BIOMATH
www.biomathforum.org/biomath/index.php/biomath
Biomath Forum

ORIGINAL ARTICLE

A rewinding model for replicons with DNA-links

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

[1]Mathematics Section, Department of Mathematical and Physical Sciences
College of Arts and Sciences, University of Nizwa
mohamad@unizwa.edu.om
[2]Independent Mathematical Institute
Miyota, Kitasaku, Nagano, Japan
t-yashiro@dokusuken.com

Received: 9 October 2019, accepted: 4 January 2020, published: 23 February 2020

***Abstract*—A double strand DNA has a double helical structure and it is modeled by a thin long twisted ribbon fixed at the both ends. A DNA-link is a topological model of such a DNA segment in the nuclear of a eukaryotic cell. In the cell cycle, the DNA is replicated and distributed into new cells. The complicated replication process follows the semi-conservative scheme in which each backbone string is preserved in the replicated DNA. This is interpreted in terms of splitting process of the DNA-link. In order to split the DNA-link, unknotting operations are required. This paper presents a recursive unknotting operations, which efficiently reduce the number of twistings.**

***Keywords*-DNA, replication, link, topological model, replicon**

Mathematics Subject Classification (2010) 92B99

I. INTRODUCTION

It is known that DNA is a polymer consisting of a set of base pairs and sugar-phosphate backbones (Watson-Click model (1953), see [2] [11]). The backbones form a double-helix structure with opposite directions induced from the ordered pair of $5'$ and $3'$. The $360°$-rotation of the helical strings is joined by about 10.5 base pairs. Therefore, it is natural that the double strand DNA (ds-DNA) is modeled by a long, thin strip twisted around the centre curve of the strip (see Figure 1, also [2][11]). The boundary curves of the strip form a double helix which correspond to the backbones of DNA [2][11]. A *full twist* of the ds-DNA is interpreted as the $360°$-rotation of the strip about the centre curve (see Figure 1).

The replication of DNA has been studied since the helical double strand structure was discovered (see [1][3][4][10][12]). In the interphase of the

Citation: Abdul Adheem Mohamad, Tsukasa Yashiro, A rewinding model for replicons with DNA-links, Biomath 9 (2020), 2001047, http://dx.doi.org/10.11145/j.biomath.2020.01.047

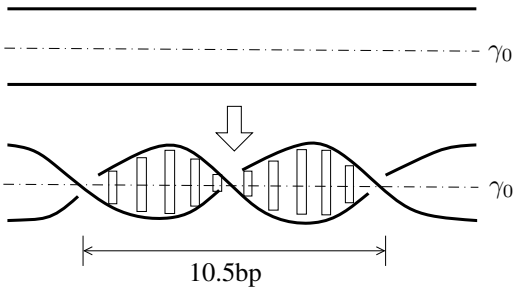

Fig. 1. The DNA double helix is modeled by a long thin strip.

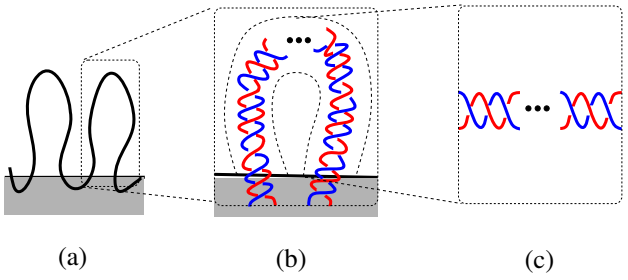

Fig. 2. (a) The linear DNA forms a set of loops attached to the nuclear matrix. (b) Each loop is anchored at the nuclear matrix (NM). (c) The anchored loop is modeled by twisted strings fixed at their ends of the boundary of a 3-dimensional ball, called a DNA-link.

eukaryotic cell cycle, there is a substructure of nucleus called the *nuclear matrix* (NM) (see [8][9]). It is believed that the ds-DNA is organized as a set of loop-shaped structures, called *replicons*. Each replicon has a specific site in which the replication starts. This is called a *replication origin* or simply, *origin*.

It is observed that two ends of the replicon (loop) are anchored at NM [8][9][14]. Thus it forms a pair of twisted strings fixed at the ends of the boundary of a 3-dimensional ball, which is a two-component tangle (see [6] for a definition). (see Figure 2 (c)); in this paper, we call this a DNA-link (see Section III for the definition).

In the DNA replication process, the sequence of base pairs along each backbone acts as a template to reproduce the base pairs in the synthesized DNA. This means that each backbone with base pairs is preserved in the synthesized DNA. This is called the *semi-conservative scheme*, we interpreted this scheme in terms of DNA-link in Lemma III.1.

As the replication process follows the semi-conservative scheme, the DNA-link must be split at the end. Topologically, it should be done by unknotting operations. Biologically, it is believed that enzymes topoisomerases TopoIA, TopoIB and TopoII are responsible for this operation. Our model uses only TopoIA.

This paper is organised as such: In Section II knots, links and linking numbers are introduced. In Section III, DNA-link is defined and the semi-conservative scheme is interpreted in terms of DNA-link (Lemma III.1). In order to rewind the ds-DNA it is natural to require the minimal energy and maximal length of rewound segment. We obtained Theorem III.1. Section IV describes a topological model for rewinding replicon. Our model is equipped with a recursive unknotting operations. This model can reduce the twisting number of

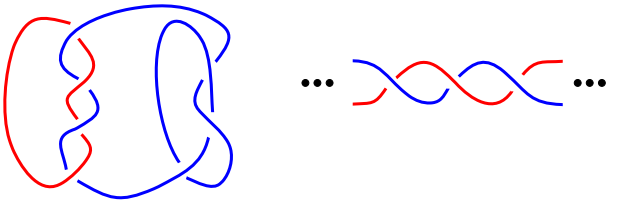

Fig. 3. Link diagrams of a usual link (left) and linearly very long double helix (right). Both are admitted as 2-component links.

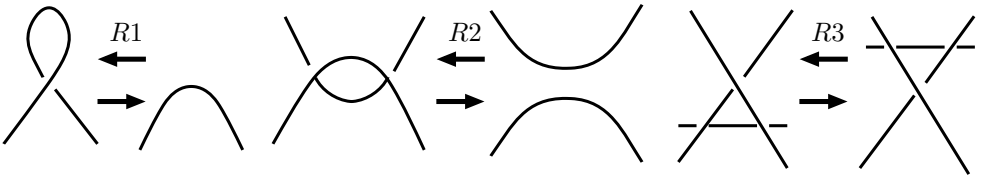

Fig. 4. Three types of Reidemeister moves.

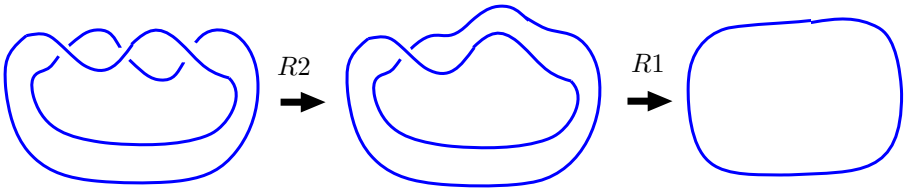

Fig. 5. The left diagram is deformed into the right by $R2$. The middle diagram is deformed into the very right by $R1$.

the DNA-link efficiently. The recursive unknotting operation resolve the unsoloved problem in [5].

II. Knots and Links

A *link* is a disjoint union of circles embedded in $\mathbf{R}^3$; K_1, $K_2, \ldots, K_n$ denoted by $L(K_1, K_2, \ldots, K_n)$. Each K_i $i = 1, 2, \ldots, n$ is called a *component* of the link L and the link L is called an *n-component link*. A 1-component link is called a *knot*. A knot is *trivial* if there is an embedded disc in $\mathbf{R}^3$ bounded by the knot. A link is *trivial* if the link consists of mutually split trivial link components. A *link diagram* D_L of a link L is a projected image of the orthogonal projection $(x_1, x_2, x_3) \mapsto (x_1, x_2)$ with crossing information at the crossings (see [6] for details).

There are three types of local moves on link diagrams depicted in Figure 4, called *Reidemeister moves*.

Two links L_1 and L_2 are *equivalent* if the diagram D_{L_1} is deformed into the diagram D_{L_2} by a finite sequence of three types of Reidemeister moves shown in Figure 4 [6]

Example II.1. The diagrams in Figure 5 are equivalent. The left diagram is deformed into the right by applying $R2$ move.

A. Split link

Let A and B be compact disjoint subsets of $\mathbf{R}^3$. The pair $\{A, B\}$ is said to be *split* if there exists an embedded sphere S in 3-space such that S bounds one of the subsets from the other. A 2-component link is called a *split link* if the components are split. A link L is said to be *oriented* if each component is oriented. A *link diagram* is the

image of L in the plane $\mathbf{R}^2$ under the orthogonal projection $(x_1, x_2, x_3) \mapsto (x_1, x_2)$ with crossing information; that is, at a crossing formed by two short arcs, one arc is *upper* and the other *lower* (see Figure 3). A link diagram of a link L is denoted by D_L.

B. Linking number

Let L be an oriented link and let D_L be a link diagram of 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 6).

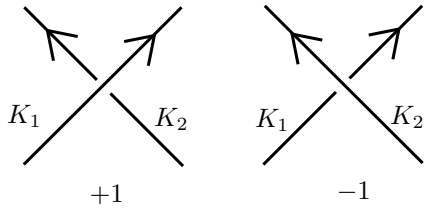

Fig. 6. Crossings with signs

Let $L(K_1, K_2)$ be an oriented DNA-link with distinct link components K_1 and K_2. Let $\mathcal{C}(D_L)$ be the set of crossings of the diagram D_L. The *linking number* is defined by

$$\mathrm{lk}(K_1, K_2) = \frac{1}{2} \sum_{c \in \mathcal{C}(D_L)} \varepsilon(c) d(c)$$

where c is a crossing of the link diagram, $\varepsilon(c)$ is the sign ± 1 according to the diagrams in Figure 6, 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 II.1. *The linking number does not depend on the choice of the diagram of L (see [6]). 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 [6]). The writhe depends on the choice of diagram (see [6]).*

C. Unknotting operations

There is an operation to exchange the over arc and the under arc, called an *unknotting operation* (see Figure 7). For every non-trivial knot K, it is

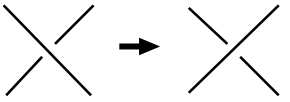

Fig. 7. Unknotting operation in mathematical topology.

modified into a trivial knot by applying a finite number of unknotting operations (see Proposition 4.4.1 in [6]).

III. DNA AS A 2-COMPONENT LINK

For a linear DNA, the length of the double strands is much longer than its diameter and also in the eukaryotic cells, the ends of a linear segment of DNA are fixed at the nuclear matrix. Therefore, it is possible to view the double strand DNA as a special 2-component link.

A. DNA-link

A ds-DNA can be modeled as the boundary components $\{S_1, S_2\}$ of a long thin twisted strip with the centre curve 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. We call L a *DNA-link*. Let $L_0 = L_0(S_1, S_2; \gamma_0)$ 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 the original DNA. $\bar{S}_1$ and $\bar{S}_2$ represent synthesized single strands from 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.

From this observation, the authors described in [5] that the semi-conservative scheme is interpreted in terms of DNA-links as:

Lemma III.1 ([5])**.** *The semi-conservative scheme is interpreted in terms of links: a deformation of the DNA-link $L_0(S_1, S_2; \gamma)$ into a split link $\{S_1', S_2'\}$, where S_1' and S_2' are inherited single strands from S_1 and S_2 respectively.*

The complicated replication process can be interpreted in terms of DNA-links. As it is shown by Lemma III.1, the semi-conservative scheme is described as follows: the DNA-link L_0 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.

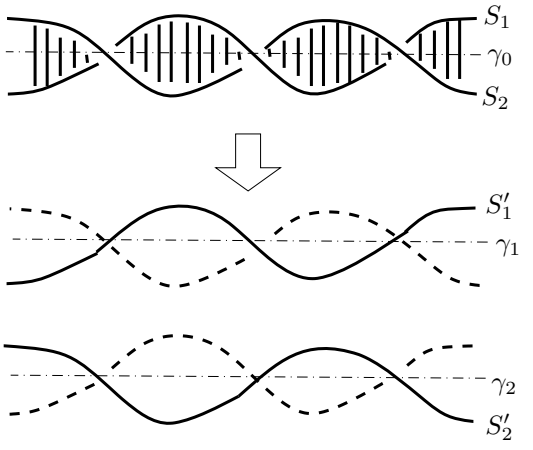

Fig. 8. A topological model of the semi-conservative scheme.

White proved in [13] the following formula of the linking number $\mathrm{lk}(S_1, S_2)$:

Lemma III.2 (White [13])**.** *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.

Lemma III.3. *Suppose that a DNA-link $L(S_1, S_2; \gamma)$ has unknotted (trivial) γ. Then $\mathrm{lk}(S_1, S_2) = 0$ if and only if $L(S_1, S_2; \gamma)$ is split.*

Proof: If $L(S_1, S_2; \gamma)$ is split, then $\mathrm{lk}(S_1, S_2) = 0$.

Suppose $\mathrm{lk}(S_1, S_2) = 0$. Consider a diagram D_L of L. Since γ is trivial, we can modify γ into a trivial circle γ' so that $\mathrm{Wr}(\gamma') = 0$. Let $L(S_1', S_2'; \gamma)$ be the modified link. Note that this modification does not change the linking number. Then

$$\begin{aligned} \mathrm{lk}(S_1', S_2') &= \mathrm{Tw}(S_1, S_2) + \mathrm{Wr}(\gamma) \\ &= \mathrm{Tw}(S_1', S_2') \\ &= 0 \end{aligned}$$

The diagram D_L is depicted in Figure 9.

Since the writhe $\mathrm{Wr}(\gamma') = 0$, the twisting number $\mathrm{Tw}(S_1, S_2)$ is cancelled by the twistings obtained from the writhe $\mathrm{Wr}(\gamma)$. Thus the diagram gives split trivial circles S_1' and S_2'. ■

B. Efficient unknotting operations

If an unknotting operation is applied to one crossing of the DNA-link $L(S_1, S_2; \gamma)$ with $\mathrm{Tw}(S_1, S_2) = m$, then the resulting link has the number of twistings $m-1$. Consider the following operation.

(*) To change consecutive k crossings of a DNA-link with $2m$ crossings, where $m = \mathrm{Tw}(S_1, S_2)$.

This operation eliminates $2k$ crossings (see Figure 10).

When the DNA-link is modified and moved so that it is partially rewound, some energy is used. The operation is expressed by a finite number of Reidemeister moves (see Figure 4). Thus it is possible that the necessary energy can be indicated by the number of steps consisting of those local moves. Of course, the number of local moves depends on the diagram. As these moves always occur around the center curve of ds-DNA, we can assume that the link L is linearly projected onto the plane can be used for describing this operation. Under this situation, the Reidemeister moves of $R2$ are used as shown in Figure 10. Once some $R2$ are applied along the center curve, count this as one step.

Then we have the following theorem.

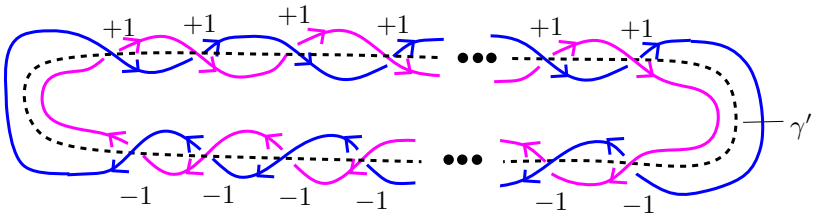

Fig. 9. Since $\mathrm{lk}(S_1', S_2') = 0$ and $\mathrm{Wr}(\gamma') = 0$, $\mathrm{Tw}(S_1, S_2)$ is cancelled by the opposite twistings induced from the writhe. The negative twistings come from the writhe $\mathrm{Wr}(\gamma)$.

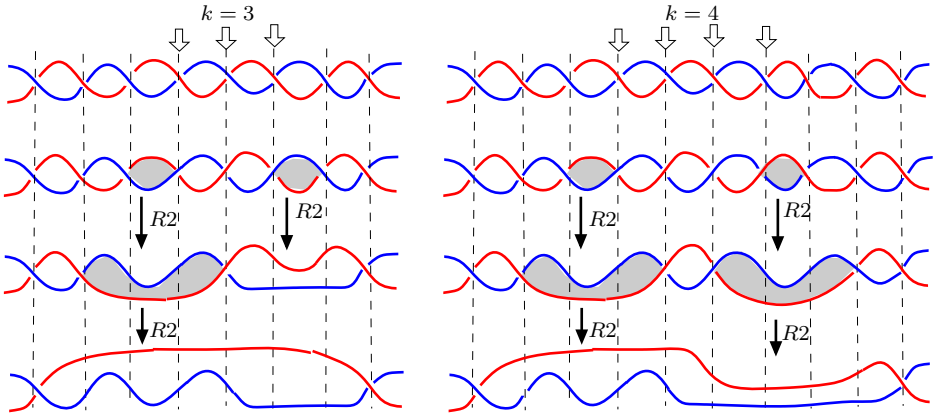

Fig. 10. The white arrows indicate the places where unknotting operations are applied; $k = 3$ (left) and $k = 4$ (right). The black arrows indicate the places that the Reidemeister moves $R2$ are applied. The both cases need 2 steps.

Theorem III.1. *For a linear DNA-link, applying unknotting operations to 2 consecutive crossings requires the minimal energy and implies the maximal length of rewound segment of the link.*

Proof: Let $\mu(k)$ denote the number of the steps of the deformation which is a function of k, expressed by

$$\mu(k) = \left\lceil \frac{k}{2} \right\rceil, \tag{4}$$

where the function $\lceil\ \rceil$ is the ceiling function. On the other hand, the outcome of the operation is indicated as the length of the rewound segment of the DNA. This can be counted as the number of crossings eliminated by the operation. Let $\nu(k)$ be the number of the eliminated crossings, expressed by

$$\nu(k) = 2k \tag{5}$$

An efficient operation should minimize μ and maximize ν. In order to see the efficiency of the operation, take the ratio:

$$\frac{\nu}{\mu} \tag{6}$$

It is easy to see that this ratio is 2 or 4. Therefore, it can be justified to take the values $\mu = 1$ and $\nu = 4$ to make the operation efficient. This means $k = 2$.

This is the required result. ■

IV. A REWINDING MODEL FOR REPLICONS

It is believed that the replication process starts from a specific domain called a replicon which has a looped shape (see Figure 2 (a)). The replicon (loop) is modeled by the 2-component tangle (link) shown in Figure 2 (c). We described a topological model for DNA replication in [5] but it requires

some over twist absorbing mechanisms which are unknown. If the process described in this section is used in the replication model, then the over twisting absorbing mechanism is not needed. The model in this section can be a solution for the problem unsolved in [5].

A. Estimate of the writhe

From Lemmas III.1, III.2 and III.3, in order to split the DNA-link $L_0(S_1, S_2; \gamma_0)$, it is necessary to eliminate the linking number $\mathrm{lk}(S_1, S_2)$. The linking number can be estimated from the length of the DNA segment of the replicon. Let ℓ_{γ_0} be the length of γ_0 in terms of base pairs and let b be the number of base pairs within one full twist of the double strands; b is approximately 10.5. Then the linking number is given by:

$$\mathrm{lk}(S_1, S_2) = \mathrm{Tw}(S_1, S_2) + \mathrm{Wr}(\gamma_0) \tag{7}$$

$$\approx \frac{\ell_{\gamma_0}}{b} + \mathrm{Wr}(\gamma_0) \tag{8}$$

$$\approx \frac{\ell_{\gamma_0}}{10.5} + \mathrm{Wr}(\gamma_0). \tag{9}$$

The DNA in eukaryotic cell has a form called *nucleosome*, where the DNA is wrapped into two negative supercoils around a histone octamer (see [11] for details). The DNA string around the histone octamer forms a double loop with $\mathrm{Wr} = -2$ (see Figure 11). The total writhe in the replicon

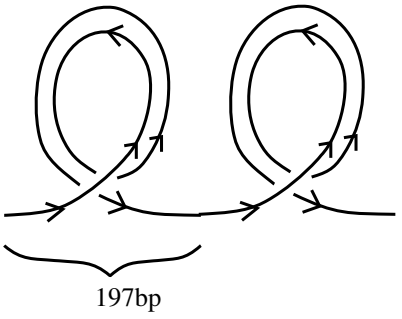

Fig. 11. The nucleosome has the length about 197bpand its writhe -2.

induced from nucleosomes in it is calculated as follows. Each nucleosome has the writhe -2 with 147 bp and a linear piece of DNA called a *linker* which has 50 bp (see [11] for details). Therefore, we may assume that a nucleosome appears at least every 197 base pairs. If we compare the number of full twists within the 197 base pairs:

$$\frac{197}{10.5} = 18.76 \quad \text{(full twists.)} \tag{10}$$

The ratio between the writhe of a nucleosome and the number of full twists of the nucleosome is:

$$\frac{|\mathrm{Wr}|}{\mathrm{Tw}} \approx \frac{2}{18.76} \approx 0.107. \tag{11}$$

This means that the ratio of the number of fully twisted supercoils is formed by nucleosomes; that is, the negative writhe $\mathrm{Wr}(\gamma)$ induced from the nucleosomes is at most about 10.7%.

Therefore, by Lemma III.2, the equation $\mathrm{lk}(S_1, S_2) = \mathrm{Tw}(S_1, S_2) + \mathrm{Wr}(\gamma)$ implies that about 90% of $\mathrm{Tw}(S_1, S_2)$ should be reduced by some operations.

For instance, it is estimated in [7] that the length of the replicon (loop) is about 100Kbp. Then the number of full-twists in one replicon (loop) is given by:

$$\frac{100\mathrm{Kbp}}{10.5\mathrm{bp}} \approx 9.5\mathrm{K}. \tag{12}$$

The length for each nucleosome is about 197bp. Suppose that for every 197bp, there is one nucleosome. Then the number of nucleosomes is

$$100\mathrm{Kbp}/197\mathrm{bp} \approx 0.5\mathrm{K}. \tag{13}$$

Each nucleosome has the writhe -2. Therefore, $\mathrm{Wr}(\gamma_0)$ is estimated as $-1\mathrm{K}$.

This observation implies that about 8.5K of full-twisting should be reduced by some operations.

B. Recursive unknotting operations

The enzymes TopoIA and TopoII act as unknotting operators to single and double strands respectively.

Theorem III.1 suggests the following operation. TopoIA is activated at the adjacent two crossings as shown in Figure 13.

This modification releases exactly two and a half twists. This decreases the twisting number 2.5. This relaxation does not require the counter

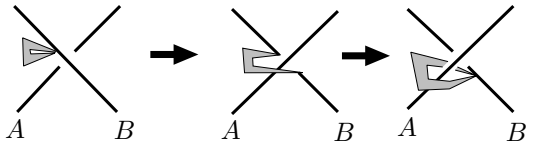

Fig. 12. TopoIA acts if A and B are single strings. TopoII acts if A and B are double strings.

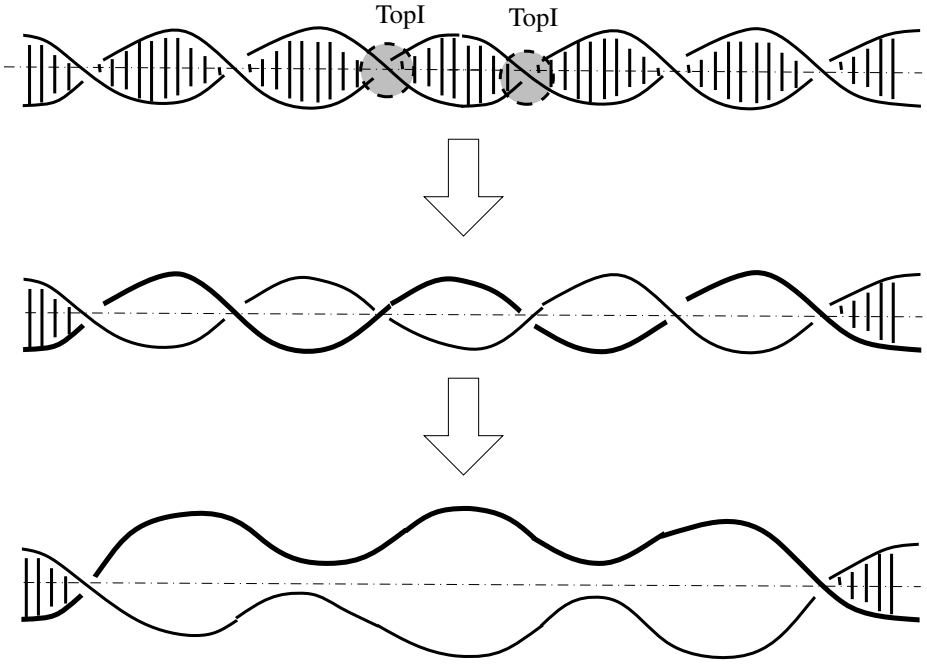

Fig. 13. The topological model of the rewinding at the replication origin. TopIA is activated at the shaded circles in the top diagram. This releases 2.5 full twists. This model keeps the writhe of the single strands. It does not require too many extra rotations.

rotation. This is different from the topological model proposed in [5].

The next problem to solve is to know how many times TopoIA should be applied to change the crossings during the replication in the replicon. From (12), the number of full twists in the replicon is about 9500 denoted by Tw_0. Thus the number of places on the replicon where the TopoIA is activated is:

$$\frac{9500}{2.5} = 3800 \tag{14}$$

The resulting relaxed DNA has now about 3800 crossings. For each replicon, about 3800 of activations of TopoIA will be needed to reduce the linking number. After activating TopoIA to one place, the linking number is reduced from 2.5 to 0.5. From this observation, if TopoIA is pairwisely activated at every 2.5 full-twists, the twistings will be reduced to 20% of the twistings Tw_0 of the replicon. It is denoted by Tw_1.

$$\mathrm{Tw}_1 = 0.5 \times 3800 \approx 1900$$

If TopoIA is applied again at every 2.5 full-twists of the relaxed DNA again, the twistings will be

reduced to 4% of Tw_0 denoted by Tw_2.

$$\mathrm{Tw}_2 = \frac{1900}{2.5} \times 0.5 = 380$$

This means that after applying the unknotting operation twice, only 380 twisings remain.

Therefore, the unknotting operation gives Tw_2 which is 4% of Tw_0. From Section IV-A, the writhe induced from the nucleosomes is at most 10% of Tw_0. The twisting number Tw_2 is less than half of 10%. Therefore, from the formula (3): and Lemma III.3, the link is very close to be split.

V. Conclusion

In this paper, we studied a possible topological model for a replicon equipped with unknotting operations. It is proved that the model is efficient to make the linking number of the replicon close to zero. However, it remains that the existence of this kind of mechanism is unknown. This should be examined through experiments.

References

[1] M. Barbi, J. Mozziconacci, H. Wong, and J. Victor, *DNA topology in chromosomes: a quantitative survey andits physiological implication*, J. Math. Biol. **68** (2014), 145–179.

[2] A. Bates and A. Maxwell, *DNA topology*, Oxford University Press, 2005.

[3] A. Kornberg and T. A. Baker, *DNA replication, second edition*, University Science Book, 2005.

[4] S. D. Levene, C. Donahue, T. C. Boles, and N. R. Cozzarelli, *Analysis of the structure of dimeric DNA catenanes by electron microscopy*, Biophysical Journal **69** (1995), 1036–1045.

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

[6] K. Murasugi, *Knot theory and its applications*, Modern Birkhäuser Classics, Birkhäuser, 2008.

[7] H. Nakamura, T. Morita, and C. Sato, *Structural organizations of replicon domains during dna synthetic phase in the mammalian nucleus*, Exp. Cell Res. **165** (1986), 291.

[8] S. V. Razin, A. A. Gavrilov, E. S. Ioudinkova, and O. V. Iarovaia, *Communication of genome regulatory elements in a folded chromosome*, FEBS Letters **587** (2013), 1840–1847.

[9] J. C. Rivera-Mulia, R. Hernández-Munõz, F. Martínez, and A. Aranda-Anzaldo, *DNA moves sequentially towards the nuclear matrix during DNA replication in vivo*, BMC Cell Biology **12:3** (2011), 16 pages.

[10] V. V. Rybenkov, N. R. Cozzarelli, and A. V. Vologodskii, *Propability of DNA knotting and the effective diameter of the DNA double helix*, Natl. Acad. Sci. USA **90** (1993), 5307–5311.

[11] R. R. Sinden, *DNA structure and function*, Academic Press, 1994.

[12] A. Vologodskii, *Bridged DNA circles: A new model system to study DNA topology*, Macromolecules **45** (2012), 4333–4336.

[13] J. H. White, *Self-linking and gauss integral in higher dimensions*, Amer. J. of Math. **91** (1969), 693–728.

[14] R. H. C. Wilson and D. Coverley, *Relationship between DNA replication and the nuclear matrix*, Gens to Cells **18** (2013), 17–31.