The classifying vector for a rational tangle is an integer entry vector (a_1, a_2, \dots, a_n) of odd or even length, with all entries (except possibly the last) nonzero and having the same sign and with $|a_1| > 1$. Due to the requirement that $|a_1| > 1$ in the classifying vector convention for rational tangles, the corresponding tangle projection must have at least two nodes. There are four rational tangles $\{(0); (0; 0); (1); (-1)\}$ that are exceptions to this convention ($|a_1| = 0$ or 1).

We will use tangles to build a model that will compute the topology of the pre- and post-recombination synaptic complex in a single recombination event, given knowledge of the topology of the substrate and product. In site-specific recombination of circular DNA substrate, two kinds of geometric manipulation of the DNA occur. The first is a global ambient isotopy, in which a pair of distant recombination sites are juxtaposed in space and the enzyme binds to the molecule(s), forming the synaptic complex. Once synapsis is achieved, the next move is local and is entirely due to enzymatic action. Within the region occupied by the enzyme, the substrate is broken at each site, and the ends are recombined. We will model this local move. Within the region controlled by the enzyme, it breaks the DNA at each site and recombines the ends by exchanging them. We model the enzyme as a 3-ball. The synaptosome consisting of the enzyme and bound DNA forms a 2-string tangle.

THEOREM 1. *Let U and R be tangles such that $N(U + iR) = 4\text{-plat}$ for some $i \geq 2$, and $N(U + jR) \neq N(U + iR)$ for some j . Then R is a rational tangle. If $i \geq 3$, then R is an integral triangle.*

PROOF. If R were locally knotted, then $N(U + iR)$, $i \geq 2$ would be composite. Since 4-plats are prime, R cannot be locally knotted. Suppose R is a prime tangle. By tangle properties $U + (i - 1)R$ is rational or locally knotted and R prime implies that $(i - 1)R$ prime and U must be ∞ -tangle or locally knotted.

Now U cannot be an ∞ -tangle. If U were the infinity tangle, then $N(U + iR) = D(iR) = D(R) \# \dots \# D(R)$. Since 4-plats are prime, $D(R) = \text{unknot}$. But $N(U + iR) = D(iR) = \text{unknot} = D(jR) = N(U + jR)$, a contradiction. Thus, if U is locally unknotted, R must be rational.

If $i \geq 3$, then R does not have parity ∞ since 4-plats have at most two components. If $i \geq 3$, U is locally unknotted, and R is not integral, then if U is not integral, $U + R$ and $(i - 1)R$ are prime. But $N(U + iR) = 4\text{-plat}$ would then contradict the tangle property. If U is integral and if R is rational, then $N(U + iR) = 4\text{-plat}$, $i \geq 3$, if and only if R is integral. Thus, if U is locally unknotted and $i \geq 3$, then R must be integral.

Suppose U is locally knotted. Then, if U' is the tangle formed from U by removing the local knot, then $N(U' + iR) = \text{unknot}$, since 4-plats are prime. $N(U + jR) \neq N(U + iR)$ implies that $N(U' + jR) \neq N(U' + iR)$. Since the unknot is a 4-plat and U' is locally knotted, R is rational if $i \geq 2$ and integral if $i \geq 3$. ■

THEOREM 2. *If $N(U + P) = 4\text{-plat}$ and $N(U + R) = 4\text{-plat}$ where $P = a_1/b_1$, $R = a_2/b_2$, $a_1b_2 - a_2b_1 \neq \pm 1$, then U is either a rational tangle or ambient isotopic to a sum of two rational tangles.*

PROOF. If $N(U + P) = 4\text{-plat}$ and $N(U + R) = 4\text{-plat}$ where $P = a_1/b_1$, $R = a_2/b_2$, $a_1b_2 - a_2b_1 \neq \pm 1$, then the cyclic surgery theorem implies that the double branched cover of the tangle U is a

Seirfert fibered space. Ernst proved that this means that U is ambient isotopic to a Montesinos tangle. ■

THEOREM 3. *Let U and R be tangles such that $N(U + iR) = K_1$ for $0 \leq I \leq 3$, where K_i s are 4-plats, and $\{K_1, K_2, K_3\}$ represent at least two different link or knot types. Then there is at most one solution for U and U is either rational or the sum of two rational tangles.*

THEOREM 4. *Let $E = t/w$ - tangle, $(w, t) = 1$, and $ay - bx = 1$. Then, the following are equivalent for $|t| \geq 2$. For $t = \pm 1$, (2) and (3) are equivalent and imply (1).*

- (1) $d_R(N(a/b), N(z/v)) \leq 1$.
- (2) If $w = \pm 1 \pmod t$, $N(z/v) = N((tb + w)a)/(-ty + wx)$ or $N((-tx + (tk + w)a)/(-ty + (tk + w)b))$.
 Else $w \cong \pm 1 \pmod t$ and $N(z/v) = N((tp^2b + sa)/(-tp^2y - sx))$ or $N((-tp^2x + sa)/(-tp^2y + sb))$.

Where $s = tp(-q + pk) \pm 1$, $(p, q) = 1$, $p > 0$

- (3) $N(a/b) = N(U + 0)$ and $N(z/v) = N(U + t/w)$ has the following solutions when $|t| \geq 2$.
 If $w \neq \pm 1 \pmod t$, then U must be rational and $U = a/(b + ka)$ or $a/(-x + ka)$.
 If $w = \pm 1 \pmod t$, then U must be ambient isotopic to a sum of at most two rational tangles and $U = (U_1 + U_2) \circ (h, 0)$ where $U_1 = (-bja(d - kj))/(pb + a(pk - q))$ or $(xj + a(d - kj))/(-px + a(pk - q))$ and $U_2 = j/p$, $pd - qj = 1$, and $h = (-w \pm 1)/t$ if $(-w \pm 1) \in Z$. If $t = \pm 1$, then the above list contains all solutions when U is ambient isotopic to a sum of rational tangles.

PROOF. $d_R(N(a/b), N(z/v)) \leq 1$ if and only if there exists a U such that $N(a/b) = N(U + 0)$ and $N(z/v) = N(U + t/w)$. By Ernst's Theorem 2, U is either a rational tangle or ambient isotopic to the sum of two rational tangles. If U is a rational tangle, $N(U + 0) = N(a/b)$ implies by tangle fact that $U = a/(b + ka)$ or $a/(-x + ka)$ and $N(z/v) = N((tb + w)a)/(-ty + wx)$ or $N((-tx + (tk + w)a)/(-ty + (tk + w)b))$. If U is ambient isotopic to the sum of two rational tangles, $U_1 + U_2$, then since $N(U + 0)$ is a 4-plat, $U = (U_1 + U_2) \circ (h, 0)$. Solving $N((U_1 + j/p) \circ (h, 0) + 0) = N(a/b)$ implies $U_1 = (-bja(d - kj))/(pb + a(pk - q))$ or $(xj + a(d - kj))/(-px + a(pk - q))$ and $U_2 = j/p$, $pd - qj = 1$. If $N((U_1 + j/p) \circ (h, 0) + t/w) = N((U_1 + j/p + t/(ht + w)))$. If U_1 or U_2 are nonintegral, $N((U_1 + U_2 + t/(ht + w)))$ is a 4-plat if and only if $ht + w = \pm 1$, i.e., $w = \pm 1 \pmod t$, in which case, $h = (-w \pm 1)/t$ if $(-w \pm 1) \in Z$.

Again by tangle properties, if $s = tp(-q + pk) \pm 1$, $N(U_1 + U_2 + \pm t) = N((tp^2b + sa)/(-tp^2y - sx)) = N((-tp^2x + sa)/(-tp^2y + sb))$. ■

THEOREM 5. *If $N(U + f_1/g_1) = \text{unknot}$ and $N(U + f_2/g_2) = N(2z/1)$ where $f_1g_2 - f_2g_1 = \pm 1$, then U is rational.*

LEMMA 1. *If $N(U_i + P) = \text{unknot}$, $I = 1, 2$, and $U_1 \neq U_2$, then P is rational.*

THEOREM 6. *If $N(U + 0/1) = N(1/0)$ and $N(U + 1/w) = N(2k/1)$, then U is rational.*

COROLLARY 1. *Suppose $bx - ay = 1$, $N(U + 0/1) = N(a/b)$ and $N(U + t/w) = N(z/v)$ where $N(a/b)$ and $N(z/v)$ are unoriented 4-plats. If $w \neq \pm 1$ or if U is rational, then $t/w = (xz - av')/(bv' - yz - kt)$ and $U = a/(b + ka)$ OR $t/w = (bz - av')/(xv' - yz - kt)$ and $U = a/(x + ka)$ where v' is any integer such that $v'v^{\pm 1} = 1 \pmod z$. If $w = \pm 1 \pmod t$, then t divides $z\Upsilon a$.*

2. THE MODEL

The model that will be formulated here is based upon the following assumptions.

MATHEMATICAL ASSUMPTION 1. $E = O_b + P$, where O_b contains the entire DNA that is bound to the enzyme or to the accessory proteins, except for the recombination sites that are contained in P (the parental tangle).

MATHEMATICAL ASSUMPTION 2. *The recombinase action corresponds to a tangle surgery where the tangle P is changed by the tangle R (the recombinant tangle), i.e., the enzymatic action corresponds to $N(O + P) = \text{substrate}$ and $N(O + R) = \text{product}$.*

MATHEMATICAL ASSUMPTION 3. *Processive recombination acts by tangle addition.*

BIOLOGICAL ASSUMPTION 1. *The recombination mechanism is constant, independent of the geometry (supercoiling) and topology (knotting and linking) of the substrate population, i.e., P , R , and O_b are constants.*

BIOLOGICAL MODEL FOR RECOMBINASES.

- Initial configuration.
- The accessory proteins fix three negative crossings in the domain. Topoisomerase III and IV bind to the two recombination sites.
- Idea: the proteins and the three negative crossings remain fixed.
- One round of recombination produces one negative crossing in the domain.
- After recombination, the enzyme releases the molecule.

BIOLOGICAL MODEL FOR UNKNOTTED SUBSTRATES.

- Substrate = unknotted circular DNA with sites in direct repeat.
- $K_1 = b(1, 1) = \langle 1 \rangle$ [where K s are 4-plats].
- Product = 4-crossings right-handed torus link with antiparallel sites.
- $K_1 = b(4, 3) = \langle 1, 2, 1 \rangle$.

BIOLOGICAL MODEL FOR CATENATED SUBSTRATES.

- Substrate = 6-crossings right-handed torus link with antiparallel sites.
- $K_1 = b(6, 5) = \langle 1, 4, 1 \rangle$.
- Product = 7-crossings knot or link.

3. MATHEMATICAL EQUATIONS BASED ON THE MODEL

Tangle equations for unknotted substrates can be listed as

- (i) $N(O + P) = \langle 1 \rangle = b(1, 1)$, and
- (ii) $N(O + R) = \langle 1, 2, 1 \rangle = b(4, 3)$,

together with the assumptions

- (a) $P = (0)$,
- (b) $R = (k)$, k nonzero integer, and
- (c) O is rational or sum of two rational tangles.

We have to solve for O and R .

Tangle Equations for Catenated Substrates

In the case of catenated substrates, our task will be to solve O and R from the equations

- (i) $N(O + P) = \langle 1, 4, 1 \rangle = b(6, 5)$, and
- (ii) $N(O + R) = K_2 = 7\text{-crossings knot or link}$,

in which

- (i) $P = (0)$,
- (ii) $R = (k)$, k nonzero integer,
- (iii) O is rational or sum of two rational tangles,
- (iv) K_2 is a 4-plat.

While solving for O and R , we encounter the following problems.

- i. Xer recombination is not processive. The action on substrates with a single topology provides only two tangle equations.

- ii. For known P (rational) and K (4-plat): $N(O + P) = K$ has infinitely many solutions for O , and
- iii. for known P (rational), K_1 and K_2 (both 4-plats), $N(O + P) = K_1$, $N(O + R) = K_2$ do not lead a unique solution.

In order to solve the tangle equations, we intend to make use of the minimum possible assumptions, with an aim to put forward the results as realistic as possible.

4. RESULTS

Unknotted Substrates

By using the properties of tangle calculus given in the Appendix, we find that when O is rational, the solutions to the tangle equations are

- $O = (-3, 0)$ and $R = (-1)$.
- $O = (-5, 0)$ and $R = (+1)$.
- $O = (1)$ and $R = (3)$.
- $O = (-1)$ and $R = (5)$.

The last two cases produce 4-crossing links with wrong side alignment and hence are to be discarded.

When O is the sum of two rational nonintegral tangles, there exist no solutions.

The results may now be summarized as follows.

- (i) $N(O + P) = b(1, 1)$.
- (ii) $N(O + R) = b(4, 3)$ with sites in antiparallel $P = (0)$, $R = (k)$.
- (iii) O is rational or the sum of two rational tangles.
- (iv) The only solutions to the system are

$$\begin{aligned}
 O &= (-3, 0), & R &= (-1); \\
 O &= (-5, 0), & R &= (+1).
 \end{aligned}$$

For catenated substrates, when O is rational, the solutions to the tangle equations are

- (a) $O = (6)$ and $R = (+1)$, $K_2 = b(7, 6)$.
- (b) $O = (6)$ and $R = (-13)$, $K_2 = b(7, 1)$.
- (c) $O = (6, 2, 0)$ and $R = (-1)$, $K_2 = b(7, 6)$.
- (d) $O = (-5, -1)$ and $R = (4)$, $K_2 = b(14, 9)$.
- (e) $O = (-5, -1)$ and $R = (-1)$, $K_2 = b(11, 9)$.
- (f) $O = (-5, -1, -2, 0)$ and $R = (+1)$, $K_2 = b(11, 9)$.

Solutions (a)–(c) have to be discarded, since they correspond to torus knots, while Solution (d) is to be discarded because it corresponds to a link of parental genotype. Solutions (e) and (f) are the only acceptable ones, since they correspond to twist knots of recombinant genotype.

When $O = X + A$, X and A being rational nonintegral, the solutions to the tangle equations are as follows.

- (1) $X = (-4, 0)$, $A = (-2, 0)$ and $R = (+3)$, $K = b(18, 13)$.
- (2) $X = (-4, 0)$, $A = (-2, 0)$ and $R = (-1)$, $K = b(14, 9)$.
- (3) $X = (-3, 0)$, $A = (-3, 0)$ and $R = (-1)$, $K = b(15, 11)$.
- (4) $X = (-3, 0)$, $A = (-3, 0)$ and $R = (+3)$, $K = b(21, 13)$.

Here, (1) and (2) are to be discarded as they correspond to a link, while (3) and (4) constitute the only acceptable solutions as they correspond to knots of recombinant genotype.

Remarks

The tangle is modeled assuming that, for a given enzyme, the tangles P and R are constants, independent of the topology of the substrate. We performed tangle analysis of two recombination events mediated by recombinases. It has been shown that if $P = (0)$, R is integral and O is rational or the sum of two rational tangles, and K_2 is a 4-plat, then there are only three solutions that explain the observed products in both the experiments (unknotted and catenated).

5. CONCLUSIONS

The tangle model treats the circular DNA substrate and products as knots or links. The site-specific recombinase and its accessory proteins are considered as a ball that intersects the DNA knot or link in two strands. The interior of the ball is divided into two regions. One of them is restricted to strand exchange and corresponds to a parental tangle P . This tangle can be chosen to be $P = (0)$. P represents the only region in the synaptic complex that changes upon recombination. The region outside P but inside the ball, called O_b , traps all the conformation that, together with the change from P to R , determines the topology of the recombination products. Finally, the region outside the ball, O_f , detects the variation between substrates with different topology. The tangle model assumes that the synaptic complex can be expressed as $N(O + P) = K_0$ where $O = O_f + O_b$ is called the outside tangle. Recombination is modeled by a tangle surgery that replaces P by the recombinant tangle R , thus leading to a product equation $N(O + R) = K_1$. The assumption of constant mechanism implies that P and R are constants uniquely determined by the enzyme. In the cases when there are both topological selectivity and specificity (e.g., Tn3 resolvase, Gin, Xer, Topoisomerase III and IV), the tangle O is also determined uniquely by both the enzyme and the topology of the substrate. If there is no topological selectivity (e.g., λ -Int, mutant Gin, and FLP) then, for a fixed substrate, P and R are constants but O can vary. Furthermore, processive recombination is modeled by tangle addition. A recombination event that consists of n -rounds of processive recombination is translated to a system of $(n + 1)$ equations with unknowns $O^{(i)}$, P , and R . The tangle $O^{(i)}$ is allowed to change from one equation to another if and only if there is no topological selectivity. This introduces more unknowns to the system, and the analysis becomes much more difficult.

APPENDIX

- (1) Both $N(a/b)$ and $D(a/b)$ are 4-plats. The knot/link $N(a/b)$ is the 4-plat $S(a, -b)$. The knot/link $D(a/b)$ is the 4-plat $S(b, a)$.
- (2) The tangle corresponding to $a_1 \setminus b_1$ is the same as the tangle corresponding to $a_2 \setminus b_2$ if and only if $a_1 \setminus b_1 = a_2 \setminus b_2$.
- (3) $a_1 \setminus b_1 + a_2 \setminus b_2 \cong$ a rational tangle unless either $b_1 = \pm 1$ or $b_2 = \pm 1$.
- (4) $a/b + t = (a + bt)/t$.
- (5) $N(A + C) = N(C + A)$ where A and C are arbitrary tangles (a lemma).
- (6) $N(A + C) = 4\text{-plat}$ implies at least one of A and C is rational or locally knotted.
- (7) $D(A + C) = D(A) \cong D(C)$.
- (8) $N(AN(c_1, \dots, c_n) + B) = N(A + BN(c_n, \dots, c_1))$ where n is odd (a lemma).

REFERENCES

1. C. Ernst and D.W. Summers, A calculus for rational tangles: Applications to DNA recombination, *Math. Proc. Camb. Phil. Soc.* **108**, 489–515, (1990).
2. I. Darcy, Biological distances on DNA knots and kinks: Applications to Xer Recombination, *J. Knot Theory Ramifications* **10**, 269–294, (2001).
3. J.H. Conway, An enumeration of knots and links and some of their related properties, In *Computational Problems in Abstract Algebra*, Proc. Conf. Oxford, 1967, pp. 329–358, Pergamon Press, Oxford, (1970).
4. I. Darcy and D. Summers, A strand passage metric for Topoisomerase action, In *Knots '96*, pp. 267–278, World Scientific, (1997).

5. I. Darcy and D. Sumners, Rational tangle distances on knots and links, *Math. Proc. Cambridge Philos. Soc.* **128**, 497–510, (2000).
6. C. Ernst, Tangle equations, *J. Knot Theory Ramifications* **5**, 145–159, (1996).
7. C. Ernst and D.W. Sumners, Solving tangle equations arising in a DNA recombination model, *Math. Proc. Camb. Phil. Soc.* **126**, 23–36, (1999).
8. J.R. Goldman and L.H. Kaufman, Rational tangles, *Adv. in Appl. Math* **3**, 300–332, (1997).
9. D.W. Sumners, C. Ernst, S.J. Spengler and N.R. Cozzarelli, Analysis of the mechanism of DNA recombination using tangles, *Quarterly Reviews of Biophysics* **28**, 253–313, (1995).