BIOMATH
https://biomath.math.bas.bg/biomath/index.php/biomath
Biomath Forum

ORIGINAL ARTICLE

Topological Process of Splitting DNA-Links

Abdul Adheem Mohamad[1], Tsukasa Yashiro[2]

[1]College of Arts and Sciences, University of Nizwa, Oman
mohamad@unizwa.edu.om

[2]Independent Mathematical Institute, Miyota, Kitasaku, Nagano, Japan
t-yashiro@dokusuken.com

Received: 30 November 2021, accepted: 28 March 2022, published: 16 May 2022

***Abstract*—A DNA replicon is modeled by a special type of 2-component link, called a DNA-link, in which two circles form a double helix around a trivial center core curve. The DNA replication process is semi-conservative, which is interpreted as a splitting process of the DNA-link. To split this non-trivial link, the linking number must become zero, and thus an unknotting operation is necessary. Some families of enzymes act as the unknotting operation. The present paper considers two topological problems; one is to know how the linking number is reduced and the other, how the enzymes are allocated at appropriate places. For the first problem, we suggest a reduction system of the linking number of a DNA-link. From this system, the number of repetitions of the procedure is obtained and this could be reduced when the DNA is previously relaxed by type I topoisomerases. For the second problem, we propose a possible conformation of the DNA-link in which the unknotting operation does not change the knot type of the core curve but decreases the writhe. This conformation could allocate type II topoisomerases to appropriate places. These models suggest that the combination of type I and type II topoisomerases efficiently reduces the linking number and it is possible to allocate enzymes by the conformation of DNA strands.**

***Keywords*-DNA; Replication; Link; Topological model; Replicon**

***MSC2010*-92B99**

I. Introduction

A DNA molecule has a double helical structure [24] along a curve that causes several topological problems when it is unwound [4], [1], [23]. During the transcription or replication process, tangled (catenated) strands occur. As the replication forks advance, the axial rotation introduces positive supercoils ahead of each of the replication forks, while the negatively supercoiled daughter DNAs are introduced behind the forks [7], [16], [18], [23]. It is known that topoisomerases are responsible for reducing stress and supercoils [1], [2], [16], [17], [22], [23]. Rybenkov et. al. [16] revealed the ability of the enzyme topoisomerase type II to simplify DNA topology. It has been pointed out that there are several topological problems caused by the double helical structure of DNA itself (see [23]). In this paper, we focus on the unwinding process of the double-strand DNA introducing supercoils and their reduction during

Citation: Abdul Adheem Mohamad, Tsukasa Yashiro, Topological Process of Splitting DNA-Links, Biomath 11 (2022), 2203288, https://doi.org/10.55630/j.biomath.2022.03.288

DNA replication.

In knot theory, a *link* is a set of embedded disjoint circles in 3-dimensional space [12], and its projected image into a plane is called a *link diagram* (Section IV-A) [12]. It is known that for any link diagram, there is a finite sequence of crossing changes (unknotting operations), which deforms the link into separated trivial circles called a *trivial link*. However, this does not mean that we can specify the set of crossings that should be changed to obtain a trivial link (Example IV.1). Mathematically, if we are given a piece of global information such as a link diagram in a plane, it is possible to find such a set, but it is not known how to find such a set from only local information.

Since each replicon is fixed at substructure of the nucleus, topologically, it is modeled by a special 2-component link called a DNA-link (see Section IV) and also, the semi-conservative replication is interpreted in terms of the DNA-link; that is, the DNA-link is deformed into a disjoint union of trivial circles (Lemma IV.2). An oriented link is characterized as an algebraic invariant called the *linking number* (see Section IV-C for the definition) which is obtained as the sum of the total number of full-twists and the writhe of the core curve (Lemma IV.3). To split the DNA-link is equivalent to having the linking number zero (Lemma IV.4). Therefore, the following problem arises.

P1: How is the linking number of the DNA-link reduced?

To solve problem P1, we propose a procedure to reduce the linking number of DNA-link. This procedure considers the situation that only type II topoisomerase is used for the reduction. Using this procedure, we will obtain a formula for the number of repetitions of the procedure (Proposition V.2), by this formula, it is at least 11.5. This observation suggests that a relaxation of the double strand by type I topoisomerase should accelerate the reduction of the linking number. The procedure above focuses only on the numeral calculations, not the location of unknotting operations. It should be noted that a random choice of crossings to apply the unknotting operations does not always reduce the linking number (see Example 5). Thus the next problem arises.

P2: How are the enzymes allocated to the right positions of crossings that need to be changed to resolve a tangled DNA?

To answer P2, we propose a topological conformation of DNA in which the chromatin fibre forms a zig-zag shaped (juxtaposed) formation (see Figure 8). This model suggests that a special local conformation enforces the enzyme to locate the right crossing.

This paper is organized as such: Section II is a brief description of DNA replication. Section III briefly describes topoisomerases type I and II. Section IV introduces DNA-links and unknotting operations. Section V introduces a possible procedure to reduce the linking number and estimate the number of repetitions of the procedure. Section VI introduces a topological model for unknotting operations. Section VII discusses the previously obtained results. Section VIII states the conclusion.

II. DNA REPLICATION

The DNA replication is done along the loop-shaped DNA called a *replicon* [1], [4], [20], the ends of which are anchored at a substructure of the nucleus called a *nuclear matrix* (NM) [14], [15]. A group of replicons forms a *replicon domain* or simply *domain* [20]. Each replicon has a specific site called a *replication origin* or simply *origin* to unwind (relax) and to separate the double strand into a pair of single strands. We sometimes write the double strand DNA as a *ds-DNA*, and the single strand DNA as an *ss-DNA*. The separation occurs at two branching points called *replication forks* or simply *forks*. It is known that there are several models of forks whether the forks move or not [4], [15]. In this paper, 'ahead of a fork' means one side of the fork in which the ds-DNA has not yet separated into ss-DNAs.

As synthesizing proceeds, it is either continuous along a single strand, called a *leading strand* or discontinuous along with the other, called a

lagging strand. Along the lagging strand, short synthesized segments (100 bp to 200 bp) called Okazaki fragments [13] are constructed. The separated single strands of their parent DNA are preserved in its daughter DNAs. This is said to be *semi-conservative* [8].

A nucleosome is an essential element of the chromatin and it is formed with the double strand DNA (ds-DNA) wrapping around the histone octamer [1], [20]. The DNA is stored in the nucleus through a compactification involving nucleosomes [3], [5], [6]. In the interphase of the eukaryotic cell cycle, the DNA exists in the format with nucleosomes called *chromatin fibre* [3], [5], [11], [6]. Although a variety of models of organizations for chromatin fibre has been suggested [3], [5], [6], [11], [21], the exact conformation of the chromatin fibre has not yet been settled.

III. Topoisomerases

There are two types of topoisomerases, type I and type II, and they have several subfamilies such as types IA and IB [1], [2]. Type I is to release the rotational stress of single strands while type II releases the stress of the double strands [1], [7], [16], [17], [23]. In this section, we briefly describe type IA, type IB, and type II. The topoisomerase type IA acts on a ds-DNA to cut one single strand of it and lets the other single strand pass through the gap and reseal the gap. This reduces one full-twist of the ds-DNA. The topoisomerase type IB acts on a ds-DNA to cut one single strand so that it creates free ends and let one of the ends rotate around the other complete strand multiple times and reseal the gap at the end. This reduces multiple full-twists of the ds-DNA. The topoisomerase II acts on a pair of ds-DNAs to capture the pair of ds-DNA segments and cut one of them to make a gap, then lets the other ds-DNA pass through the gap and reseal it. Finally, those ds-DNAs are released (see Figure 1).

It is observed in [17] that topoisomerase type II is more efficient than topoisomerase type I to untangle DNA strands. Note that when the type II topoisomerase acts on the pair of ds-DNA segments, these two segments must be close enough so that the enzyme can capture both segments (see Figure 1-(a)).

IV. DNA-links

A. Links

A *link* is a disjoint union of circles embedded in $\mathbf{R}^3$. Each of the circles is called a *component*. If the number of components is n, then the link is called an *n-component link*. A 1-component link is called a *knot*. If a link L can be deformed into a link L' without any cut or intersecting in $\mathbf{R}^3$, then L and L' are in the same *link type* and L and L' are said to be *equivalent* and they have the same type (see [12] for details). If a knot bounds a disc in 3-space, then it is said to be *trivial*. If a link L is equivalent to a disjoint union of trivial knots, then the link is said to be *trivial*.

The image of a link L under the orthogonal projection from $\mathbf{R}^3$ to $\mathbf{R}^2$ by omitting the last coordinate, is a diagram with finite number of crossings every which is formed by two short segments; one is higher and the other is lower with respect to the last coordinate. This is called *crossing information*. The image of L under the orthogonal projection with the crossing information is called a *link diagram* denoted by D_L. If every component of a link L has an orientation, (oriented), L is called an *oriented link*.

On a link diagram, there are three elementary moves Ω_1, Ω_2 and Ω_3, called Reidemeister moves (see Figure 2).

Lemma IV.1 ([12])**.** *Two diagrams of equivalent links are deformed into each other by a finite sequence of Reidemeister moves.*

Proof: A proof can be found in [12]. ■

B. Topological semi-conservative scheme

A ds-DNA can be viewed as the boundary components $\{S_1, S_2\}$ of a long thin twisted strip with the trivial centre circle γ. We assume here, the centre curve γ is oriented and the components are parallelly oriented along γ. We write this as $L = L(S_1, S_2; \gamma)$, where S_1 and S_2 represent the single strands of the DNA. To define a topological

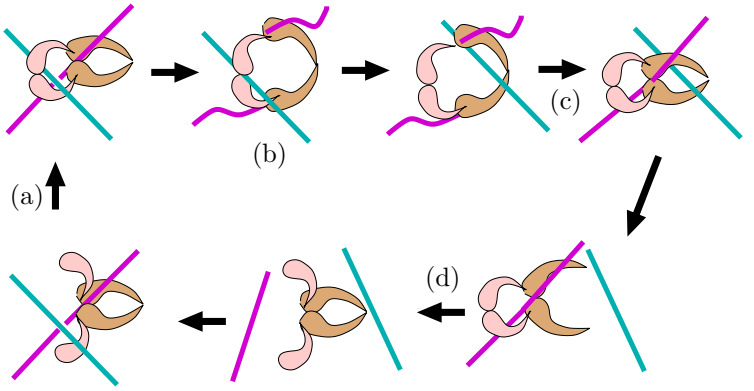


Fig. 1. (a) Topoisomerase II captures the pair of ds-DNA strands. (b) It cuts one of those strands to make a gap. (c) The other strand passes through the gap. (d) Reseals the gap and releases them.

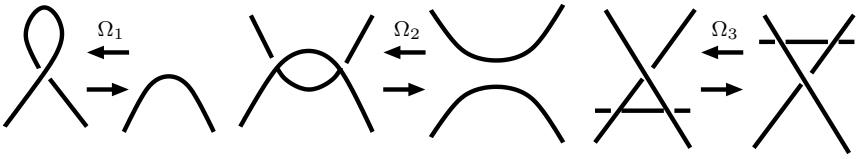


Fig. 2. Two diagrams of equivalent links are moved to each other by a finite sequence of three moves Ω_1, Ω_2 and Ω_3.

model of a replicon, we assume that the model satisfies the following conditions.

A1. The core curve γ is unknotted (trivial) and oriented.

A2. the link $L(S_1, S_2; \gamma)$ is an oriented 2-component link in which S_1 and S_2 are parallely oriented along the orientation of γ and they form a positive double helical structure around γ.

We call $L(S_1, S_2; \gamma)$ satisfying all the assumptions above, a *DNA-link*.

Let $L_0 = L_0(S_1, S_2; \gamma_0)$ be a DNA-link. After the DNA is replicated and distributed into two daughter cells, there are two identical DNA-links representing daughter DNAs,

$$L_1 = L_1(S_1', \bar{S}_1, \gamma_1), \tag{1}$$
$$L_2 = L_2(S_2', \bar{S}_2, \gamma_2), \tag{2}$$

where S_1' and S_2' are single strands (templates) inherited from L_0. $\bar{S}_1$ and $\bar{S}_2$ represent counterparts of S_1 and S_2 respectively and γ_1 and γ_2 are centre curves of the strips for L_1 and L_2 respectively.

The semi-conservative scheme is interpreted in terms of DNA-link $L(S_1, S_2; \gamma)$:

Lemma IV.2 ([10])**.** *The semi-conservative scheme is interpreted as such: the DNA-link $L_0(S_1, S_2; \gamma)$ is deformed into the split 2-component link $\{S_1', S_2'\}$, where S_i' is obtained from S_i $(i = 1, 2)$ by applying unknotting operations to L_0.*

C. Linking number

Let D_L be a link diagram of an oriented link L. At a crossing point of D_L, there are two types of crossings formed by short subarcs of D_L; positive and negative crossings (see Figure 4).

Let $L(S_1, S_2)$ be an oriented link with link components S_1 and S_2. Let $\mathcal{C}(D_L)$ be the set of crossings of the diagram D_L. The *linking number* is defined by

$$\mathrm{Lk}(S_1, S_2) = \frac{1}{2} \sum_{c \in \mathcal{C}(D_L)} \varepsilon(c) d(c)$$

where c is a crossing of the link diagram, $\varepsilon(c)$ is the sign ± 1 according to the diagrams in Figure 4,

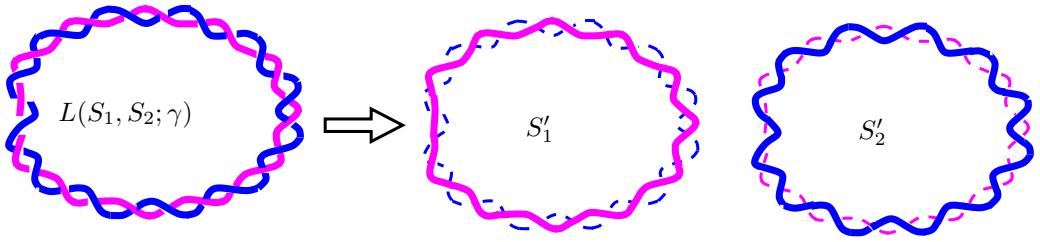


Fig. 3. The semi-conservative scheme is topologically interpreted as splitting the DNA-link.

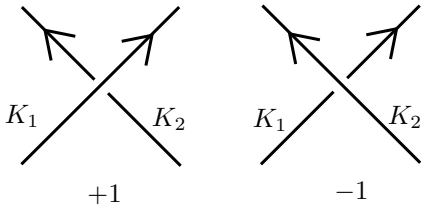


Fig. 4. Crossings with signs

and also

$$d(c) = \begin{cases} 1 & \text{if the crossing } c \text{ consists of} \\ & \text{distinct components,} \\ 0 & \text{otherwise.} \end{cases}$$

Let K be a knot and let D_K be a knot diagram. The total sum of signs $w(D_K)$:

$$w(D_K) = \sum_{c \in \mathcal{C}(D_K)} \varepsilon(c)$$

is called a *writhe* of K.

Note IV.1. *The linking number does not depend on the choice of the diagram of L (see [12]). If a 2-component link is split, then the linking number between the components is of course zero but the converse is not always true (see [12]). The writhe depends on the choice of diagram (see [12]).*

White proved in [25] the following formula of the linking number $\mathrm{Lk}(S_1, S_2)$:

Lemma IV.3 (White [25])**.** *For a DNA modeled by the DNA-link $L(S_1, S_2; \gamma)$, the following formula of the linking number holds:*

$$\mathrm{Lk}(S_1, S_2) = \mathrm{Tw}(S_1, S_2) + \mathrm{Wr}(\gamma), \qquad (3)$$

where $\mathrm{Tw}(S_1, S_2)$ is the number of full-twists of the curves $\{S_1, S_2\}$ along the centre curve γ and $\mathrm{Wr}(\gamma)$ is the writhe of γ.

Proof: A proof can be found in [25]. ■

The following is easily verified (see [9]).

Corollary IV.1. *Suppose that a DNA-link $L(S_1, S_2; \gamma)$ has a trivial γ. Then $\mathrm{Lk}(S_1, S_2) = 0$ if and only if $L(S_1, S_2; \gamma)$ is split.*

Therefore, the splitting process of the DNA-link is equivalent to that making the linking number zero.

Lemma IV.4. *The semi-conservative scheme is interpreted to make the linking number of the DNA-link zero.*

Proof: Combining Lemma IV.2 and Corollary IV.1, the result follows. ■

We understand that the contribution to the writhe from the conformation of DNA of a degree higher than the nucleosomes should be considered. However, as this paper focuses on an individual replicon, and for sake of simplicity, we ignore the contribution from the conformation of a degree higher than the nucleosomes.

D. Unknotting operations

There is an operation to exchange the over arc and the under arc, called an *unknotting operation* (see Figure 5-(a)).

Lemma IV.5 (Proposition 4.4.1 [12])**.** *For any knot diagram of a non-trivial knot, it is deformed into a knot diagram of a trivial knot by applying a finite number of unknotting operations.*

Proof: A proof of this lemma can be found in [12]. ■

Note that this lemma gives a guarantee to modify every knot into a trivial knot by applying a finite number of unknotting operations. However, it should be emphasized that Lemma IV.5 does not say which crossings should be changed to obtain the trivial knot. It should be noted that a random choice of a sequence of the deformation does not work. For instance, as we can see in the diagram in Figure 5-(b), it is a trivial knot but even one crossing change creates a non-trivial knot. Conversely, if a non-trivial knot is given, if we choose a wrong sequence of crossings to exchange, then we cannot reach a trivial knot. This observation suggests that the conformation of a supercoil must have a certain format and size to allocate the enzymes to the right place.

Example IV.1. The diagram (b) in Figure 5 is a diagram of a trivial knot γ with loops and some crossings. If we change the crossing at the top, then we obtain a non-trivial knot (the figure-eight knot). On the other hand, if we change two crossings at the bottom (Figure 5(c)), then it keeps the triviality and decreases the writhe by 4.

As we can see in the proof of Proposition 4.4.1 in [12], if a link diagram is given, then there is a way to specify the set of crossings to be changed to obtain a diagram of a trivial link. However, we do not know how to detect such a set of crossings from only local information such as a set of crossings.

V. A REDUCTION PROCESS

We make the following assumption.

A3 The number of unwound full twists is equal to the number of crossings in the positive supercoil ahead of the fork.

The semi-conservative scheme implies that the DNA-link must be deformed into a trivial link. By Corollary IV.1, to split a DNA-link, the linking number must be zero. We define the following procedure to reduce the linking number consisting of the following steps.

S1 Unwind n full twists at a specified point of the DNA-link to create a pair of forks.
S2 Create positive n crossings (supercoil) in front of each fork.
S3 Apply the unknotting operations on the n crossings of the supercoil to obtain $-n$ crossings.
S4 If the linking number is not zero, then go back to S1.

To determine the number n, it is natural that n is proportional to the initial twisting number Tw_0. So,

$$n = c\mathrm{Tw}_0, \qquad 0 < c < 1 \tag{4}$$

We start the process with the initial linking number Lk_0:

$$\mathrm{Lk}_0 = \mathrm{Tw}_0 + \mathrm{Wr}_0, \tag{5}$$

where Wr_0 is the initial writhe of γ. Following the steps from S1 to S4, we obtain the following sequence of values Tw_k and Wr_k, $k = 0, 1, \ldots$.

$$\mathrm{Tw}_1 = \mathrm{Tw}_0 - c\mathrm{Tw}_0 = \mathrm{Tw}_0(1-c) \tag{6}$$
$$\mathrm{Wr}_1 = \mathrm{Wr}_0 - c\mathrm{Tw}_0 \tag{7}$$
$$\mathrm{Lk}_1 = \mathrm{Tw}_1 + \mathrm{Wr}_1 \tag{8}$$
$$= \mathrm{Tw}_0(1-c) + \mathrm{Wr}_0 - c\mathrm{Tw}_0 \tag{9}$$
$$= \mathrm{Lk}_0 - 2c\mathrm{Tw}_0 \tag{10}$$

If the procedure ends at this stage; that is, $\mathrm{Lk}_1 = 0$, then the number n is half of Lk_0. This is almost half of Tw_0. However, this is not possible without changing Tw_0 because, to form a crossing in which two segments of double strand DNA become very close, a DNA segment with a certain length must necessarily be bent. This implies that the number of unwound twists becomes much larger than the possible number of crossings introduced ahead of the fork.

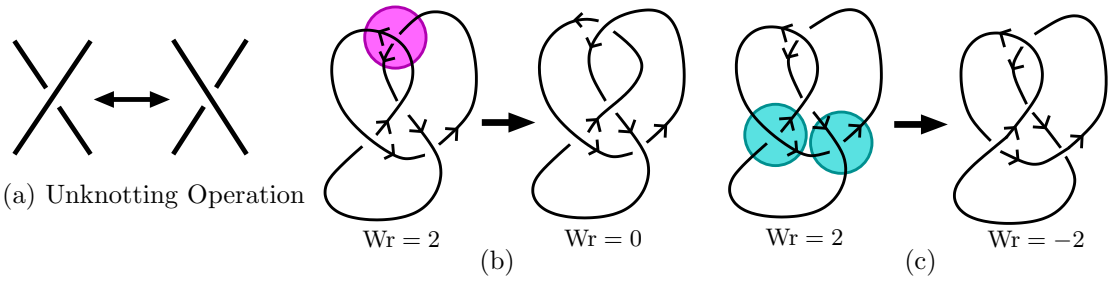


Fig. 5. We cannot change the randomly selected crossings to obtain a required knot type. (a) The unknotting operation. (b) The unknotting operation at the top crossing leads to a non-trivial knot. (c) The uknotting operations at the bottom crossings keep decreasing the writhe and keep the triviality of γ.

A. The number of repetitions

Next, we consider how many times we have to repeat the procedure.

We assume that the number of crossings unwound in each cycle is the same as the number of nucleosomes existing ahead of the fork. This assumption will be justified in Section VI. Then we have the identity:

$$n = c\mathrm{Tw}_0 = \frac{\mathrm{Tw}_0}{l}, \tag{11}$$

where l is the number of full twists within the DNA around a nucleosome and its linker DNA. Note that if the nucleosomes are distributed uniformly along a replicon, then $\frac{\mathrm{Tw}_0}{l}$ is the number of nucleosomes in the replicon. We denote this by τ_0.

Although, a recent study in [19] shows that the writhe contributed to each nucleosome is -1.26, for sake of generality, here we use $\alpha > 0$ as the contribution of writhe for each nucleosome. Thus the initial writhe for the replicon is

$$\mathrm{Wr}_0 = -\alpha\tau_0 = -\alpha c\mathrm{Tw}_0 \tag{12}$$

Proposition V.1. *Let L be a DNA-link with the initial twists Tw_0 and the initial writhe Wr_0. If the reduction process is applied to L k times, then the linking number Lk_k is given by the following.*

$$\mathrm{Lk}_k = \mathrm{Tw}_0\left[2\left(1-c\right)^k - \left(1+\alpha c\right)\right], \tag{13}$$

where α is the writhe contribution to each nucleosome, and c is the rate of unwond full twists to Tw_0.

Proof: After applying the unknotting operations to the chromatin fibre at the first stage, the number of full-twists Tw_1 is given by

$$\begin{aligned}\mathrm{Tw}_1 &= \mathrm{Tw}_0 - \tau_0\\ &= \mathrm{Tw}_0 - c\mathrm{Tw}_0\\ &= \mathrm{Tw}_0\left(1-c\right)\end{aligned}$$

The number of nucleosomes τ_1 at the second stage is given by

$$\tau_1 = c\mathrm{Tw}_1 = c\mathrm{Tw}_0\left(1-c\right)$$

After applying the unknotting operations to the chromatin fibre at the second stage,

$$\begin{aligned}\mathrm{Tw}_2 &= \mathrm{Tw}_1 - \tau_1\\ &= \mathrm{Tw}_0\left(1-c\right) - c\mathrm{Tw}_0\left(1-c\right)\\ &= \mathrm{Tw}_0\left(1-c\right)^2\end{aligned}$$

At the kth stage,

$$\mathrm{Tw}_k = \mathrm{Tw}_0\left(1-c\right)^k, \tag{14}$$

whrere k is the number of repetition of the deformation cycles.

On the other hand, the initial writhe Wr_0 is given by the following.

$$\mathrm{Wr}_0 = -\alpha\tau_0 = -\alpha c\mathrm{Tw}_0$$

For further stages,

$$\begin{aligned}\mathrm{Wr}_1 &= \mathrm{Wr}_0 - \tau_0 \\ &= -\alpha c\mathrm{Tw}_0 - c\mathrm{Tw}_0 \\ &= c\mathrm{Tw}_0(-\alpha - 1) \qquad (15)\\ \mathrm{Wr}_2 &= \mathrm{Wr}_1 - \tau_1 \\ &= c\mathrm{Tw}_0(-\alpha - 1) - \tau_1 \\ &= c\mathrm{Tw}_0\left(-\alpha - 1 - (1-c)\right) \qquad (16)\\ \mathrm{Wr}_3 &= c\mathrm{Tw}_0\left(-\alpha - 1 - (1-c) - (1-c)^2\right) \qquad (17)\end{aligned}$$

$$\vdots$$

$$\begin{aligned}\mathrm{Wr}_k = c\mathrm{Tw}_0\big(&-\alpha - 1 - (1-c) - \cdots \\ &- (1-c)^{k-1}\big) \qquad (18)\end{aligned}$$

Applying the formula of a geometric series, we obtain the following.

$$\begin{aligned}\mathrm{Wr}_k &= -c\mathrm{Tw}_0\left(\alpha + \frac{1-(1-c)^k}{1-(1-c)}\right) \\ &= -\mathrm{Tw}_0\left(1+\alpha c\right) + \mathrm{Tw}_0\left(1-c\right)^k \qquad (19)\end{aligned}$$

Therefore, the sum of (14) and (19) is the linking number after applying the procedure k times.

$$\begin{aligned}\mathrm{Lk}_k &= \mathrm{Tw}_k + \mathrm{Wr}_k \\ &= 2\mathrm{Tw}_0\left(1-c\right)^k - \mathrm{Tw}_0\left(1+\alpha c\right) \\ &= \mathrm{Tw}_0\left[2\left(1-c\right)^k - \left(1+\alpha c\right)\right] \qquad (20)\end{aligned}$$

■

Proposition V.2. *Suppose that the reduction system is applied to a DNA-link multiple times to obtain the linking number zero. The number of the repetitions k is given by*

$$k = \frac{\ln\left(\frac{1+\alpha c}{2}\right)}{\ln\left(1-c\right)} \qquad (21)$$

Proof: Suppose $\mathrm{Lk}_k = 0$, we obtain

$$\begin{aligned}\mathrm{Tw}_0\left[2\left(1-c\right)^k - \left(1+\alpha c\right)\right] &= 0 \\ 2\left(1-c\right)^k &= \left(1+\alpha c\right) \\ k\ln\left(1-c\right) &= \ln\left(\frac{1+\alpha c}{2}\right) \\ k &= \frac{\ln\left(\frac{1+\alpha c}{2}\right)}{\ln\left(1-c\right)} \qquad (22)\end{aligned}$$

■

Since the number α is a constant, k is determined by the parameter c, which is given by Tw_0.

We consider the nucleosome in which the DNA wraps around the histone core as a unit of the bending. The diameter of the histone core is about $6.4\,\mathrm{nm}$, and the diameter of DNA is $2\,\mathrm{nm}$ [1], [20]. The DNA wraps around a histone core about 1.8 times [1], [20] and each nucleosome is associated with a linker DNA. The total length is $197\,\mathrm{bp}$. Therefore, we obtain:

$$l = \frac{197}{10.5} \approx 18.8 \qquad (23)$$

The number of unwound full twists depends on the parameter l; that represents the relaxation of the ds-DNA. A recent study in [19] shows that the writhe contributed to each nucleosome is -1.26. Substituting (23) and $\alpha = 1.26$ to the formula (21) of k, we obtain $k \approx 11.5$.

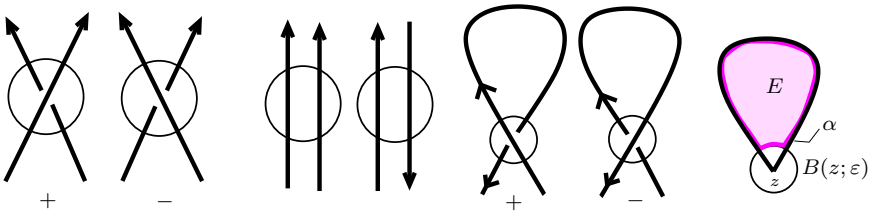
This implies that if we apply the reduction process on the core curve of the DNA-link, then we need to repeat the process 11.5 times. The linking number Lk_k in (20) is a function of $c = 1/l$ which is determined by the relaxation of the double strand DNA done by type I topoisomerase. In fact, from the formula (21) with $\alpha = 1.26$, if a DNA-link is relaxed so that $l = 3.26$, then $k = 1$.

VI. Topological Model

A. ε-crossings

Let γ be an oriented knot. Let $x, y \in \gamma$ be two distinct points, and let $\varepsilon > 0$ be some number.

Let $B(x;r)$ denote an open 3-ball in $\mathbf{R}^3$, centered at the point x and with radius r. Suppose that there is a point $z \in \mathbf{R}^3 \setminus \gamma$ such that $x, y \in B(z;\varepsilon/2)$. If $\gamma \cap B(z;\varepsilon/2)$ is a pair of line segments e_1 and e_2 such that $x \in e_1$ and $y \in e_2$, then we say e_1 and e_2 form an *ε-crossing*. The ε-crossing has a sign $+$ or $-$ according to the orientation (see left two diagrams in Figure 6). We do not admit the exact parallel cases (the middle two diagrams in Figure 6) as the ε-crossings. A *loop* is a simple sub-arc of a knot, from an ε-crossing to itself (see Figure 6-(c)).



(a) ε-crossings (b) (c) loops (d)

Fig. 6. Singed ε-crossings and loops.

Suppose that the boundary of a disc E, ∂E is the union of portion of $\alpha \setminus B(z;\varepsilon)$ and a simple arc on $\partial B(z;\varepsilon)$ (see Figure 6-(d)), then we say the loop α based at an ε-crossing bounds a disc E.

Note that we have the following fact.

Lemma VI.1. *Let γ be an oriented knot. Let α be a loop of γ based at an ε-crossing. If α bounds an embedded disc E in $\mathbf{R}^3$, and the interior of E does not meet γ, then applying an unknotting operation at the ε-crossing does not change the knot type of γ.*

Proof: Without loss of generality, we can assume the ε-crossing of the oriented loop is positive. Since the loop α with the ε-crossing bounds a disc E in $\mathbf{R}^3$, the loop can be deformed into the diagram shown in the left diagram of Figure 6-(c). Then we can apply the reversed Reidemeister move I to remove the ε-crossing and apply the Reidemeister move I to create a loop based at the negative ε-crossing (see the right diagram of Figure 6-(c)). The resulting diagram is equivalent to that obtained by a crossing change at the ε-crossing. Since the move does not change the knot type (see [12] for details), the unknotting operation at the ε-crossing does not change the knot type. ■

B. A modeling policy

As Example IV.1 demonstrates, it is not easy to deform a non-trivial knot into a trivial knot. Therefore, it is natural to assume that the center curve γ remains trivial during the replication. We assume the following.

A4 The imaginary core curve γ keeps its triviality during the replication process.

Note that although the core curve itself is trivial, it may have a certain complexity.

C. Elementary twists

In order to solve the problem P2, a special type of conformation of DNA needs to be introduced so that the conformation allocates the enzymes to the suitable positions.

Suppose a short segment of DNA is U-shaped (see Figure 7 (a)) and one end is fixed while the other side is rotated around the axial curve. This rotational stress will introduce a loop, based at ε-crossing (Figure 7 (b)) so that the writhe is increased by 1. Then apply the unknotting operation at the crossing (Figure 7). Here the writhe is decreased by -1. Then the segment returns to the initial position. During the move, the segment is fully twisted twice around the axis. This deformation will be called an *elementary twist*.

The chromatin fibre has a sequence of nucleosomes [20]. We suppose that the chromatine fibre has a zig-zag shape (juxtapositioned) shown as the diagram in Figure 8 (see also [17]). Then near a nucleosome, it has a U-shape region. We can apply the elementary twist around each nucleosome so that the obtained ε-crossings near nucleosomes is changed (see Figure 8 (c) and (d)). In this conformation, as each ε-crossing is close enough to the histone core, we can assume the following.

Supposition 1. *The loop based at the ε-crossing near the histone core bounds an embedded disc E in $\mathbf{R}^3$ and the interior of E does not meet the DNA strand.*

This means that the histone core plays the

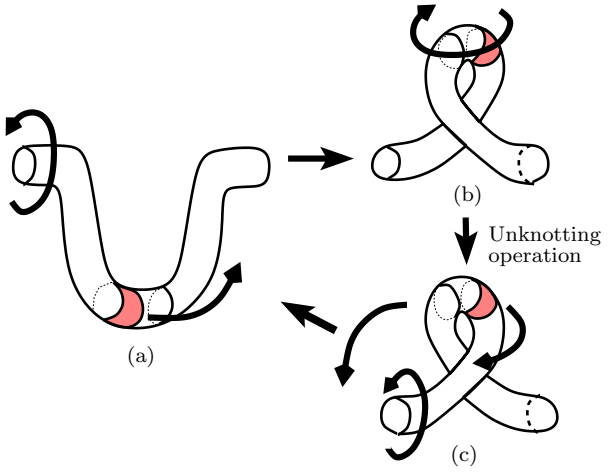


Fig. 7. (a) The U-shaped string is twisted around the axial curve. (b) A positive loop is created. (c) Applying unknotting operations to the crossing, the negative crossing is obtained.

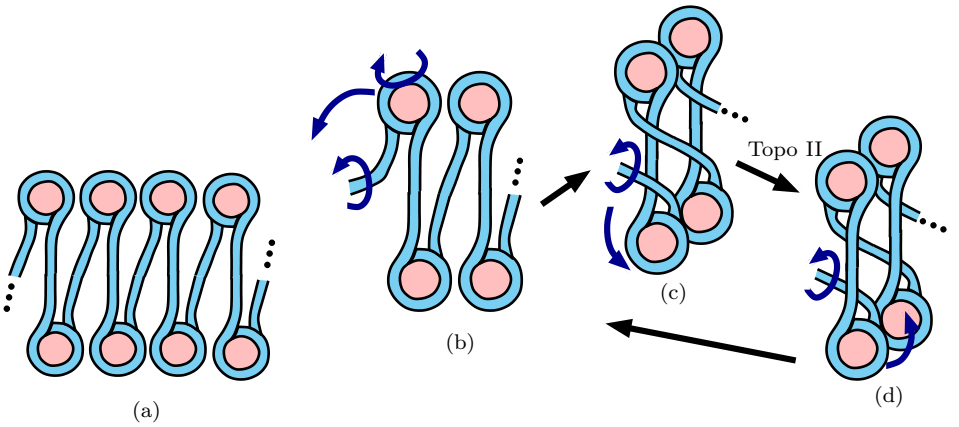


Fig. 8. Schematic diagrams showing the unknotting operation done by Topo II at crossings near the nucleosomes.

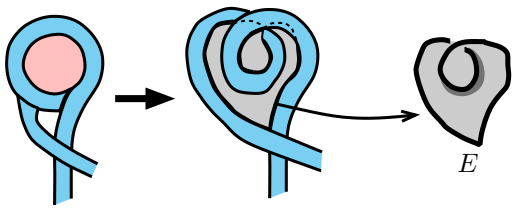


Fig. 9. The loop around the histone octamer with a ε-crossing close enough to the histone octamer is supposed to bound an embedded disc E.

role of a disc E in Lemma VI.1. Therefore, by Lemma VI.1, the unknotting operation at the ε-crossing does not change the knot type of γ. We call this process an *unknotting process* of a DNA-link (replicon).

The process follows the steps below.

- S1 Unwinding n full-twists at the replication fork will introduce a positive supercoil with n crossings ahead of the fork (see Figure 8c)).
- S2 When the number of the crossings reaches the maximum, activate the type II topoisomerases to the crossings.
- S3 Continue the move to return the conformation to the original shape.

This unknotting process keeps γ trivial. As discussed in Section V, the number of repetitions of the deformation above depends on the numbers l and α. Using this process in the reduction process explained in Section V, the assumption that the number of crossings unwound in each cycle is the same as the number of nucleosomes existing ahead of the fork, is justified.

VII. Discussion

To solve problem P1, a schematic procedure of reducing the linking number of DNA is considered. The procedure gives the formula (13) in Proposition V.1 to obtain the reduced linking number. On the other hand, Proposition V.2 gives the formula (21), and it implies that if only the unknotting operations, namely, the type II topoisomerase, is used, the number of repetitions is about 11.5. Although it is difficult to say whether this number is large or small, the formula (21) depends on two parameters α and $c = 1/l$, which may vary under the relaxation of the double strand DNA [1]. This suggests that if the ds-DNA is relaxed in prior, then the number of repetitions could be much smaller. For instance, the authors proposed in [10] the reduction of Tw with type I topoisomerase, in which Tw is reduced to $20\,\%$ of the initial twists by the end of the first stage. This gives $l' = l/5$; that is, $c' = 5c$. Assuming the same contribution of writhe from the nucleosome, $\alpha = 1.26$, substituting c' and $\alpha = 1.26$, to (21), the number of repetitions is about 1.3. Combining type I and II topoisomerases to simplify DNA has been pointed out in researches [2], [7], [22], [26] from different viewpoints.

Next, in order to solve P2, we proposed a model of a mechanism to allocate type II topoisomerases to suitable crossings. As we have seen in Example IV.1, we cannot randomly choose the set of crossings to apply unknotting operations to deform a non-trivial knot into a trivial knot. Therefore, it is natural to assume that the model does not change the triviality of the core curve of a DNA-link. The model has zig-zag shaped (juxtaposed) nucleosomes in which the axial rotation introduces a trivial loop with a crossing near each of the nucleosomes. For this crossing, the unknotting operation reduces the writhe, but it does not change the triviality of the core curve. Also, this guarantees that the core curve is always trivial during the replication process.

VIII. Conclusion

From the observation of the proposed procedure, we obtained 11.5 as the necessary number of repetitions to make the DNA-link split. This number is parametrized by two parameters, α and $c = 1/l$, and c will be changed by relaxation of double strand DNA by type I topoisomerase. Therefore, a combination of two types of topoisomerases efficiently reduce the linking number. As we have seen in Example IV.1, specifying the location of topoisomerase II is an essential issue to make the linking number zero. Our model provides the mechanism that allocates enzymes to the right position and the action of type II topoisomerase does not change the knot type of the core curve. From the arguments about the procedure and the model, we can conclude that the linking number is efficiently reduced when two types of topoisomerases are combined, and it is possible to allocate type II topoisomerase to the appropriate places by the conformation of DNA.

In this research, we have not considered the reduction process of the negative supercoils behind the forks. This should be done in further study.

ACKNOWLEDGEMENT

The authors would like to thank Dr Jagir Hussan for giving valuable suggestions for the earlier version of the paper.

REFERENCES

[1] A.D. Bates, A. Maxwell, et al. *DNA topology.* Oxford University Press, USA, 2005.

[2] A. Hanke, R. Ziraldo, and S.D. Levene. "DNA-Topology Simplification by Topoisomerases". In: *Molecules* 26.11 (2021). ISSN: 1420-3049. DOI: 10.3390/molecules26113375.

[3] Y. Joti et al. "Chromosomes without a 30-nm chromatin fiber". In: *Nucleus* 3.5 (2012), pp. 404-410.

[4] A. Kornberg and T.A. Baker. *DNA Replication.* University Science Books, 2005.

[5] V.R. Lobbia, C.T. Sanchez, and H. van Ingen. "Beyond the nucleosome: nucleosome-protein interactions and higher order chromatin structure". In: *Journal of Molecular Biology* (2021), p. 166827.

[6] K. Maeshima, S. Ide, and M. Babokhov. "Dynamic chromatin organization without the 30-nm fiber". In: *Current opinion in cell biology* 58 (2019), pp. 95-104.

[7] A.K. McClendon, A.C. Rodriguez, and N. Osheroff. "Human topoisomerase IIα rapidly relaxes positively supercoiled DNA: implications for enzyme action ahead of replication forks". In: *Journal of Biological Chemistry* 280.47 (2005), pp. 39337-39345.

[8] M. Meselson and F.W. Stahl. "The replication of DNA". In: *Cold Spring Harbor symposia on quantitative biology.* Vol. 23. Cold Spring Harbor Laboratory Press. 1958, pp. 9-12.

[9] A.A. Mohamad and T. Yashiro. "A rewinding model for replicons with DNA-links". In: *BIOMATH* 9.1 (2020), p. 2001047.

[10] A.A. Mohamad and T. Yashiro. "A topological model of DNA replication with DNA-links". In: *Far East J. Mathematical Sciences* 107 (2018), pp. 241-255.

[11] M. Moraru and T. Schalch. "Chromatin fiber structural motifs as regulatory hubs of genome function?" In: *Essays in biochemistry* 63.1 (2019), pp. 123-132.

[12] K. Murasugi. *Knot theory and its applications.* Springer Science & Business Media, 2007.

[13] R. Okazaki. "Mechanism of DNA replication; possible discontinuity of DNA chain growth. (Symposium on biosynthesis of nucleic acids)". In: *Japanese Journal of Medical Science and Biology* 20.3 (1967), pp. 255-260.

[14] S.V. Razin et al. "Communication of genome regulatory elements in a folded chromosome". In: *FEBS letters* 587.13 (2013), pp. 1840-1847.

[15] J.C. Rivera-Mulia et al. "DNA moves sequentially towards the nuclear matrix during DNA replication in vivo". In: *BMC cell biology* 12.1 (2011), pp. 1-16.

[16] V.V. Rybenkov et al. "Simplification of DNA topology below equilibrium values by type II topoisomerases". In: *Science* 277.5326 (1997), pp. 690-693.

[17] J. Salceda, X. Fernández, and J. Roca. "Topoisomerase II, not topoisomerase I, is the proficient relaxase of nucleosomal DNA". In: *The EMBO journal* 25.11 (2006), pp. 2575-2583.

[18] J.B. Schvartzman et al. "Closing the DNA replication cycle: from simple circular molecules to supercoiled and knotted DNA catenanes". In: *Nucleic acids research* 47.14 (2019), pp. 7182-7198.

[19] J. Segura et al. "Intracellular nucleosomes constrain a DNA linking number difference of -1.26 that reconciles the Lk paradox". In: *Nature communications* 9.1 (2018), pp. 1-9.

[20] R.R. Sinden. *DNA structure and function.* Gulf Professional Publishing, 1994.

[21] A.K. Singh and F. Mueller-Planitz. "Nucleosome positioning and spacing: from mechanism to function". In: *Journal of Molecular Biology* 433.6 (2021), p. 166847. DOI: 10.1016/j.jmb.2021.166847.

[22] A.S. Sperling et al. "Topoisomerase II binds nucleosome-free DNA and acts redundantly with topoisomerase I to enhance recruitment of RNA Pol II in budding yeast". In: *Proceedings of the National Academy of Sciences* 108.31 (2011), pp. 12693-12698. issn: 0027-8424. DOI: 10.1073/pnas.1106834108.

[23] A. Vologodskii. "Disentangling DNA molecules". In: *Physics of life reviews* 18 (2016), pp. 118-134.

[24] J.D. Watson and F.H.C. Crick. "Genetical implications of the structure of deoxyribonucleic acid". In: *Nature* 171.4361 (1953), pp. 964-967.

[25] J.H. White. "Self-linking and the Gauss integral in higher dimensions". In: *American journal of mathematics* 91.3 (1969), pp. 693-728.

[26] G. Witz, G. Dietler, and A. Stasiak. "DNA knots and DNA supercoiling". In: *Cell Cycle* 10.9 (2011). PMID: 21393995, pp. 1339-1340. DOI: 10.4161/cc.10.9.15293.