1

DNA and Knot Theory

Koya Shimokawa

Mariel Vazquez

1　Introduction

DNA has a double helical structure discovered by Watson and Crick [33]. The axis of the double helix can be seen as a string in 3-dimentional space, and its "topology" can be studied. Closed circular DNA forms may be knotted or linked. Linear DNA forms are topologically trivial unless the two ends are fixed in space, in which case we can consider local topology. In this article, we consider enzymes which change the topology of DNA. In particular we deal with site-specific recombinases and topoisomerase and show how knot theory is applied to their study.

DNA topology is the study of geometrical and topological properties of circular DNA. Essentially all reactions involving DNA are influenced by its topology. [4] is a good reference book for general DNA topology. Some knot theory books include expository chapters on the applications of low-dimensional topology and knot theory to the study of DNA ([1, 16, 22]). These books focus on the tangle method of Ernst and Sumners and the analysis of site-specific recombination [14].

2　Knots, tangles and Dehn surgery

Here a *knot* means an image of an embedding of the 1-dimensional sphere into a 3-manifold. A disjoint union of knots are called *link*. We mainly deal with knots and links in the 3-dimensional sphere.

A 2-*string tangle*, or simply *tangle*, means a pair of a 3-dimensional ball and two arcs properly embedded in it. *Rational tangles* are the simplest class of tangles. See Figure 1 for examples. There is a one-to-one correspondence between a set of rational tangles and $\mathbb{Q} \cup \{1/0\}$. A rational tangle can be untangled by a finite number of horizontal and vertical twistings.

The *numerator* $N(T)$ of a tangle T is a knot or link obtained by connecting two upper endpoints of T by an arc, and the two lower endpoints of T by another arc as illustrated in Figure 2. From two tangles S and T, a new tangle can be obtained by connecting the two East endpoints of S to the two West endpoints of T as is shown in Figure 3. The obtained tangle is called the *sum* of S and T, and is denoted by $S + T$.

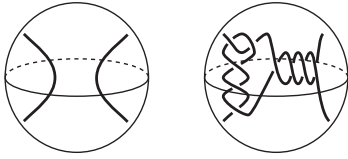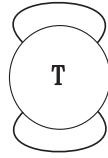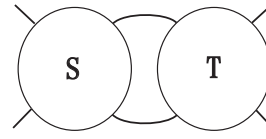*Dehn surgery* is a method of constructing 3-manifolds using knots and links. Let K be a

図1：Rational tangles.

図2：The numerator $N(T)$

図3：The tangle sum $S + T$

knot in a 3-manifold M and let $E(K)$ be the exterior $M-\mathrm{int}N(K)$, where $N(K)$ is the regular neighborhood of K. Let γ be an isotopy class of a simple closed curve on $\partial E(K)$. We attach a solid torus $D^2 \times S^1$ to $E(K)$ so that γ bounds a meridian disk of $D^2 \times S^1$. Let $K(\gamma)$ denote the obtained 3-manifold. We say that $K(\gamma)$ is obtained from M by Dehn surgery along K.

3 DNA knots and catenane and tangle method for site-specific recombination

Circular genomes and naturally occurring plasmids are subject to knotting and catenation. These circular DNA forms are common in prokaryotes (e.g. the genome of the bacterium Escherichia coli is circular). Even though DNA in higher order organisms is commonly linear, it often appears subdivided into loops as a consequence of the tight organization of DNA in the cell nucleus. DNA knots and links can be detected experimentally by using gel electrophoresis or electron microscopy (EM) (reviewed in [4]).

Site-specific recombinases are enzymes able to change the topology of DNA. These enzymes bind two specific DNA sites, introduce one double-stranded break on each site, recombine the open ends, and reseal the ends. Their cellular role is to change the genetic code of an organism by integration or excision of a DNA segment, by moving a DNA segment to a new location, by inverting a DNA segment within a genome, or by changing its topological form.

Changes in topology can be observed experimentally by taking closed circular DNA substrates and by incubating them with the enzyme of choice. The products are analyzed by gel electrophoresis or electron microscopy, and the mechanism of recombination is analyzed using tangles. The tangle method was introduced by Ernst and Sumners [14] and it has been used to characterize topologically the action of several site-specific recombinases (e.g.[6, 9, 14, 26, 30, 31]).

In the tangle method, the enzymes and the bound DNA can be modeled as a tangle E. The tangle O_f is the exterior of E and contains the DNA not bound by the enzyme. The following biologically reasonable assumptions are made.

(1) The enzyme mechanism is constant and independent of topology of the substrate. The action occurs in a 2-string tangle E which is a tangle sum of two tangles O_b and P. The tangles O_b and O_f remain unchanged during the recombination reaction.

(2) Recombination is modeled by tangle surgery, where the tangle P is changed into the tangle R. P can be assumed to be rational. Sometimes R can be proven to be rational (e.g. [14, 30]). In other cases R can be assumed to be rational on biological grounds.

The tangle $O = O_f + O_b$ is called the outside tangle, and it remains unchanged during the recombination reaction.

The length of the part of DNA in the tangle P is short (32bp for Xer system [28]), so the tangles P and R cannot be very complicated (see assumption 2).

Suppose the substrate has the knot or link type K_1 and the product K_2. Then we have the following tangle equation (1).

$$\begin{cases} N(O_f + O_b + P) = N(O + P) = K_0 \\ N(O_f + O_b + R) = N(O + R) = K_1 \end{cases} \tag{1}$$
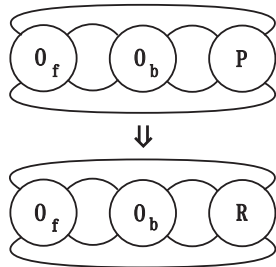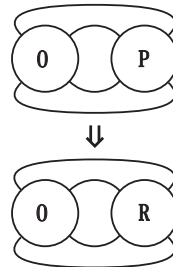

図 4 : Tangle model of recombination

図 5 : Tangle model of recombination II

4 Tangle surgery and Dehn surgery

When P and R are rational, a site-specific recombination reaction in Figure ?? can be modeled as a rational tangle surgery. The double branched covering of a rational tangle is a solid torus. Hence a rational tangle surgery corresponds to a Dehn surgery. In many cases, the substrates are trivial knots and the products are 2-bridge knots and links. The double branched covering of the 3-sphere branched along the trivial knot is the 3-sphere and the double branched covering of the 3-sphere branched along a 2-bridge knot or link is a lens space. Hence these rational tangle surgeries correspond to Dehn surgeries along a knot in the 3-sphere yielding lens spaces. By the construction, the knot is strongly invertible. Hence results such as [8, 18] have direct applications to the study of site-specific recombination.

5 Applications

In [14], Ernst and Sumners characterized the action of Tn3 resolvase. The action of Tn3 resolvase on DNA is processive, i.e. the enzyme may mediate more than one round of recombination before dissociating from its DNA substrate.

The tangle method has an additional assumption in this case: (3) In procesive recombination, each round of recombination adds a copy of the recombinant tangle R to the bound DNA, and recombination acts by tangle addition.

The tangle equation for n-th round is the following.

4　　　　　　　　　　　　　　　論　　　　説

$$N(O_f + O_b + nR) = N(O + nR) = K_n$$

Here nR denotes a tangle obtained by the tangle sum of n copies of the tangle R. Experimental data [34, 35] shows K_1 is a Hopf link, K_2 a figure eight knot, and K_3 Whitehead link. Ernst and Sumners solved this tangle eqation and predicted K_4 will be 6_2 knot.
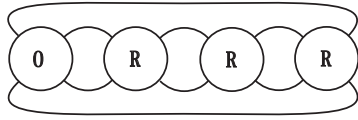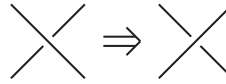

図 6 : $N(O + 3R)$

図 7 : Crossing change

In [31], the action of Xer system which yield 4-cat ($(2, 4)$-torus link) from a trivial knot is characterized. There they show that the tangle O in the tangle equation is rational using a result in [18]. In [17] unlinking of DNA catenane by Xer system is reported. The corresponding tangle analysis can be found in [26].

There are softwares which solve tangle equations under suitable assumptions [11, 25].

6　Type-2 topoisomerases

Type-2 topoisomerases are enzymes essential to every living organism. The play key roles in important cellular processes such as DNA replication. The enzyme acts by a 2-gate model where it binds a DNA segment, introduces a double-stranded break, transports a second segment through the break, and reseals the transient break. The net effect of this reaction is a strand-passage which may change the topology of a circular DNA molecule as illustrated in the figure. 7.

In [23], it is shown type II topoisomerase simplify DNA topology below equilibrium values. These enzymes have been observed to unknot large numbers of complicated DNA knots (i.e. knots wit high crossing number) very efficiently [3]. These experimental results cannot be accounted for by a random mechanism of action and suggest that on average type-2 topoisomerases find the shortest topological pathway between two knots. We can conjecture that given a knots of type K_1, type-2 enzymes unknot K_1 in $u(K_1)$ steps where $u(K_1)$ is the unknotting number of K_1. More in general, type-2 enzymes turn one knot K_1 into a knot K_2 in N steps, where N is given by the shortest strand-passage distance between K_1 and K_2 as defined in [10]. Undertanding the mechanism of action of topology simplification by type-2 topoisomerases remains one of the most relevant open problems in DNA topology. Recent advances in this problem use a variety of mathematical [7, 11, 19, 36], computational [5, 12, 15, 20, 21, 32] and experimental methods [13, 29, 23, 27].

文　　　　献

[1]　C.C. Adams, The Knot Book: An Elemen-

tary Introduction to the Mathematical Theory of Knots. New York: W. H. Freeman, 1994.

[2]　J. Arsuaga, Y. Diao and M. Vazquez, Mathe-

matical methods in DNA topology: Application to chromosome organization and site-specific recombination, To appear in IMA proceeding.

[3] J. Arsuaga, M. Vazquez, S. Trigueros, D.W. Sumners, and J. Roca, Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids, Proc. Natl. Acad. Sci. U.S.A. **99** (2002) pp. 5373-5377.

[4] A.D. Bates and A. Maxwell, DNA Topology (2nd ed.), Oxford university press, 2005.

[5] Y. Burnier, C. Weber, A. Flammini, A. Stasiak, Local selection rules that can determine specific pathways of DNA unknotting by type II DNA topoisomerases. Nucleic Acids Res. **35(15)** (2007), 5223–31.

[6] D. Buck, C.V. Marcotte, Tangle solutions for a family of DNA-rearranging proteins, Math. Proc. Camb. Phil. Soc. **139** (2005), 59–80.

[7] G.R. Buck, and E.L. Zechiedrich, DNA disentangling by type-2 topoisomerases, J. Mol. Biol. **340** (2004) 933-939.

[8] M. Culler, C.McA. Gordon, J. Luecke and P.B. Shalen, Dehn surgery on knots, Ann. of Math. **125** (1987), 237–300.

[9] I.K. Darcy, Biological distances on DNA knots and links: Applications to Xer recombination, J. Knot Theory Ramifications **10** (2001), 269–294.

[10] I.K. Darcy and D.W. Sumners, Rational tangle distances on knots and links. Math. Proc. Camb. Phil. Soc., **128** (2000) 497-510.

[11] I.K. Darcy and R.G. Scharein, TopoICE-R: 3D visualization modeling the topology of DNA recombination, Bioinformatics **22** (2006), 1790-1791.

[12] I.K. Darcy, R.G. Scharein and A. Stasiak, 3D visualization software to analyze topological outcomes of topoisomerase reactions. Nucleic Acids Res. **36(11)** (2008), 3515-21.

[13] K.C. Dong, J.M Berger, Structural basis for gate-DNA recognition and bending by type IIA topoisomerases. Nature (2007) **450**(7173) 1201-1205.

[14] C. Ernst and D.W. Sumners, A calculus for rational tangles: Applications to DNA recombination, Math. Proc. Cambridge Philos. Soc. **108** (1990), 489–515.

[15] A. Flammini, A. Maritan, and A. Stasiak, Simulations of action of DNA topoisomerases to investigate boundaries and shapes of spaces of knots, Biophys J. **87** (2004) 2968-2975.

[16] E. Flapan, When Topology Meets Chemistry: A Topological Look at Molecular Chirality, Cambridge University Press, 2000.

[17] I. Grainge, M. Bregu, M. Vazquez, V. Sivanathan, S. CY Ip and D.J. Sherratt, Unlinking chromosome catenanes in vivo by site-specific recombination, EMBO J. **26** (2007) 4228–4238.

[18] M. Hirasawa and K. Shimokawa, Dehn surgery on strongly invertible knots which yield lens spaces, Proc. Amer. Math. Soc. **128** (2000), 3445-3451.

[19] X. Hua, D. Nguyen, B. Raghavan, J. Arsuaga, M. Vazquez, Random state transitions of knots: a first step towards modeling DNA unknotting by type II topoisomerases, Topology Appl. (2006).

[20] Z. Liu, J.K. Mann, E.L. Zechiedrich, and H.S. Chan, Topological information embodied in local juxtaposition geometry provides a statistical mechanical basis for unknotting by type-2 DNA topoisomerases, J. Mol. Biol. **361** (2006) 268-285.

[21] Z. Liu and H.S. Chan, Efficient chain moves for Monte Carlo simulations of a wormlike DNA model: excluded volume, supercoils, site juxtapositions, knots, and comparisons with randomflight and lattice models. J. Chem. Phys. (2008) **128(14)**:145104.

[22] K. Murasugi, Knot theory and its applications, Birkhäuser, 1996.

[23] V.V. Rybenkov, C. Ullsperger, A.V. Vologodskii, and N.R. Cozzarelli, Simplification of DNA topology below equilibrium values by type II topoisomerases, Science **277** (1997) 690-693.

[24] Y. Saka and M. Vazquez, TangleSolve: topological analysis of site-specific recombination, Bioinformatics **18** (2002) 1011-1012.

[25] K. Shimokawa, K. Ishihara, and M. Vazquez, Tangle analysis of DNA unlinking by the Xer/FtsK system. Bussei Kenkyu (2009).

[26] T. Stuchinskaya, L.A. Mitchenall, A.J. Schoeffler, K.D. Corbett, J.M. Berger, A.D. Bates, A. Maxwell, How do type II topoisomerases use ATP hydrolysis to simplify DNA topology beyond equilibrium? Investigating the relaxation reaction of nonsupercoiling type II topoisomerases. J Mol Biol. (2009) **385** 1397-1408.

[27] D.K. Summers and D.J. Sherratt, Multimerization of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerization and stability, Cell **36** (1984) 1097–1103.

[28] Trigueros, S., Salceda, J., Bermudez, I., Fernandez, X. and Roca, J., Asymmetric removal of supercoils suggests how topoisomerase II simplifies DNA topology, J Mol Biol **335** (2004) 723-731.

[29] M. Vazquez and D.W. Sumners, Tangle analysis of Gin site-specific recombination, Math. Proc. Camb. Phil. Soc. **136** (2004) 565–582.

[30] M. Vazquez, S.D. Colloms and D.W. Sumners, Tangle analysis of Xer recombination reveals only three solutions, all consistent with a single three dimensional topology pathway, J. Mol. Biol. **346** (2005) 493-504.

[31] A.V. Vologodskii, W. Zhang, V.V. Rybenkov, A.A. Podtelezhnikov, D. Subramanian, J.D. Griffith, and N.R. Cozzarelli, Mechanism of topol-

6　　　　　　　　　　　　　　　　　　　論　　　説

ogy simplification by type II DNA topoisomerases, Proc. Natl. Acad. Sci. U.S.A.**98** (2001) 3045-3049.

[32]　J.D. Watson and F.H.C Crick, A structure for deoxyribose nucleic acid. Nature **171** (1953) 737-738.

[33]　S.A. Wasserman and N.R. Cozzarelli, Determinaiton of the stereostructure of the product of Tn3 resolvase by a general method, Proc. Nat. Acad. Sci. U.S.A. **82** (1985) 1079-1083.

[34]　S.A. Wasserman, J.M. Dungan, and N.R. Cozzarelli, Discovery of a predicted DNA knot substantiates a model for site-specific recombination, Science **229** (1985) 171-174.

[35]　J. Yan, M.O. Magnasco, and J.F. Marko, A kinetic proofreading mechanism for disentanglement of DNA by topoisomerases, Nature **401** (1999) 932-935.

(Koya Shimokawa・Saitama University) (Mariel Vazquez・San Francisco State University)