

## A GENERALIZED 4-STRING SOLUTION TANGLE OF DNA-PROTEIN COMPLEXES

SOOJEONG KIM

DEPARTMENT OF APPLIED MATHEMATICAL AND COMPUTATIONAL SCIENCES, UNIVERSITY OF IOWA,  
USA

*E-mail address:* kimcrystal@gmail.com

**ABSTRACT.** An  $n$ -string tangle is a three dimensional ball with  $n$  strings properly embedded in it. A tangle model of a DNA-protein complex is first introduced by C. Ernst and D. Summers in 1980's. They assumed the protein bound DNA as strings and the protein as a three dimensional ball. By using a tangle analysis, one can predict the topology of DNA within the complex. S.Kim and I. Darcy developed the biologically reasonable 4-string tangle equations and decided a solution tangle, called  $R$ -standard tangle. The author discussed more about the simple solution tangles of the equations and found a generalized  $R$ -standard tangle solution.

### 1. INTRODUCTION

An  $n$ -string tangle  $T$  is a pair  $(B, t)$  where  $B$  is a three dimensional ball and  $t = \{t_1, \dots, t_n\}$  is a set of  $n$  arcs embedded in  $B$  such that two end points of  $t_i$  are on the boundary of  $B$  for all  $i = 1, \dots, n$ . (See Figure 6 for examples of tangles.) The concept of a tangle is introduced by John Conway [3] in the 1960's while he tabulate knots. Especially the case  $n = 2$  is well studied by many mathematicians [7, 8, 9, 19, 12, 16, 17, 18]. In the 1980's, this 2-string tangle theory is used to model a DNA-Tn3 protein complex and a DNA-Phase $\lambda$  protein complex by C. Ernst and D. Summers, motivated by Nick Cozzarelli's experiments [6, 17, 21, 22]. In this model, the protein is represented by a three dimensional ball  $B^3$  and the DNA segments within the protein are represented by arcs (See Figure 1). Ernst and Summers predicted the topology of DNA-segments within the Tn3/Phase $\lambda$  proteins [7, 8].

A tangle analysis with  $n = 3$  or more is much more complicated than 2-string tangle analysis. In 2007, I. Darcy, J. Luecke and M. Vazquez used a 3-string tangle analysis to figure out the topology of DNA trapped by Mu-proteins. This work is motivated by S. Pathania, M. Jayaram and R. Harshey's difference topology experiment[14] (See Figure 5). This experiment is also computationally analyzed in [4] by Darcy et al.

---

Received by the editors April 4, 2011; Revised June 10, 2011; Accepted in revised form June 24, 2011.

2000 *Mathematics Subject Classification.* 57M25; 92C40; 92E10.

*Key words and phrases.* DNA topology,  $n$ -string tangle, Conway notation, R-standard tangle.

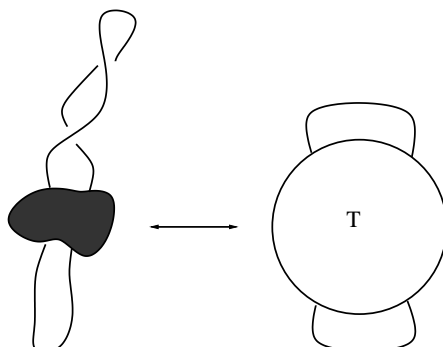


FIGURE 1. DNA topology. Left figure is redrawn from AFM image of a Cre synaptic complex with circular DNA [19]. Right figure is a tangle model corresponding to it.

Recently, in 2009, S. Kim and I. Darcy addressed a possibility that proteins bind to DNA at 4 sites [10, 11] and introduced a 4-string tangle model of DNA-protein complexes in this case. Even though there is no experimental data for this 4-string tangle model, there are many biological processes (especially DNA replication, recombination and transposition) involve multiple proteins with multiple DNA segments. Kim and Darcy determined that a biologically reasonable 4-string tangle model of DNA-protein complexes is an  $R$ -standard tangle (see section 3).

In this paper, I discussed more about the 4-string tangle solution of Kim and Darcy's equations. Also, I discovered a property that a graph which corresponds to an  $R$ -standard tangle is very similar to a 2-string rational tangle. By using a Conway notation of 2-string rational tangle, I found a generalized  $R$ -standard solution tangle.

## 2. DNA TOPOLOGY AND DNA-PROTEIN COMPLEX

As well as Human, every organism contains its genetic information in DNA. The information is encoded with 4 bases: A(adenine), G(guanine), C(cytosine) and T(thymine). In 1953, James Watson and Francis Crick discovered the double helical structure of DNA [23]. The meaning of 'double' is that DNA consists of two strands connected to each other by hydrogen bonds. One strand is made up of repeated bases (A, G, C and T) and the other strand also consists of those 4 bases in the way that  $A=T$  (A pairs with T by two hydrogen bonds) and  $G\equiv C$  (G pairs with C by three hydrogen bonds). 'Helical' is from the fact that the two strands are coiled around the helical axis once every 10.5 base pairs in right handed fashion. Simply speaking, the shape of DNA is a twisted ladder (see Figure 2 (a)).

Because of this double helical structure, when twists are added or subtracted to a long DNA molecule (this is a very common situation due to biological process such as replication, translation, recombination, etc), DNA tends to coil upon itself. Such coiling of DNA is the process

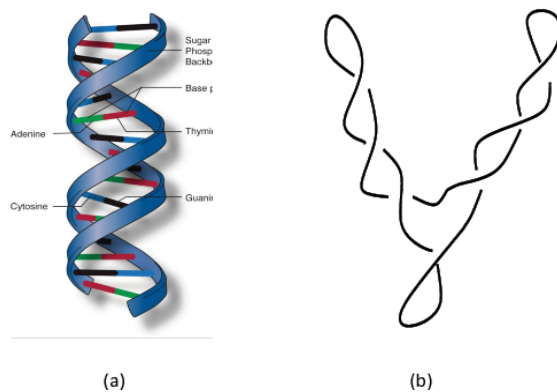


FIGURE 2. (a) The structure of double helical DNA. Courtesy: National Human Genome Research Institute (b) A typical conformation of supercoiled DNA. A double stranded DNA is assumed as a single string. Redrawn from [20]

of *supercoiling*. It is usually negatively supercoiled in most organisms. See Figure 2 (b) for a typical conformation of a supercoiled DNA.

DNA is in either linear or circular form. The typical form of bacterial DNA and cytoplasmic DNA in animal is circular [20]. Thus studying properties of a circular DNA is very interesting and important to understand biological activities of organism. To measure the supercoiling of a circular DNA, one can use twist and writhe (see [1, 2, 20] for more detail). This is one way to understand *DNA topology*. In this paper, I focus on a supercoiled circular DNA.

To be replicated, the helixes of DNA need to be unwound and the two strands should be separated from each other. And then a synthesis of new DNA will be held. From the beginning to the end of the DNA replication, lots of proteins(enzymes) act on DNA. Besides this activity, to perform a variety of biological process, many proteins(enzymes) bind to DNA to help. There are two kinds of DNA-binding proteins; one binds to non-specific DNA segments, and the other binds to specific DNA segments. I will focus on the latter. *Cre* is a well known example of a specific DNA-binding protein which binds at two specific DNA sites, *loxP*. It helps a DNA recombination and the process is the following: *Cre* binds to two *loxP* sites, cut those two sites and then rejoin them. This process can be modeled by Ernst and Sumners' 2-string tangle model as in Figure 3.

*Mu* proteins are bacteriophage which help transposition (a movement of DNA segment from one place to another within a genome) efficiently. These proteins bind to DNA at 3 specific

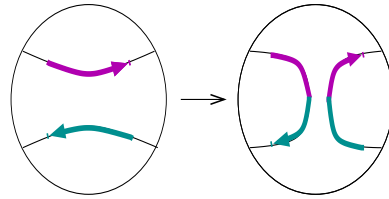


FIGURE 3. A 2-string tangle model of Cre reaction. Cre is modeled by a ball and DNA is modeled by a string. Courtesy: I. Darcy

sites, L(attL), R(attR) and E(enhancer). When Mu proteins bind to all those three sites lying on a circular DNA, we call the complex Mu-transpososome (See Figure 4 (a)). The structure of Mu-transpososome is very important to understand the transposition process. In 2002, Pathania, Jayaram and Harshey studied the topology of DNA within the complex [14]. They predicted that it has the five crossing conformation (Figure 4 (b)). This is under the assumption that the DNA bound by Mu is supercoiled and branched. They used a difference topology technique(see Figure 5) for the experiments and it's well summarized in [14, 5, 10, 11].

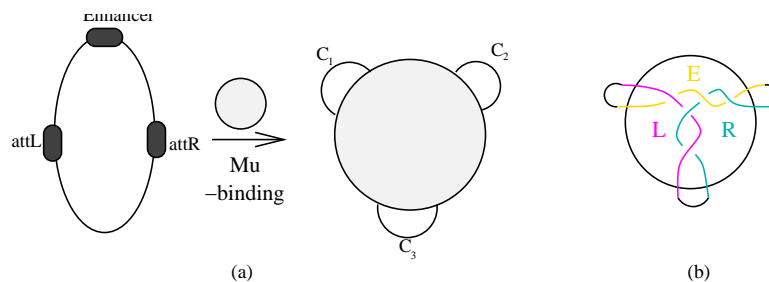


FIGURE 4. (a) Construction of Mu-transpososome. (b) 5-crossing conformation of DNA within the Mu. Figure from [11]

### 3. TANGLE ANALYSIS

An  $n$ -string tangle is a pair  $(B, t)$ , where  $B$  is a 3 dimensional ball and  $t$  is a set of arcs embedded in  $B$ . The two end points of each arc lie on the boundary of  $B$ . Two tangles are *equivalent* if they are ambient isotopic keeping the boundary of  $B$  fixed. Two tangles  $T_1, T_2$  are *freely isotopic* if there is an isotopy of the 3-ball taking  $T_1$  to  $T_2$ , which is not necessarily fixed on its boundary. If a tangle  $T = (B, t)$  is freely isotopic to a tangle with no crossing, then we say  $T$  is a *rational tangle*.  $T$  is *locally knotted* if there is a 2-sphere  $S$  in  $B$  that intersect one of the two strings transversely in two points and the string in  $S$  is knotted with end points on  $S$ . In the case that  $T$  is neither rational nor locally knotted,  $T$  is called a *prime tangle*. Examples of 3-string tangles are shown in Figure 6.

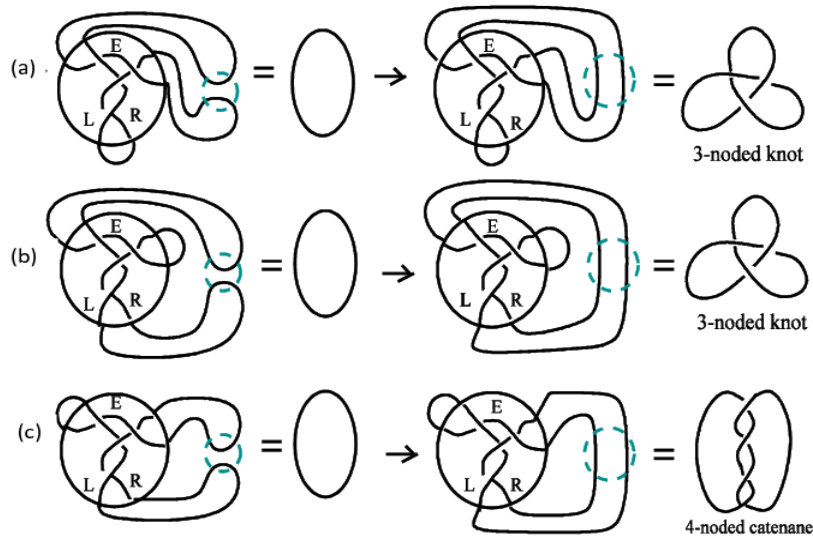


FIGURE 5. Tangle model of the difference topology experiments of Mu-transpososome. In this figure, the dotted circles represent Cre. (a) Cre binds to the two loxP sites lying on the two outside loops of DNA; one is connecting E and R, the other is connecting E and L. Before the Cre reaction, DNA conformation is unknot, whereas after the Cre reaction, it is 3-noded knot (trefoil). (b) Cre binds to the two loxP sites lying on the two outside loops of DNA; one is connecting E and R, the other is connecting L and R. Before the Cre reaction, DNA conformation is unknot, whereas after the Cre reaction, it is 3-noded knot (trefoil). (c) Cre binds to the two loxP sites lying on the two outside loops of DNA; one is connecting E and L, the other is connecting R and L. Before the Cre reaction, DNA conformation is unknot, whereas after the Cre reaction, it is 4-noded link. In this figure, I summarized the main experiments used to determine DNA topology within the Mu. In [14], there are more experiments up to the orientation of loxP sites. Thus, they did experiments with linear DNA instead of the circular DNA (which include E, L and R) to confirm their results.

Especially, 2-string tangle theory is well studied initiated by J. Conway[3]. A 2-string rational tangle is defined as a tangle freely isotopic to a 2-string tangle with no crossing, the zero tangle (see Figure 7(a)). Inversely, a rational tangle can be constructed by adding alternating horizontal and vertical half twists to the zero tangle. From this fact, Conway developed a *Conway notation* of a 2-string rational tangle as  $(x_1, \dots, x_n)$  where  $x_i$ 's are integers: Start with the zero tangle and add horizontal  $x_1$  half twists by rotating NE and SE boundary points of

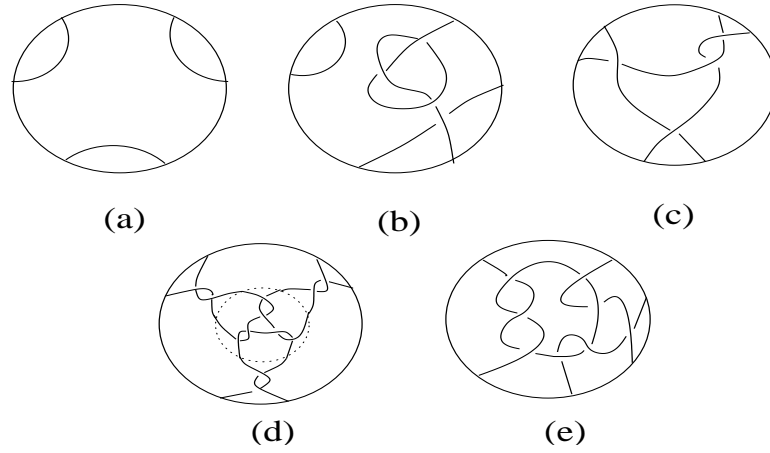


FIGURE 6. Examples of 3-string tangles

(a) Zero tangle; (b) Locally knotted tangle; (c) and (d) Rational tangles; (e) Prime tangle.

strings, add  $x_2$  half vertical twists by rotating SW and SE boundary points of strings and etc. The last integer  $x_n$  must be a number of horizontal half twists and hence  $n$  should be an odd number. For the sign, left handed horizontal twists and right handed vertical twists are positive. See Figure 7(b) for an example.

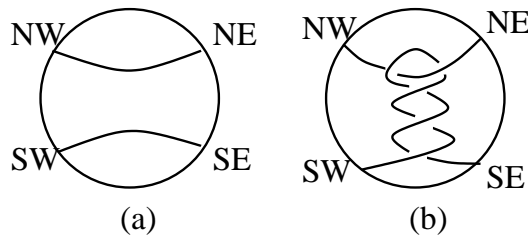


FIGURE 7. Examples of 2-string tangles

(a) The zero tangle; (b) The 2-string tangle,  $(-2,-3,0)$  In this figure, NW, NE, SW and SE are symbolic of the four endpoints of the strings.

By using the Conway notation  $(x_1, \dots, x_n)$ , we can construct the unique extended rational number  $\frac{a}{b} \in \mathbb{Q} \cup \infty$  as the following:

$$\frac{a}{b} = x_n + \frac{1}{x_{n-1} + \frac{1}{x_{n-2} + \frac{1}{\ddots + \frac{1}{x_2 + \frac{1}{x_1}}}}}$$

Conway proved that two equivalent 2-string rational tangles correspond to the same extended rational number. The inverse is also true.

A 3-string tangle analysis is much more complicated than 2-string tangle theory. Since the contour of Mu-transpososome is a ball of proteins with three outside DNA loops, it can be modeled by a 3-string tangle as in Figure 4 (a). Darcy, Luecke and Vazquez analyzed the result of difference topology experiments of Mu-transpososome by 3-string tangle theory. They concluded that the 3-branched 5-crossing conformation is the only biologically reasonable model of Mu-transpososome. This is without the 3-branched and supercoiled DNA assumption:

**Proposition 3.1.** [5] *Let  $T$  be a 3-string tangle which satisfies the system of tangle equations in Figure 8 (a). If  $T$  can be freely isotopic to a projection with less than 8 crossings, then  $T$  is the tangle in the Figure 8 (b).*

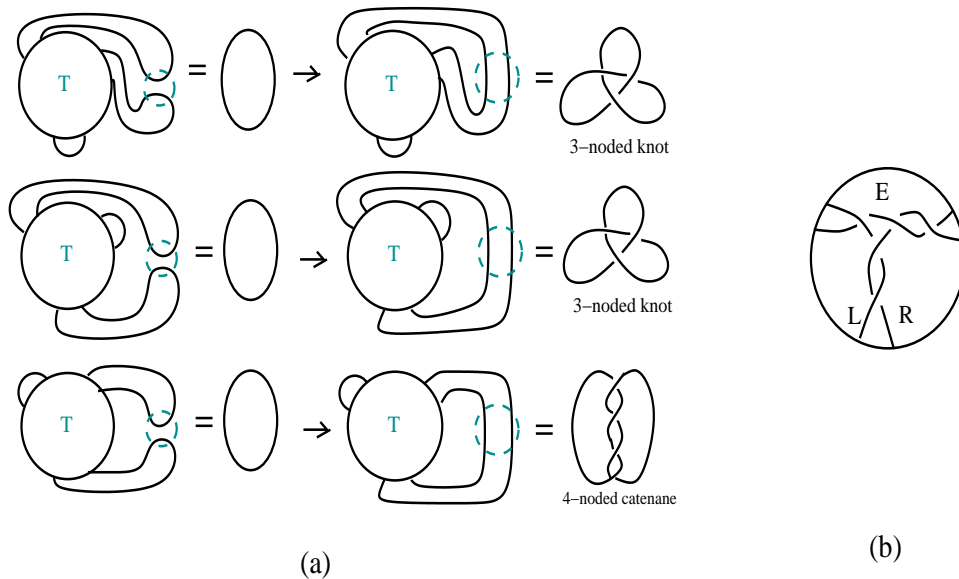


FIGURE 8. Tangle model of Mu transpososome. Figure from [11]

Note that Darcy et al also used the Pathania et al’s extra experimental result with linear DNA to exclude eight crossing solutions.

Kim and Darcy proposed the potential that proteins bind to DNA at four sites. The experimental data is not discovered yet but it’s very practical proposal. In this case, a DNA-protein complex can be modeled by 4-string tangle and the most biologically relevant model is in Figure 9.

Motivated from this model, they defined a *standard tangle* and a *standard graph*:

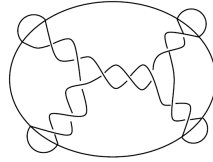


FIGURE 9. The most biologically relevant 4-string tangle model of a DNA-protein complex

**Definition 3.1.** [10] A tangle of the form shown in Figure 10 (a) will be called standard, where  $n_i$  is the number of left-handed half twists. Note that a 4-string standard tangle  $T$  can be represented by a weighted graph  $G_s$ , where  $G_s$  is as in Figure 10(b). Call this graph  $G_s$  a standard graph.

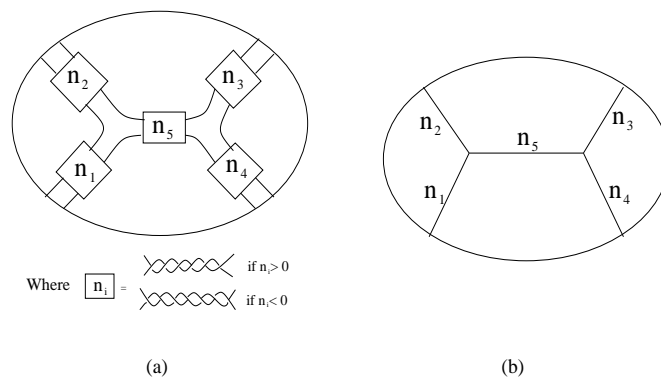


FIGURE 10. A standard tangle and a standard graph

(a) Standard tangle; (b) A weighted graph  $G_s$  representing a 4-string standard tangle. Figure from [10]

From the possibility that a pair of supercoiled DNA branches is twisted, they defined an *R-standard tangle*:

**Definition 3.2.** [10] A weighted graph  $G_R$  is an *R-standard graph* if it is isotopic to a standard graph  $G_s$  allowing the boundary of  $G_s$  to move (Figure 11). A tangle  $T$  is *R-standard* if it corresponds to an *R-standard graph*  $G_R$ .

Kim and Darcy set up a system of tangle equations based on difference topology, see Figure 12. Define a *solution tangle* as a 4-string tangle satisfies all tangle equations in Figure 12 where the products are  $(2, p_i)$  links. From both biological and mathematical reason, they restricted



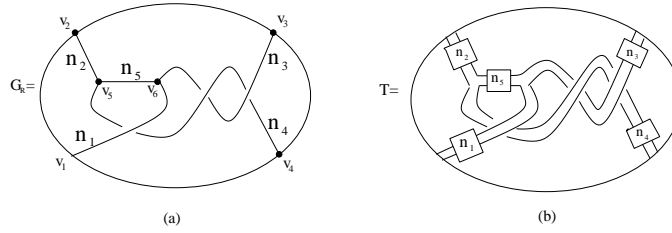


FIGURE 11. A weighted  $R$ -standard graph  $G_R$  and the corresponding tangle. Figure from [11]

the condition of the product topology of DNA. See [11] for more details. And they concluded that a biologically reasonable solution tangle is an  $R$ -standard tangle:

**Theorem 3.2.** [11] *Suppose  $T$  is 4-string tangle which has less than 8 crossings up to free isotopy. If  $T$  is a solution tangle, then  $T$  is  $R$ -standard.*

#### 4. BRANCHED SUPERCOILED DNA SOLUTIONS

In this section, we discuss more about the branched supercoiled DNA solutions. We start with the following definition:

**Definition 4.1.** *Let  $G_R$  be a graph which corresponds to an  $R$ -standard tangle. There are two vertices in the interior of the ball and 4 vertices on the boundary of the ball. Let  $v_1 = SW$ ,  $v_2 = NW$ ,  $v_3 = NE$ ,  $v_4 = SE$  be the vertices on the boundary of the ball, and  $v_5$  and  $v_6$  be the vertices in the interior of the ball.  $G_R$  is  $(2, j)$ -branched if  $v_5$  connects  $v_2 = NW$  and  $v_j$  for some  $1 \leq j \leq 4$ ,  $j \neq 2$ .*

The vertex  $v_5$  can only be connected with  $(v_1, v_2)$  or  $(v_2, v_3)$  or  $(v_2, v_4)$ ; therefore, there are only 3 different  $(i, j)$  branching as in Figure 13. For the case  $n_5 = 0$ ,  $G_R$  is  $(i, j)$ -branched for all  $(i, j)$  and the inverse is also true. Since each edge of  $G_R$  represents a branch of a supercoiled DNA, an  $(i, j)$ -branched graph and a  $(k, l)$ -branched graph represent different topologies of a DNA molecule when  $n_5 \neq 0$  and  $(i, j) \neq (k, l)$ .

We can think about the simplest example of each branching as in Figure 14 which are the most biologically relevant 4-string tangle models of DNA-protein complexes. Assume the tangle in Figure 14 (a) is a solution of equation in Figure 12. Then we have the equations in Table 1 (a) I. In this equations, the  $p_1, \dots, p_6$  values must be determined experimentally, and we solved the equations for  $n_1, \dots, n_6$  as in Table 1 (a) II. For the tangles in Figure 14 (b) and (c), similar calculations are done as in Table 1 (b) and (c). For more information about these simple solution tangles, see [11].

The concept of  $R$ -standard tangle was developed because of the possibility that a pair of supercoiled DNA branches can be twisted. An  $R$ -standard tangle  $T$  corresponds to a weighted graph  $G_R$  as in Figure 11. This implies that twisting a pair of branches of  $T$  corresponds to twisting a pair of edges (excluding the one connecting  $v_5$  and  $v_6$ ) of  $G_R$  along the boundary of

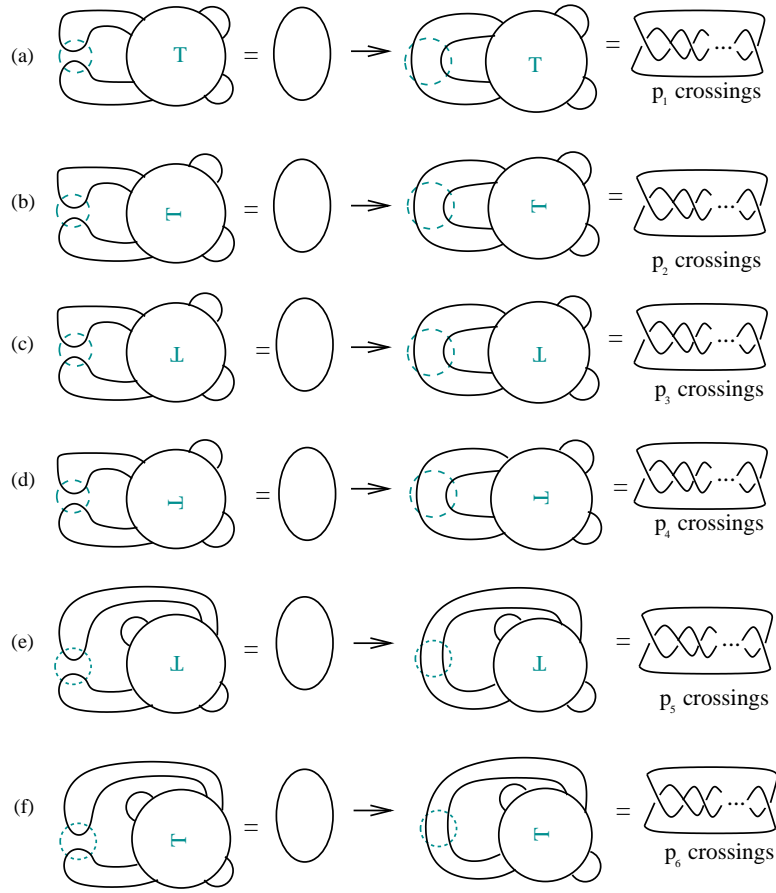


FIGURE 12. A system of tangle equations based on difference topology. In this figure, the tangle  $T$  represents a protein which binds to DNA at 4 sites. The dotted circle represents Cre. Before the Cre recombination, the topology of DNA is unknot, whereas after the reaction, it is  $(2, p_i)$  links. In (b)~(f),  $T$  is rotated. Figure from [11]

the 3-ball. This action of adding twists to  $G_R$  is exactly the same as the action of constructing a rational 2-string tangle (See section 3) except for the fact that a tangle starts from a zero tangle versus a graph starts from a standard graph (see Definition 3.1). Hence we can get an  $R$ -standard graph by adding alternating horizontal and vertical twists to a standard graph, and thus an  $R$ -standard graph can be denoted by Conway notation.

Let  $[a_1, b_1, \dots, a_{n-1}, b_{n-1}, a_n]$  be a Conway notation for  $G_R$ , where  $a_i$  is the number of horizontal right-handed half twists and  $b_i$  is the number of vertical left-handed half twists. The  $a_i$ 's and  $b_i$ 's are integers. In other words, we start with a standard graph, add  $a_1$  horizontal half twists by moving vertices at NE and SE along the boundary of 3-ball, add vertical  $b_1$  half

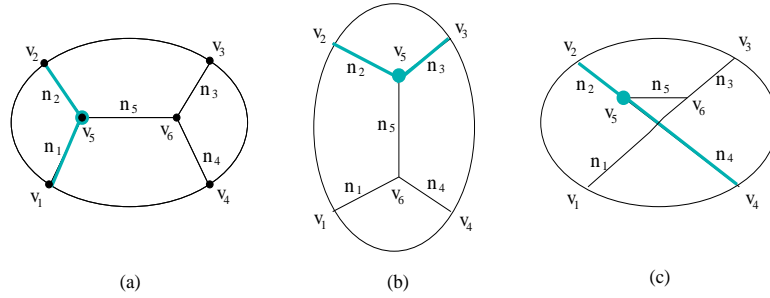


FIGURE 13.  $(i, j)$ -branched weighted graphs for  $R$ -standard tangle  
 (a)(1,2)-branched; (b)(2,3)-branched; (c)(2,4)-branched weighted graph for  $R$ -standard tangle  
 (Note that the edge connecting  $v_4, v_5$  can pass either over or under the edge connecting  $v_1, v_6$ .)

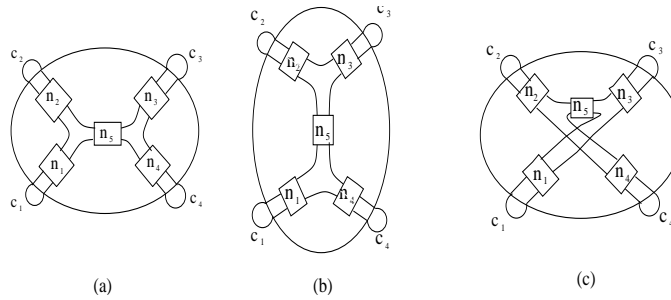


FIGURE 14. Examples of  $R$ -standard tangle model of a branched DNA-protein complex corresponding to a weighted graph  
 (a) (1,2)-branched; (b) (2,3)-branched; (c) (2,4)-branched (Similar to Figure 13 (c), there are two cases at the crossing of two branches depending on which branch is over than another.)

twists by moving vertices at SW and SE along the boundary of 3-ball, and add  $a_2$  horizontal half twists, etc. Similar to rational 2-string tangles, a unique rational number ( $\in \mathbb{Q} \cup \infty$ ) is decided by a continued fraction and two  $R$ -standard graphs are the same (i.e. ambient isotopic) if the two rational numbers from each graph are the same. See Figure 15 for an example.

An interesting property of the continued fraction is that the numbers corresponding to a Conway notation can always be even numbers. Without loss of generality, let  $\frac{A}{B}$  be a reduced ( $A, B$  are relatively prime) rational number which corresponds to  $G_R$  with  $B = A \cdot q + r$  ( $0 < r < A$ ). We can assume  $A < B$  since if  $A > B$ , then  $\frac{A}{B} = q_0 \pm \frac{C}{D}$  where  $C < D$  and  $q_0$  is even which is an entry of Conway notation. Then  $\frac{A}{B} = \frac{1}{\frac{B}{A}} = \frac{1}{\frac{A \cdot q + r}{A}} = \frac{1}{q + \frac{r}{A}}$ . If  $q$  is an even number, it will be an entry of the Conway notation for  $G_R$ . If  $q$  is an odd number, then  $q + 1$

TABLE 1. Equations and solutions related to the three simple branched supercoiled solution tangle

Equations		Solution	
(a) I.	$n_1 + n_2 = p_1$	II.	$n_1 = \frac{p_1 + p_4 - p_6}{2}$
	$n_2 + n_3 + n_5 = p_2$		$n_2 = \frac{p_1 - p_4 + p_6}{2}$
	$n_3 + n_4 = p_3$		$n_3 = \frac{p_2 + p_3 - p_6}{2}$
	$n_1 + n_4 + n_5 = p_4$		$n_4 = \frac{-p_2 + p_3 + p_6}{2}$
	$n_1 + n_3 + n_5 = p_5$		$n_5 = \frac{-p_1 + p_2 - p_3 + p_4}{2}$
	$n_2 + n_4 + n_5 = p_6$		$p_2 + p_4 = p_5 + p_6$
(b) I.	$n_1 + n_2 + n_5 = p_1$	II.	$n_1 = \frac{-p_3 + p_4 + p_5}{2}$
	$n_2 + n_3 = p_2$		$n_2 = \frac{p_3 + p_4 - p_5}{2}$
	$n_3 + n_4 + n_5 = p_3$		$n_3 = \frac{-p_1 + p_2 + p_5}{2}$
	$n_1 + n_4 = p_4$		$n_4 = \frac{p_3 + p_4 - p_5}{2}$
	$n_1 + n_3 + n_5 = p_5$		$n_5 = \frac{p_1 - p_2 + p_3 - p_4}{2}$
	$n_2 + n_4 + n_5 = p_6$		$p_1 + p_3 = p_5 + p_6$
(c) I.	$n_1 + n_2 = p_1$	II.	$n_1 = \frac{p_1 - p_2 + p_5}{2}$
	$n_2 + n_3 + n_5 = p_2$		$n_2 = \frac{p_1 - p_4 + p_6 \pm 2}{2}$
	$n_3 + n_4 = p_3$		$n_3 = \frac{-p_1 + p_2 + p_5}{2}$
	$n_1 + n_4 + n_5 \pm 2 = p_4$		$n_4 = \frac{-p_1 + p_4 + p_6 \pm 2}{2}$
	$n_1 + n_3 = p_5$		$n_5 = \frac{p_2 + p_4 - p_5 - p_6 \pm 2}{2}$
	$n_2 + n_4 = p_6$		$p_1 + p_3 = p_2 + p_4 \pm 2$

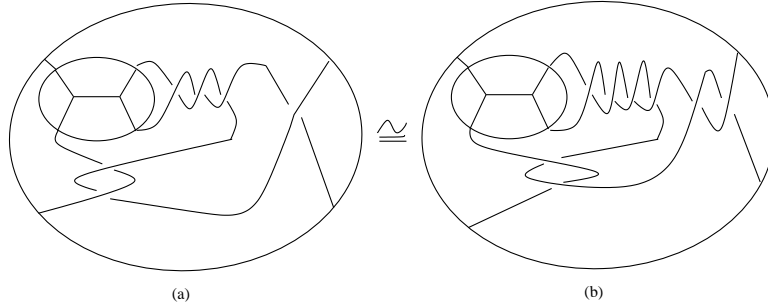


FIGURE 15. Two ambient isotopic  $R$ -standard graphs  
 (a) $[-3,-2,-1] \leftrightarrow -\frac{10}{7} = -1 + \frac{1}{-2+\frac{1}{-3}}$ ; (b) $[-4,2,-2] \leftrightarrow -\frac{10}{7} = -2 + \frac{1}{2+\frac{1}{-4}}$

will be an entry of the Conway notation for  $G_R$  since  $\frac{A}{B} = \frac{1}{\frac{B}{A}} = \frac{1}{\frac{A \cdot (q+1) + (r-A)}{A}} = \frac{1}{(q+1) + \frac{r-A}{A}}$ .  
 In this manner, we can always obtain a Conway notation for an  $R$ -standard graph with all even

entries. See Equation 4.1 for an example:

$$\frac{12}{65} = \frac{1}{\frac{65}{12}} = \frac{1}{\frac{12 \cdot 6 + (5-12)}{12}} = \frac{1}{6 + \frac{-7}{12}} = \frac{1}{6 + \frac{1}{\frac{7 \cdot 2 + (5-7)}{-7}}} = \frac{1}{6 + \frac{1}{-2 + \frac{2}{7}}} = \frac{1}{6 + \frac{1}{-2 + \frac{1}{4 + \frac{1}{-2}}}} \quad (4.1)$$

Kim and Darcy proved that a solution tangle (with less than 8 crossings up to free isotopy) of the equations in Figure 12 is an  $R$ -standard tangle. Let  $T$  be a solution tangle and  $G_R$  be the weighted  $R$ -standard graph which corresponds to  $T$ . Then  $G_R$  has a Conway notation  $[a_1, b_1, \dots, a_{n-1}, b_{n-1}, a_n]$  with all even integral entries (see Figure 16 (a)). Hence  $G_R$  can be obtained by adding alternating full horizontal twists and full vertical twists to a weighted standard graph. Since we add full twists to a weighted graph, the vertices are still  $v_1 = SW$ ,  $v_2 = NW$ ,  $v_3 = NE$ ,  $v_4 = SE$ . This is a very important clue to find a solution tangle.

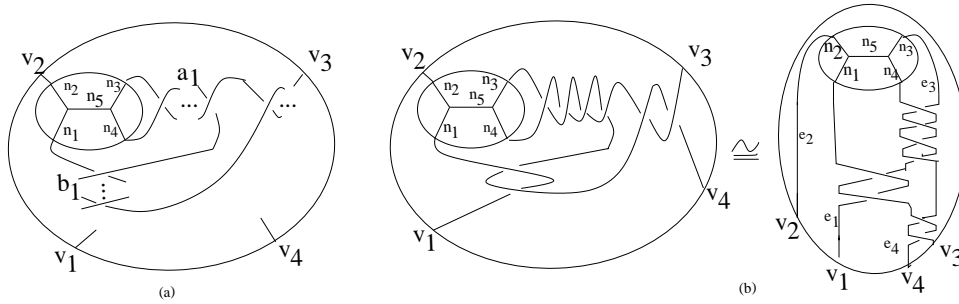


FIGURE 16.  $G_R$   
 (a)  $G_R$ ; (b)  $G_R$  with a Conway notation  $[-4, 2, -2]$  and a different view of it

Let  $p_1, \dots, p_6$  be the numbers from the tangle equations in Figure 12. In other words,  $p_i$  is the number of half twists on the link from Cre recombination on two outside loops of a solution tangle  $T$  which corresponds to  $G_R$  in Figure 16 (a). For the standard tangle,  $p_i$  values are related to the  $n_i$ 's as in Table 1(a). Since we didn't move  $v_2$  at all when adding twists to a standard graph, the equation involving the  $p_1, p_2, p_5, p_6$  are the same as in Table 1 (a), and only  $p_3, p_4$  are changed. This change can be easily seen in the different view of  $G_R$  in Figure 16 (b). Let  $e_i$  be the edge of  $G_R$  with weight  $n_i$ . Then  $e_2$  has no crossings with any other edge.  $e_1$  and  $e_3$  may have crossings with only  $e_3$  because the number of added twists are all even numbers. For example, in the graph of Figure 16 (b), if we walk along the edge  $e_3$  from  $v_3$ , we cross only  $e_4$ . Same is true for  $e_1$ . This implies that we get  $(a_1 + \dots + a_n)$  (sum of the numbers of all horizontal half twists added to a standard tangle) many crossings when we connect  $v_3$  and  $v_4$  by using an arc. Hence the link obtained from Cre recombination on  $c_3$  and  $c_4$  of a solution tangle  $T$  has  $(a_1 + \dots + a_n)$  writhe which can be converted to  $2 \cdot (a_1 + \dots + a_n)$  half twists [2]. I.e.,  $p_3 = n_3 + n_4 + 2 \cdot (a_1 + \dots + a_n)$ . Similarly, the link obtained from Cre recombination on  $c_1$  and  $c_4$  of  $T$  has  $(b_1 + \dots + b_{n-1})$  writhe which can be converted to  $2 \cdot (b_1 + \dots + b_{n-1})$

half twists. Hence  $p_4 = n_1 + n_4 + n_5 + 2 \cdot (b_1 + \cdots + b_{n-1})$ . We can summarize all  $p_i$  values of a solution tangle  $T$  in Equation 4.2.

$$\begin{aligned}
 n_1 + n_2 &= p_1 \\
 n_2 + n_3 + n_5 &= p_2 \\
 n_3 + n_4 + 2 \cdot (a_1 + \cdots + a_n) &= p_3 \\
 n_1 + n_4 + n_5 + 2 \cdot (b_1 + \cdots + b_{n-1}) &= p_4 \\
 n_1 + n_3 + n_5 &= p_5 \\
 n_2 + n_4 + n_5 &= p_6.
 \end{aligned} \tag{4.2}$$

We can solve the system of equations in Equation 4.2, and the solution is the following:

$$\begin{aligned}
 n_1 &= \frac{p_1 + p_4 - p_6}{2} \\
 n_2 &= \frac{p_1 - p_4 + p_6 + 2 \cdot (b_1 + \cdots + b_{n-1})}{2} \\
 n_3 &= \frac{p_2 + p_3 - p_6 - 2 \cdot (a_1 + \cdots + a_n)}{2} \\
 n_4 &= \frac{-p_2 + p_3 + p_6 - 2 \cdot (a_1 + \cdots + a_n)}{2} \\
 n_5 &= \frac{-p_1 + p_2 - p_3 + p_4 + 2 \cdot (a_1 + \cdots + a_n) - 2 \cdot (b_1 + \cdots + b_{n-1})}{2} \\
 p_2 + p_4 - 2 \cdot (b_1 + \cdots + b_{n-1}) &= p_5 + p_6.
 \end{aligned} \tag{4.3}$$

We can do the similar work for equations in Table 1 (b) and (c). Hence we found a generalized branched supercoiled DNA tangle solution (i.e.,  $R$ -standard tangle solution) of equations in Figure 12.

#### ACKNOWLEDGMENTS

The author wishes to acknowledge the help of Dr. I. Darcy in deep discussion about this research.

#### REFERENCES

- [1] A.D. Bates and A. Maxwell, *DNA Topology*, IRL Press Oxford, 1993.
- [2] W. R. Bauer, F. H. C. Crick, and J. H. White, *Supercoiled DNA*, Scientific American **243** (1980), 118–133.
- [3] J.H. Conway, *An enumeration of knots and links and some of their related properties*, Computational Problems in Abstract Algebra (John Leech, ed.) Pergamon Press, Oxford and New York 1969.
- [4] I. K. Darcy, A. Bhutra, J. Chang, N. Druivenga, C. McKinney, R. K. Medikonduri, S. Mills, J. Navarra Madsen, A. Ponnusamy, J. Sweet, and T. Thompson, *Coloring the Mu transpososome*, BMC Bioinformatics, **7:Art. No. 435** (2006).
- [5] I.K. Darcy, J. Luecke, and M. Vazquez, *Tangle analysis of difference topology experiments: Applications to a Mu protein-dna complex*, Algebraic and Geometric Topology, **9** (2009), 2247–2309.

- [6] F.B. Dean, A. Stasiak, T. Koller, and N.R. Cozzarelli, *Duplex DNA knots produced by Escherichia Coli topoisomerase I*, J. Biol. Chem., **260** (1985), 4795–4983.
- [7] C. Ernst and D. W. Sumners, *A calculus for rational tangles: applications to DNA recombination*, Math. Proc. Camb. Phil. Soc., **108** (1990), 489–515.
- [8] C. Ernst and D. W. Sumners, *Solving tangle equations arising in a DNA recombination model*, Math. Proc. Camb. Phil. Soc., **126** (1990), 23–36.
- [9] M. Hirasawa and K. Shimokawa, *Dehn surgeries on strongly invertible knots which yield lens spaces*, Proc. Amer. Math. Soc., **128** (2000) 3445–3451.
- [10] S. Kim and I. Darcy, *Topological Analysis of DNA-protein complexes*, Mathematics of DNA Structure, function and interactions, The IMA volumes in Mathematics and its applications, Springer Science + Business Media, LLC, New York, 2009
- [11] S. Kim and I. Darcy, *A 4-string tangle analysis of DNA-protein complexes based on difference topology*, Preprint (2009)
- [12] P. Kronheimer, T. Mrowka, P. Ozsvath, and Z. Szabo, *Monopoles and lens space surgeries*, Ann. of Math., **165(2)** (2007), 457–546.
- [13] B. Lewin, *Genes*, Oxford University Press, 5th ed., New York, 1994.
- [14] S. Pathania, M. Jayaram, and R. M. Harshey, *Path of DNA within the Mu transposome: transposase interactions bridging two Mu ends and the enhancer trap five DNA supercoils*, Cell, **109(4)** (2002), 425–436.
- [15] D. Rolfsen, *Knots And Links*, AMS Chelsea Publishing, Providence, Rhode Island, 1976.
- [16] Y. Saka and M. Vazquez, *Tanglesolve: topological analysis of site-specific recombination*, Bioinformatics, **18** (2002), 1011–1012.
- [17] S.J. Spengler, A. Stasiak, and N.R. Cozzarelli, *The stereostructure of knots and catenanes produced by phage  $\lambda$  integrative recombination : implications for mechanism and DNA structure*, Cell, **42** (1985), 325–334.
- [18] D. W. Sumners, C. Ernst, N.R. Cozzarelli, and S.J. Spengler, *Mathematical analysis of the mechanisms of DNA recombination using tangles*, Quarterly Reviews of Biophysics, **28** (1995).
- [19] Alexandre A. Vetcher, Alexander Y. Lushnikov, Junalyn Navarra-Madsen, Robert G. Scharein, Yuri L. Lyubchenko, Isabel K. Darcy, and Stephen D. Levene, *DNA topology and geometry in Flp and Cre recombination*, J. Mol. Biol., **357** (2006), 1089–1104.
- [20] Alexander V. Vologodski <http://www.biophysics.org/btol/supramol.html#13> Online Biophysical Chemistry Textbook, ed. V. Bloomfield, 1999
- [21] S.A. Wasserman and N.R. Cozzarelli, *Determination of the stereostructure of the product of Tn3 resolvase by a general method*, Proc. Nat. Acad. Sci. U.S.A., **82** (1985), 1079–1083.
- [22] 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.
- [23] J. Watson and F. Crick, *A structure for deoxyribose nucleic acid*, Nature, **171**, 737-738.