REPUBLIC OF TURKEY
ERCİYES UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF MATHEMATICS AND COMPUTER

TOPOLOGY AND GEOMETRY OF DNA

Prepared by
Havva Nur KARABAY

Supervisor
Assoc. Prof. Nazmiye ALEMDAR

Master Thesis

October 2022
KAYSERİ

REPUBLIC OF TURKEY
ERCİYES UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF MATHEMATICS AND COMPUTER

TOPOLOGY AND GEOMETRY OF DNA

(Master Thesis)

Prepared by
Havva Nur KARABAY

Supervisor
Assoc. Prof. Nazmiye ALEMDAR

October 2022
KAYSERİ

SCIENTIFIC ETHICS CONFORMITY

I declare that all information in this thesis was obtained in accordance with academic and ethical rules. All results and materials which are not at the essence of this thesis were cited and the related references were provided by the academic and ethical rules.

Havva Nur KARABAY

SUITABILITY FOR THE INSTRUCTION GUIDE

The master thesis entitled **'Topology and geometry of DNA'** has been prepared in accordance with Erciyes University Graduate School of Natural and Applied Sciences Institute Thesis Preparation and Writing Guide.

Prepared by
Havva Nur KARABAY

Supervisor
Assoc. Prof. Nazmiye ALEMDAR

Chairman of Department of Mathematics
Prof. Dr. Hikmet ÖZARSLAN

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor Assoc. Prof. Nazmiye ALEMDAR whose door to office was always open whenever I ran into a trouble spot or had a question about my research or writing. She consistently allowed this paper to be my own work and steered me in the right direction whenever she thought I needed it.

I would also like to thank dearest departed friend of mine, Pakize Uslu, with whom I always dreamt about doing master degree by motivating each other. I am gratefully indebted to her for her very valuable motivation and want to dedicate this thesis in the memory of her.

Finally, I must express my very profound gratitude to my husband for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without him.

Havva Nur KARABAY

October 2022, KAYSERİ

TOPOLOGY AND GEOMETRY OF DNA

Havva Nur KARABAY

Erciyes University, Graduate School of Natural and Applied Sciences

Master Thesis, October, 2022

Supervisor: Assoc. Prof. Nazmiye ALEMDAR

ABSTRACT

This thesis consists of 7 Chapters. In Chapter 1; the definitions of Knot Theory and basic knot moves were given. The core of Chapter 2 is about the structure of DNA and knots on a more biochemical context schemed by DNA and the topological structure of DNA. In Chapter 3; the modeling of DNA in terms of being a topological ribbon, Serret-Frenet framing, and the presentation of principal connection Lk = Tw+Wr is dealt with. Chapter 4 emphasizes on the significance of DNA supercoiling biologically and compactification and the topogical structure of DNA together with the topoisomerases. In Chapter 5; a discourse on the alteration to DNA topology. In chapter 6; a concise sketch of the tangle model and 4-plats are provided.

Keywords: Topology, Topology and Geometry of DNA, Knots, Knot Theory,4-plats

DNA'NIN TOPOLOJİ VE GEOMETRİSİ

Havva Nur KARABAY

Erciyes Universitesi, Fen Bilimleri Enstitüsü
Yüksek Lisans Tezi, Ekim 2022
Danışman: Doçent Dr. Nazmiye ALEMDAR

ÖZET

Bu tez 7 Bölümden oluşmaktadır. 1. Bölümde; Düğüm Teorisi ve temel düğüm hareketlerinin tanımları verildi. Bölüm 2'nin özü, DNA ve DNA'nın topolojik yapısı tarafından düzenlenen daha biyokimyasal bir bağlamdaki DNA ve düğümler hakkındadır. 3. Bölümde; DNA'nın topolojik şerit olması bakımından modellenmesi, Serret-Frenet çerçevelemesi ve Lk = Tw+Wr ana bağlantısının sunumu ele alınmaktadır. Bölüm 4, DNA'nın biyolojik olarak süper sarmallanmasının ve kompaktlaştırmanın önemi ve topoizomerazlarla birlikte DNA'nın topolojik yapısı üzerinde durur. 5. Bölümde; DNA topolojisindeki değişiklik üzerine bir söylevi içerir. 6. bölümde; arapsaçı modelinin kısa bir taslağı ve 4 plat sağlanmıştır.

Anahtar Kelimeler: Topoloji, DNA'nın Topolojisi ve Geometrisi, Düğümler, Düğüm Teorisi, 4-plats

TABLE OF CONTENTS

TOPOLOGY AND GEOMETRY OF DNA

LIST OF FIGURES

INTRODUCTION

Knot theory is a theory for the investigation of three-dimensional closed curves in mathematics. No matter how easy it is to name a knot as the closed curve formed by connecting the two ends of a rope in mathematical terms, its theory, namely knot theory, is not as unsophisticated as that, since many studies have been carried out and articles have been written over the last decade about it. Chemical discussions, on the other hand, have always driven theorists who have put forward knot theory since ancient times, as well. There is an indispensable and as much connection between mathematics and knot theory, since mathematics, just as in other fields, is used to find answers to questions to arise from knot theory. Knot theory has existed within a variety of fields from genetics to quantum mechanics. As it is well known that science has the secrets of its own and that knots are just one of those attempts to have been implemented to solve them.

What can be regarded as the most basic question within the knot theory is nothing but whether two knots are identical or not. Knots are generally viewed as same by mathematicians so that both look just like one another, in case of a deformation, damage, stretch or bend. The evidence behind two knots' being identical stems from the one which becomes just like the other even after being deformed in mathematical sense. Therefore, proving two knots to be disparate is not possible, since deformations are everlasting.

Knots turned out to be more of a study of mathematics than aesthetics with a significant revelation made by Johann Carl Friedrich Gauss (1777 – 1855), the prodigious mathematician. At the present time, we know that 'intertwinings are numerical invariants and that these invariants pass through each other thanks to Gauss who discovered how many intertwinings integral can calculate in 1883.

Gauss used to confer knots with Johann Benedict Listing (1808 – 1882), one of his postgraduate students. Listing, later on, invented the term topology by merging the words

'topos', namely 'place', and logos, namely 'logic' of Greek origin on the purpose of construing the location of geometry. Both Gauss and Listing were inquisitive on knots, nevertheless more of strides were necessarily to be made by someone who shared the same enthusiasm about knots.

This curious mathematician and physicist was no one but William Thomson who was inconsistently self-reliant. While as a student at the University of Cambridge, he was so convinced that he was going to be the top mathematician. He had a wide variety of interests.

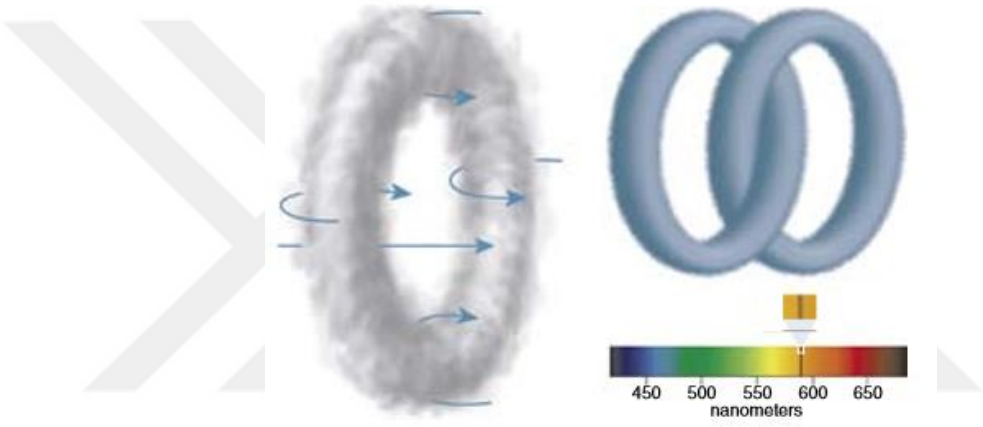

Figure 1. According to William Thomson demeanors of atoms can be explained as long as they are ether's knots transmitted by vortex motion. The arrows on the left around the smoke circle illustrate the steering of the vortex motion. The exterior piece of the circle is thrust into the bigger surface area of the circle by air, thus maneuvering the circle in the steering of the air which runs through interior of the circle.

Three Scottish physicists, Peter Guthrie Tait, William Thomson and James Clerk Maxwell, were of the opinion that the application of knot theory into the main questions arising from many of the fields of the science would give answers to questions about the universe in the 19th century. William Thomson and Peter Guthrie Tait, also had a conviction that the existence and functioning of atoms could well be described by chemical elements, which, they believed, were merely ether, provided that the current condition of physics have to be implemented. James Clerk Maxwell is another Scottish scientist and he emboldened Thomson and Tait for the advancements in vortex theory of the atom. The vortex theory of the atom was kind of futile, since in the end it failed. It,

concisely, was based on the endeavor to describe as to why atoms came in large numbers though in small diversity. Yet, its contributions to mathematics are undeniable, since the endeavor to use knot theory to the main queries regarding science, the structure of matter is an outstanding anecdote.

James Clerk Maxwell was also interested in knots and topology and he expounded his views by putting forward that linking number of two knots had a physical significance. A magnetic field, he explained, is created when electrical current is dashed through one knot. Molecule which is charged follows the route by the other knot thereby creating the linking number. Therefore, according to Maxwell the linking number is a double integral which had already been discovered by Gauss.

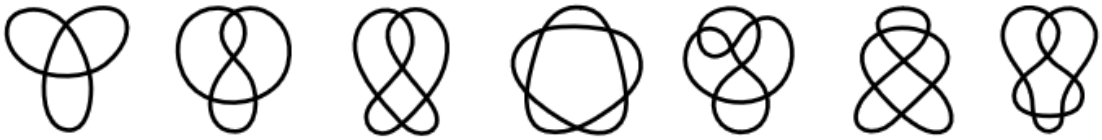

Figure 2. Tait made a categorization of knots according to the number of crossings each one has so as to create a table of elements. Above are seven knots which stretch through six crossings. Ultimately, Tait was able to construct tables including ten-crossing knots. Today crossings count up to 16, including 1.6 million entries.

Tait's eagerness together with Thomson's assurance towards knots has led to the creation of the first knot theory research program. Maxwell's who kept on providing inputs, on the other hand, were also indispensable to the knot theory. Today knot theory is still studied; however, vortex-atom theory ended.

The first endeavor in terms of physical operation of knot theory was performed by the vortex-atom theory. From that time onwards, knot theory procedures have proceeded from fluid dynamics to DNA research and calculation of quantum. Still the physical importance of knots is yet to be understood. Obvious as it may seem, some more years

may pass until the fully comprehension of knots and links. Deep physical applications will bloom when the profound essence of knots are fully grasped.

In Figure 3 below, the levitation of a circle over Mount Etna, which is not a dormant volcano, can be seen. Tiny ejections from volcanic fissures, reaching over 600 feet in diameter, lead to such circles as these and they can remain steady over 10 minutes. This phenomenon attracted the attention of the Scottish physicist Peter Gıthrie Tait who wanted to elaborate in vortex motion so that he made experiments with these smoke circles in a tiny way. Consequently, he comprehended that the most basic patterns of substance may have been steady and that they were merely stable knots of ether equipped with vortex motion. The ventures Tait advanced knot theory, which is the most effective field of mathematics.

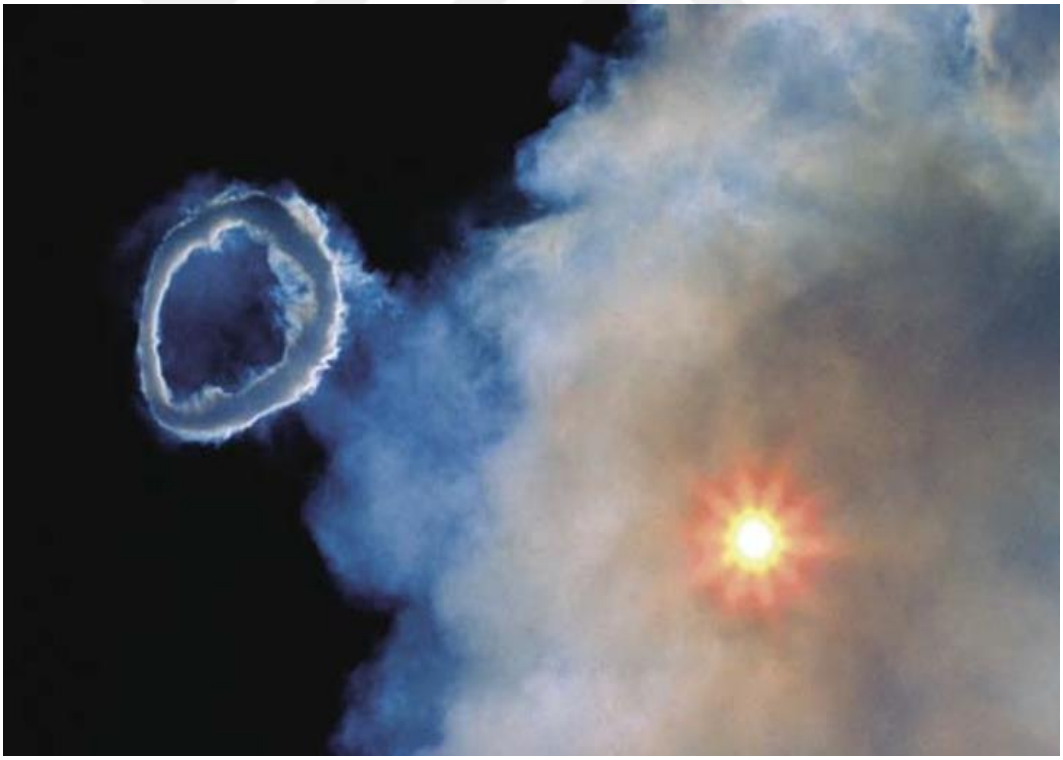

Figure 3. The levitation of a circle over Mount Etna.

The sphere of this thesis is but not limited to the DNA and its topological characteristics. Commencing with the history and origin of knot theory, which greatly helps in comprehending the characteristics, structure of DNA and how recombinases unknot DNA was studied. The connections between DNA and knot theory was also set and explained. DNA has a helical form and how this form has specific connections to the molecule, its geometry and topology has been drafted. The topology and geometry of DNA is adjusted by proteins, the comprehension of which leads to mathematical inquiries. Both the review of these mathematical inquiries and the proteins whose main responsibility is to adjust the geometry and topology of DNA is also what has been worked on. Conservative site-specific recombinations which are proteins altering DNA topology is also a part of our study. A topological model to anticipate the kinds of knots and links emerging from conservative site-specific recombination is also discussed. In the end, the knot model, including exemplifications of how it was beneficial in comprehending the system of a conservative site-specific recombination is delved into.

In spite of the fact that the impetus to the comprehension of knots vanished, biologists and chemists have started using knot theory anew. DNA which was linked and knotted was first discovered amid 1980s and had profound function in terms of investigating the outcomes according to topologists. This thesis overviews some portion of the area which is sketched within the current interdisciplinary of DNA Topology [75].

CHAPTER 1

BASIC CONCEPTS

1.1 Defınatıon Of Knot Theory

Definition 1.1.1. ‘Knot’ is defined as a closed curve.

A knot is a closed, non-self-intersecting curve contained in three dimensions that cannot be untangled to form an uncomplicated loop, according to mathematics (i.e., the unknot). While one or more strands might be left open on each side of a knot in ordinary usage, the mathematical theory of knots attributes to such an item as a "braid" instead of a knot. A knot is only considered a knot by a mathematician on condition that the free ends of the object are connected in some way, leading to in a single looped strand.

Let the 3-dimensional sphere be denoted by $S^3 = R^3 \cup \{\infty\}$. Any cluster that is topologically homogeneous (homeomorphic) with the circle $S^1 = \{(x,y,z): x^2 + y^2 = 1, z = 0 \text{ in } S^3\}$ is called a knot. So the knot is a simple closed curve in space. In other words, the knot is the position of the unit circle in space.

Definition 1.1.2.

Linking Number: The ‘linking number’ is a numerical invariant that describes the connection of two closed curves in three-dimensional space. Linking number is hereinafter Lk. A link, on the other hand, could be a finite, organized collection of knots that don't meet one another.

Figure 4. Linking number

As can be seen from the figure below, there are two probable schemes. In the scheme a, +1 oriented sign, that is positive crossing, in the scheme b, -1 oriented sign, that is negative crossing, has been shown.

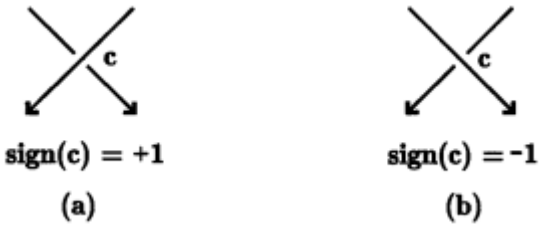

Figure 5. Crossings

Let's assume that, D is an oriented regular diagram of a 2 component link L = $\{K_1, K_2\}$ The projections of K_1 and K_2 extend from $c_1, c_2, \dots, c_m$.

$$\frac{1}{2}\{sign(c_1) + sign(c_2) + \cdots sign(c_m)\}$$

and this is what is called as the linking number.

Theorem 1.1.1. The linking number (K_1, K_2) is an invariant of L In other words, if D' is oriented diagram of L, it can be said that the value of the linking number is the same as that of D. Moreover, the linking number is free of the order of K_1 *and* K_2.

Definition 1.1.3. Crossing Number

Within the sphere of knot theory crossing number can be defined as the lowest number of crossings in any sketch for the knot.

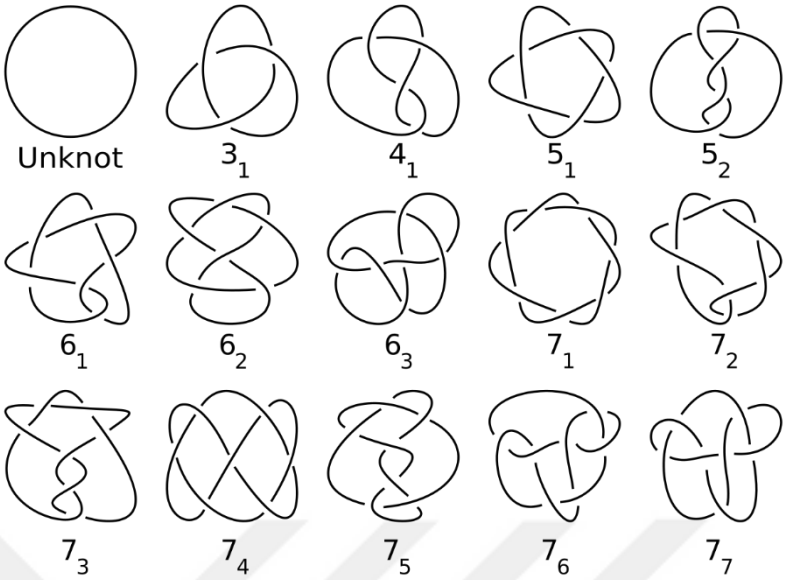

Figure 6. Crossing numbers

Definition 1.1.4. Writhe

The 'writhe' is the sum of positive and negative crossings of an oriented diagram. Writhe is hereinafter 'Wr'.

Wr = $\sum_{crossing}$

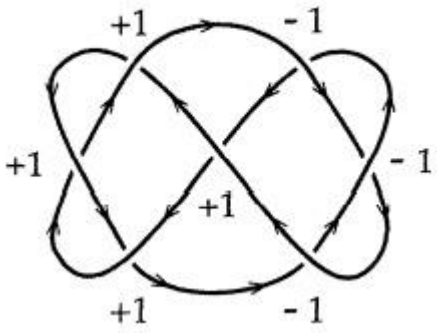

Figure 7. wr = +4- 3 = 1

Definition 1.1.5. Twist

'Twist' is the rate at which a straight strip rotates around the space curve, Twist is hereinafter 'Tw'.

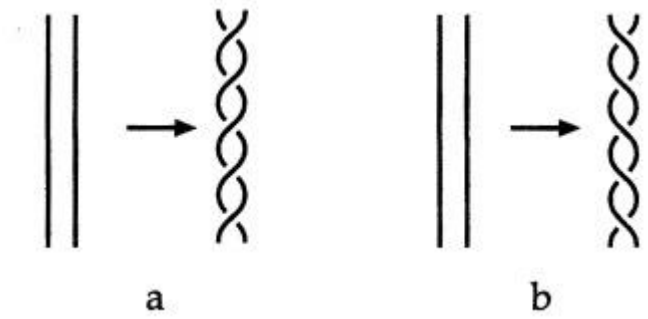

Figure 8. (a) A 5-move. (b) A 5-move.

1.2. The Basic Knot Moves

If the shape of a knot is thought as a polygon, it can be said that it is formed from the finite number of edge.

In order to describe knot mathematically, a change in the shape of the knot is required. For example, let's replace one edge AB with 2 new edges AC, CB at a K knot. This displacement process is called simple knot movements.

Definition 1.2.1. Below are four of the operations which can be carried out on a provided knot K**.**

(1) We may separate an edge, AB, in space of K into two edges, AC, CB, by setting a point C on the edge AB.

(1)' [The inverting of (1)] When AC and CB are two bordering edges of K and if C is eradicated AB gets to be a straight line, at that point we might discard the point C.

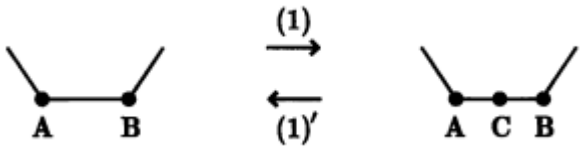

Figure 9.1. The Basic Knot moves

(2) Assume C may be a point in space that does not depend on K. In the event that the triangle ABC, shaped by AB and C, does not cross K, with the exemption of the edge AB, at that point we might discard AB and include the two edges AC and CB.

(2)' [The inverting of (2)] In case there is a triangle ABC in space involving two adjoining edges AC and CB of K, and this triangle does not meet K, but at the edges AC and CB, at that point we might discard the two edges AC, CB and include the edge AB.

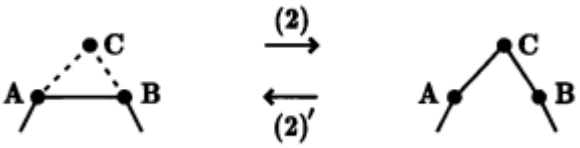

Figure 9.2.

(1), (1)', (2) and (2)' are called as the basic knot moves. [In any case, since (1) and (1)' are not "moves" in the usual understanding of this word, frequently as it were (2) and (2)' are referred to as basic knot moves.

1.3. The Equalness Of Knots

Knot won't change mathematically if basic knot movement is applied to a knot. However, if this move is applied too much in different parts and more than once, then the exurgent knot will look like a different knot. Knots in the figure are paired knots. These knots are called Perko's knot pair.

For instance, the two knots K_1 and K_2 in Figure 10 should be looked at, which might be known as Perko's pair.

Figure 10. Perko's Pair of Knots

Perko's pair of knots seems totally distinctive from the outer appearance. Actually, it hasn't occurred to anyone that the otherwise could be possible for the past 100 years. Nevertheless, if elementary knot moves are carried out in finite number of times, it is probable to alter the knot K_1 into the knot K_2.

The kinds of Knots which can be altered in between each other by administering the basic knot motion are declared to be equivalent or equal. Thus, it makes the two knots in Figure 10 equivalent.

A knot K is asserted to be equivalent (or equal to) to a knot K' on condition that we are able get K' from K by administering the basic knot motion a restricted number of times.

In the event that K is equal to K' , at that point, because K' is additionally comparable to K, it can well be said that the two knots K and K' are equivalent (or equal). We shall denote this equity by K ≈ K'.

A knot has no beginning point and no endpoint, i.e., it may be a basic closed curve (to be exact; a closed polygonal bend). In this manner, we will allot an direction to the bend. As is the tradition, we might indicate the direction of a knot by an indicator on the bend. It is instantly apparent that any knot has two conceivable directions, Figure 11.

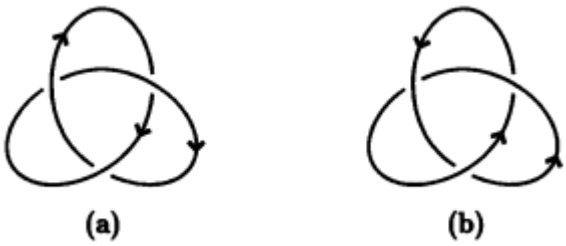

Figure 11. Directions of Knots-1

If oriented basic knot moves can change K and K' as shown in Figure 12, then it can be said that K and K' are equal thus K ≅ K' can be reached.

Figure 12. Directions of Knots-2

The two knots in Figure 10. are assuredly equivalent without a direction applied; nevertheles, it is not clear if they are equal or not with direction.

Definition 1.3.1. The two knots are equal with direction in Figure 11.

Two of the ties K_1 and K_2 can be said to be an identical, or K_1 is equal to K_2, in case there exists an direction-perpetuating homeomorphism of R^3 which outlines K_1 to K_2 [76].

CHAPTER 2

THE STRUCTURE AND TOPOLOGICAL EVALUATION OF DNA

2.1. The Formatıon Of Dna

Guidance within live organisms is provided by DNA which stands for Deoxyribonucleic Acid, the molecule (or a batch of them) whose task is to cipher all genetic information. All the innate qualities such as the determination of one's eye color and to what extent he or she is vulnerable to congenital sicknesses is ascertained by the information kept in DNA. That's why, DNA is reckoned as the 'sketch for life'.

The pattern of DNA was assembled together by Francis Crick and James Watson who built their work on the images of Rosalind Franklin. Nucleotides, each of which consist four bases (Adenine, Thymine, Guanine or Cytosine), are the units forming DNA.

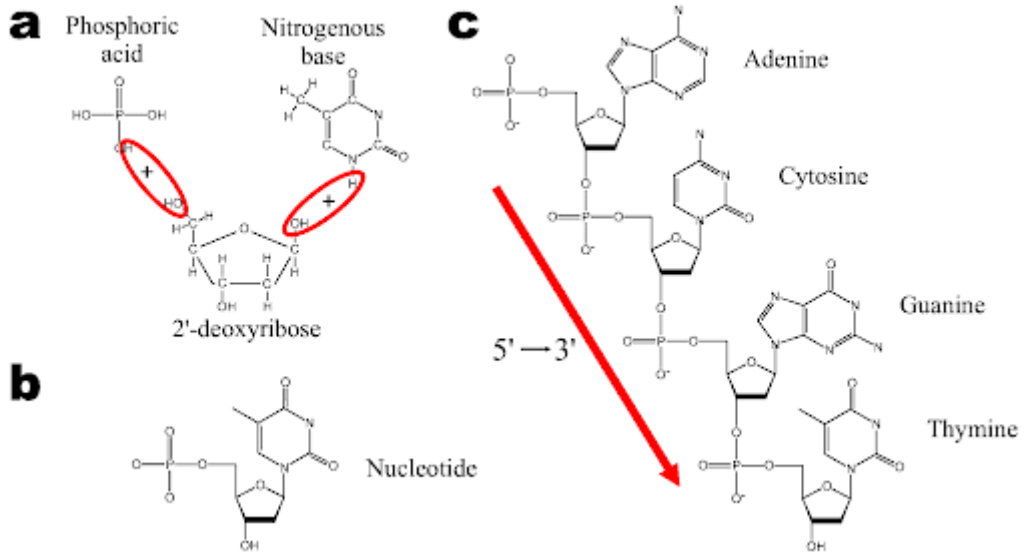

Figure 13. The structure of DNA

A DNA double helix, represented like a ladder, is modeled by these nucleotides. DNA, to put it in different way, is a double twisted molecule like a winding ladder. The exterior parts of the molecule are formed by deoxyribose sugars which are replaced with phosphates. “Backbone” is the word to refer this segment of the molecule from time to time. It should be taken into account that the twists often called the "backbone" exist inversely and indicated as “anti-parallel”. (This is presented by the 3-prime (3’) and 5-prime (5’**)** [1].

The nucleotide bases which are Cytosine, Guanine, Adenine, and Thymine form the inner part of the molecule resembled to the ‘steps’ of a ladder. C is linked to G via three hydrogen links whereas A is linked to T with two of them. The main information storing part of the molecule is these nucleotide bases which have disparities in the rerunning course of Cs, Gs, As and Ts, thereby making us one and only.

For one DNA twist to multiply and become two, it has to unfasten onward the hydrogen links. [2].

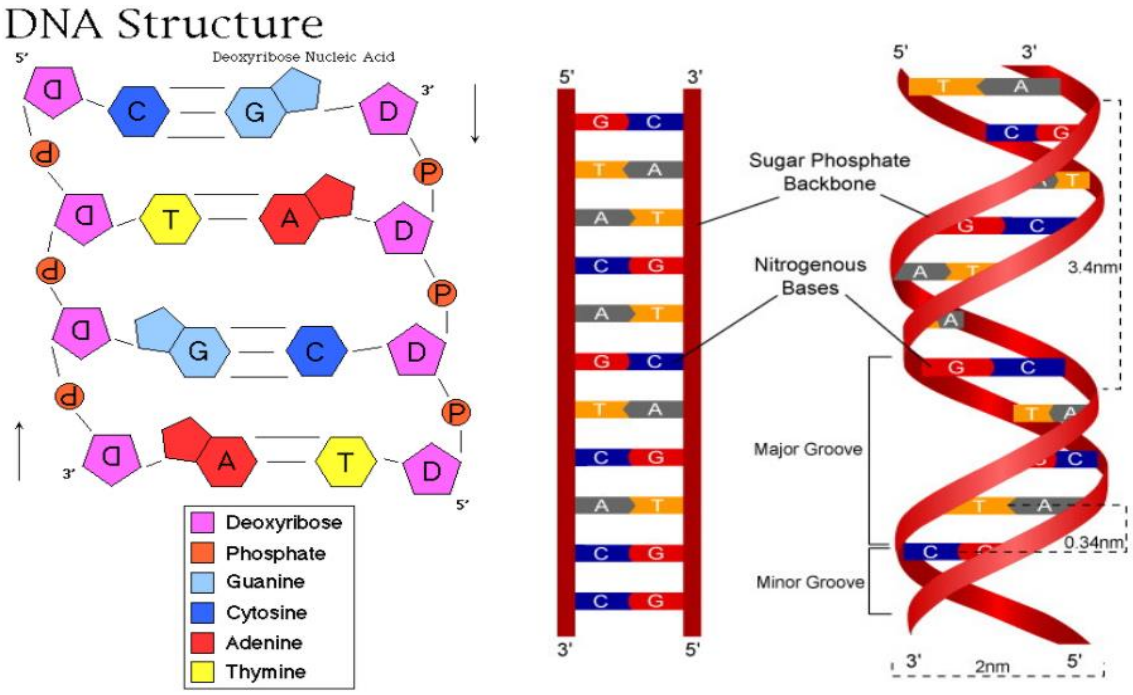

Figure 14. The prime and subsidiary structure of DNA and the pairing of bases within DNA

Tertiary structure of DNA through which the axis of DNA can just as well be twisted in space, thereby causing ‘supercoiled’ DNA to exist was discovered by Vinograd ten years after the discovery of double helix. [3].

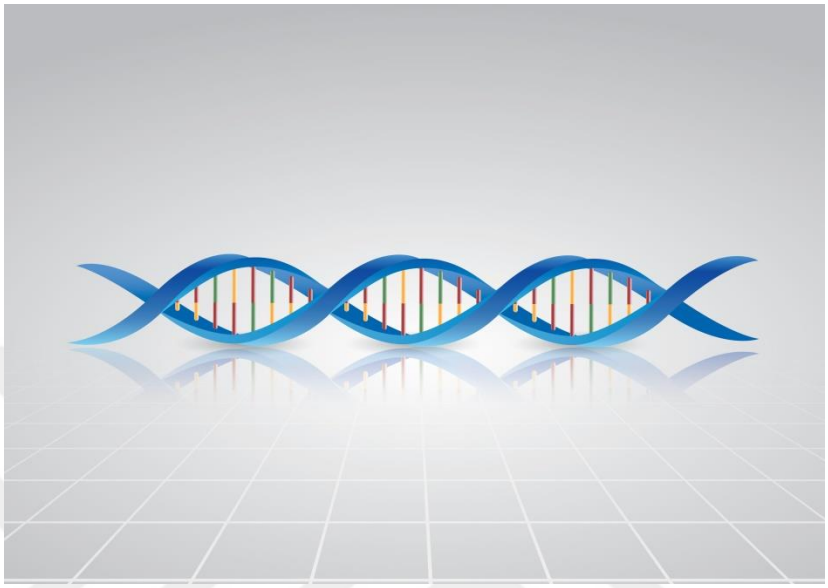

Figure 15. Linear DNA

Although linear DNA exists, its orbit is usually in a spheroid form which transpires when the covalent bond, namely the chemical bond, ties the edges of 2 backbone strands. DNA having a round orbit can be exemplified by bacterial genomic DNA, chloroplast DNA, and human mitochondrial DNA. Moreover, plasmid DNA and all that within biochemical labs have a round axis.

What might hold attention is the fact that even the orbit of DNA can be knotted or linked. We know that knots and links of DNA transpire amid the replication and recombination (copying and rearranging, consecutively) of DNA. Nontrivial topology of DNA is caused by twisting of the round orbit, namely the double helix and supercoiling is to be accentuated.

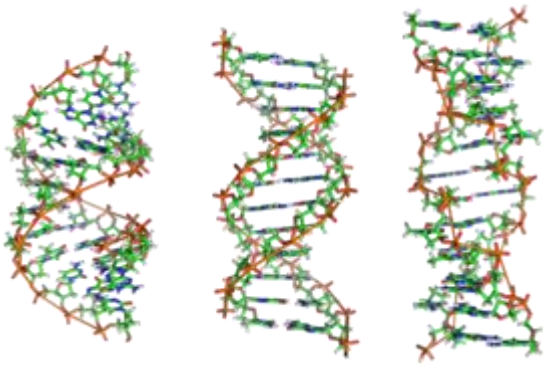

Figure 16. Tertiary structure of DNA

2.2. The Identification Of Dna Topology

The three-dimensional structure (shape) of proteins, DNA and RNA within the cell, and the affinity with the macromoleculer structure and its functional relationship are one of the most important things within the molecular biology. Normally, protein and the structure of DNA is detected by X-ray crystallography, electron and atomic force microscopy, and nuclear magnetic resonance imaging (NMR). The close packed structure necessary for crystallization, a preliminary preparation for electron or atomic forcescope and the absence of resolution of NMR does not provide a precise proof for the molecular shape within the solution. Also, some proteins (enzymes) act like molecular machines and shift in change while fulfilling its function thereby preventing and extra single static spatial snapshot to account for all the process. Topology can elucidate this significant topic. To be able to figure out how enzyme operates and what its structure is, topology and geometry together with its expressive and analytical powers are utilized, which actually is empirical and the one that is a topological way to enzymology (Fig. 17) [77].

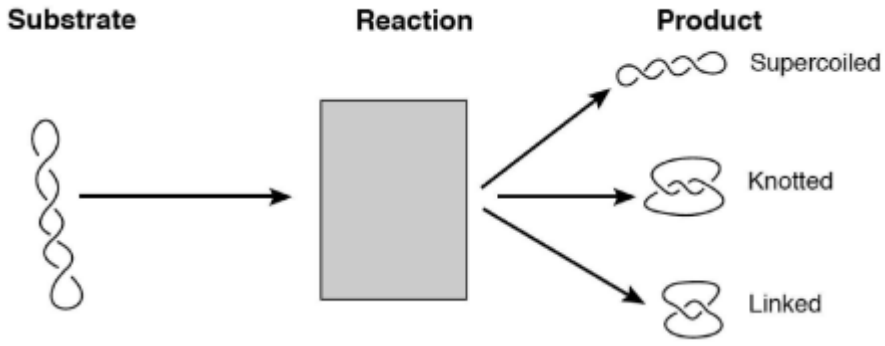

Figure 17. The Topological Approach to Enzymology

Alterations within DNA structure with the effect of enzymes necessiated the significant usage of geometry and topology in molecular biology. The usage of mathematics as an analytical means is significant since there is no empirical way to straightforwardly inspect the dynamics of enzymatic operation [77].

As discussed before, DNA knots and links have been associated in several cellular phases [18, 35]. They can take place as a consequence of replication and recombination to be discussed below. The consequence of the action of enzyme produces knots and links, as well. Besides the topoisomerases, to be discussed later, the reactions from recombinases and transposases [35, 36], which are the two other protein families, produce knots and links. Most commonly, DNA knots and links are the productions of specific laboratory experiments which are called the experiments of 'topological enzymology' on unnaturally created tiny (5–10 kb) round molecules. These knots and links can allow the relation and system of the protein to be determined which is being worked on. The fact that there are lots of DNA knots and links having been inspected made these molecules an urgent affair in terms of their biologically separation. [37, 38, 39, 40, 41, 42].

Liu and Davis [43] were the first to have discovered the knotted DNA in the lab in 1981. Empirically, there are two ways of dissolving DNA knots and links in two ways: either through electron microscopy or electrophoretic migration [44, 45, 46].

Electron microscopy makes it possible for DNA molecules to be visualized exquisitely. (See Figure below.) In this phase, the complete DNA molecule is covered in a RecA covering, which congeals and condenses the molecule. As a result, the sign of the axis crossings can be determined thereby certainly determine the exact knot or link kind. Yet, this phase can be arduous and difficult, especially while producing a comparably large amount of knotted/linked DNA and decoding the sign of crossings. Maybe, which was not so surprising at all then, there are only a couple of electron microscopy images of DNA knots and links which have been published by a few labs most noticeably that of Andrzej Stasiak who have developed the fundamental proficiency.

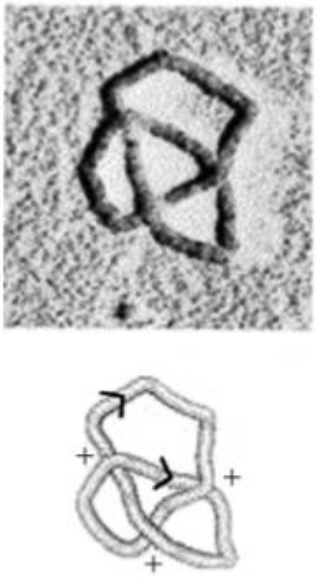

Figure 18. Electron microscope picture of knotted DNA.

Agarose gel electrophoresis is a common, though inadequate, method to be utilized to segregate DNA knot and link types. To do this, the DNA is first slit—that is, the strand of one backbone is cut to let any supercoils go. Next, the gel layer which is a substance resembling to jell-O derived from agarose, a seaweed derivative, is later soaked in water and a flow of water is run through it. After that, DNA is ducted into one end of the gel. If we remember from the previous sections, the DNA backbones are comprised of substituting sugar and phosphate groups. These phosphate groups which are negatively charged cause the DNA move to the positive end of the gel. Gel electrophoresis will flake DNA knots and links when all the slit DNA which is ducted is of the equal molecular mass and sequence. When exposed to UV light, it looks like as if an array of black crew is advancing towards the white gel. Each crew equals to several DNA molecules, and its concentration is commensurate with the quantity of DNA available.

Gel electrophoresis is direct and necessitates comparatively tiny quantities of DNA. Ordinarily, the interval a provided knot or link moves through the gel is commensurate with the minimum crossing number (MCN standing for the slightest number of crossings by which it can be drawn). Under normal conditions, knots having bigger MCN move more swiftly when compared to those which have lesser MCN [47, 48, 49].

Yet, 1,701,936 knots with MCN ≤ 16 are available; therefore, a better lamination is required to positively describe a specific knot [50]. 2-dimensional gel electrophoresis can segregate some main knots with the same MCN [51] by cunningly letting the DNA flow in a second dimension. Pitiably, in the second dimension, an apparent liaison between relative movements of knots with the same MCN does not exist.

In some circumstances, gel electrophoresis, when DNA is given the length of 1-dimension, gel electrophoresis can segregate some knots with the identical MCN. For instance, the five and seven crossing torus knots move more haltingly than the same five and seven crossing twist knots. This is not a sweeping assumption; nevertheless, current experiments in a persuading way show that knots and links may move linearly in terms of the mediocre crossing number of a specific convenience which is the *ideal configuration* (which are the conformations letting maximal radial expansion of a virtual tube placed around the knot. Initially, these configurations were calculated, yet current analytic work [54] has confirmed the presence and oneness of these resolutions of the knot or link [47, 53].

Gel electrophoresis is not complete in terms of the fact that adjoining bands detect just comparative MCN or approximate crossing number, not exact values, which make things more intricate. Therefore, let’s provide two DNA molecules A and B, when A flows further through the gel than B, then A has a higher MCN than B, yet a generalization of whether MCN(A) = MCN(B) + 1 or MCN(A) = MCN(B) + 2, etc. cannot promptly be made. When MCN(B) = 3, say, or 4 cannot be induced. If MCN(B) = 3, say, or 4 without additional examination can not be determined, as well.

As a result, gel electrophoresis and electron microscopy are frequently utilized at the same time. First DNA is sent through a gel to separate a knot with a provided MCN, and the specific band of interest is tangibly taken out of the gel. This DNA is additionally clarified, and then studied through electron microscopy to decide the exact knot or link type.

A control mechanism is to be set if deciding the certain knot or link type without utilizing electron microscopy is sought, which is possible through a knot staircase. This is a resolution of DNA molecules the exact MCN of which are known. Then, these molecules are ducted into an adjourning well to the unknown DNA within the gel. These known

knots or links later act as flags in the gel by which calibrating the DNA molecules is possible. Although this is possible in some circumstances (such as T4 topoisomerase which will generate a ladder of twist knots [55]), creating such a ladder of known knots and links from DNA of the identical length and similar sequence as the unknown knots is quite complicated.

Mathematicians together with experimentalists are interested in contemporary means for ascertaining (or determining) the exact DNA knot or link type, since not only the electron microscopy but also gel electrophoresis are challenging. In most cases, the modeling of DNA movement via a gel is an active field of investigation.

2.3. The Topological Formation Of Dna

The structure of DNA which is loose or unrestrained has 10.4 base pairs per turn of the helix. The DNA double helix can also be explained through a blending of twist and writhe whose total is reckoned as the linking number. The twist is the number of spiral tuns of one string around the other. Therefore, the measurement of the twist can be possible via calculating how many times one string covers totally the other.

For example, the loose linear DNA shown in the figure below, has three twists and zero writhe.

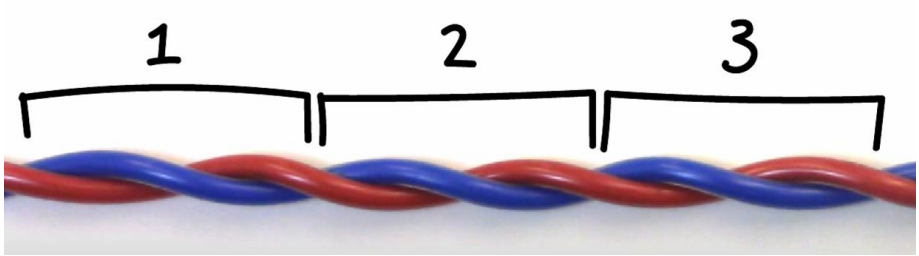

Figure 19. Twists

The explanation of the writhe, on the other hand, can be through the number of times the double helix passes over itself whether it is positive or negative relying upon the direction. For instance, -1 is the value that this circular DNA possesses and since it abides by the rule of right hand and then upper string moves left to right. Moreover, this DNA possesses a writhe of 1.

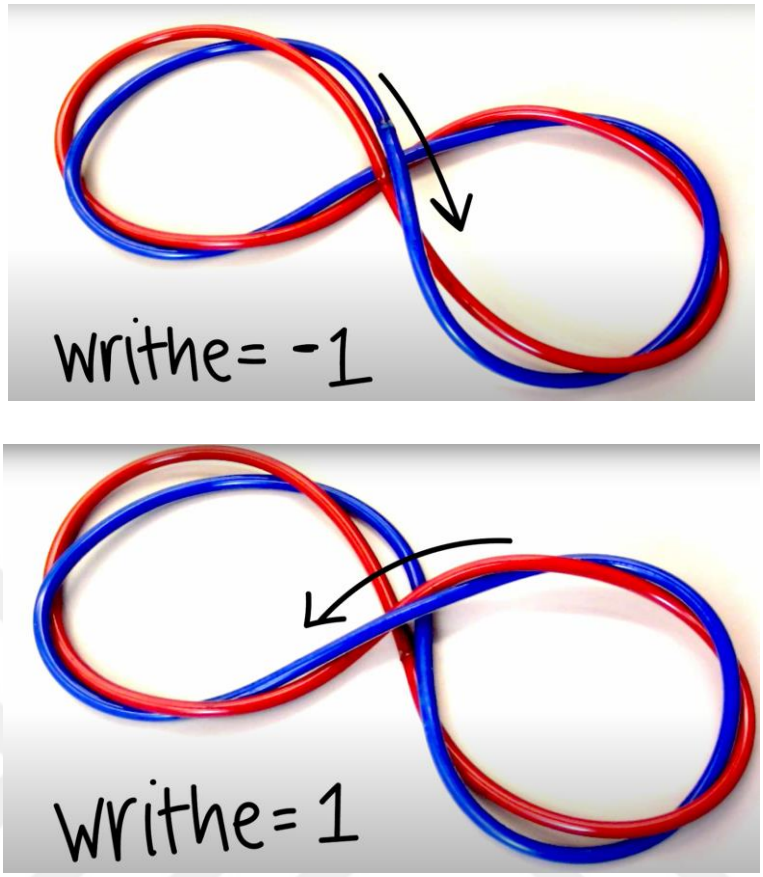

Figure 20. Writhe

Two of these instances possess twist, as well. In fact, linking numbers of such long DNA molecules could be in thousands.

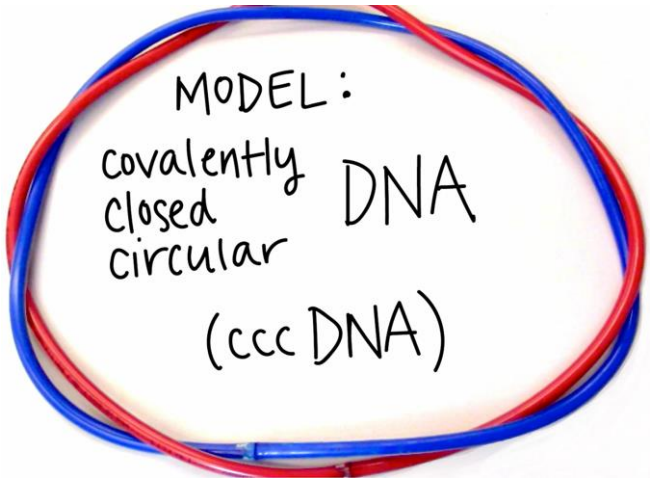

*Figure 21.*Topologically constrained DNA

The figure above is constricted in a topological way which is not alike linear DNA so that the alteration of the linking number of the DNA is not possible if one or two of the strings of DNA is not fissured.

The kind of DNA which is overcoiled possesses lower than 10.4 base pairs by spin of the helix. Therefore, overcoiled DNA is more tightly coiled than loosen DNA and possesses a far more number of twists and a higher linking number than loosen DNA. Once DNA is overcoiled and has an escalated linking number, the DNA turns into positively supercoiled and the double-stringed structure commences by covering positive writhe. Without taking these conspicuous alterations, it is a good idea to note and recall that positively overbound DNA will have more linking number than the loosen DNA, additionally the two strings of the double-string DNA are harder to detach.

DNA which is not wound enough possesses more than 10.4 base pairs by each twist of the helix. As a result, DNA which is more overcoiled possesses less wound than loosen DNA, thereby having decreased linking number. Once DNA is undercoiled and possesses a diminished linking number, the DNA becomes negatively supercoiled and the double-stringed structure begins covering around itself in the reverse orientation as positive supercoiling, thus forming negative writhe. Without taking these conspicuous alterations, it is a good idea to note and recall that negatively overbound DNA will have more linking number than the loosen DNA, additionally the two strings of the double-string DNA are easier to detach.

In brief, negative supercoiled, or undercoiled, thanks to DNA it is more effortless to disembody the double helix into two single strands. Since the double helix and negatively supercoiled DNA are disembodied, more twists are formed in the rest of the DNA, hence inducing rewinding of the underwound strands. Instead of relating strands on the DNA topology, the DNA which is still base paired is recurred to the ideal relaxed condition. Additionally, you have a piece of the disembodied sole DNA strands that can be used in the cellular progresses like reproduction and facsimile. Only if an endeveaour to disembody the double helix of relaxed DNA is made, would, then, the introduction of more twists in the DNA and end up of ovewrinding, or positively supercoiling, the double helix take place, which is adverse in terms of energy. [78]

2.4. The Study Of Dna And Its Topological Ribbon

From its earliest days, outlining supercoiled DNA necessitated a merge of mathematicians and biologists. The formulae of a basic connection for space curves was created by Călugăreanu [4] and White [5]: Lk = Tw+Wr. Twist and Write, the two geometrical sums whose quantities may alter when exposed to malformations by the curves, value to a topological quantity, known as Linking Number as an invariant when exposed to malformations as these [6].

In an endeavor to comprehend this connection, a molecule of circular DNA as a twisted ribbon should be designed. A curve in 3-space (the verge of the ribbon) is linked for each two backbones. Traditionally, backbones are called as C and W (thanks to Crick and Watson). The direction of these two ribbon edges correlate with each other, thereby being devoid of innate direction from the bottom chemical pattern.

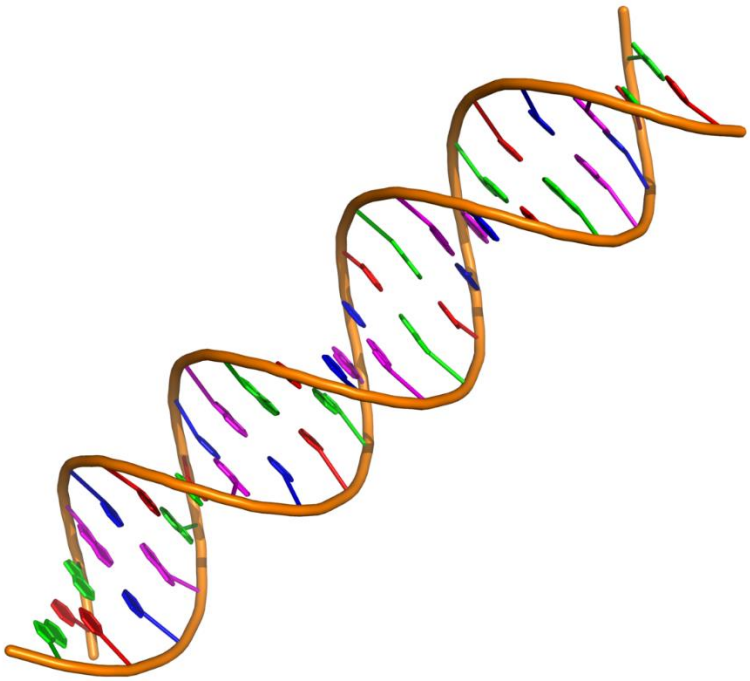

Figure 22. DNA as a topological ribbon. A reduced depiction of a double stranded DNA helix.

The curves which are called as C and W are intertwined with each other, thereby proving that the ribbon can be knotted on its own, too. Similarly, so long as more than one DNA molecule is taken into account, the association of ribbon in itself is possible. The

formalization of this is carried out with regard to the linking number, Lk. When a directed two-component link is taken into account, a ±1 is to be assigned for each crossing which relies on the steering of the two strands. Lk, the linking number, is the half of their sum:

$$Lk = \frac{1}{2} \sum_{crossings} \pm 1$$

As Adams discuss in his article, Lk has no boundaries with the specific projection and that it is a topological invariant of the enclosing of the ribbon. (Adams, 2008) [7].

The two DNA backbones – C and W consecutively – make up a (2,m)-torus link. It is possible to draft a (2,m)-torus knot or link, thus letting its crossings transpire in line with m (standing for either positive or negative) could be derived so all its intersectings take place in succession of m (positive or negative plectonemic) crossings. Given that m is odd, then T(2,m) is a knot and in cases when m is even, then T(2,m) is a link. There could be varied definitions of Lkas illustrated in Figure 10. Note that if m is odd, then T(2,m) is a knot and if m is even, then T(2,m) is a link. Lk can be defined in a variety of correspondent ways. [8]. One depiction of Lk is with regards to Gauss' integral:

$$Lk = \frac{1}{4} \int_c \int_w \frac{(x'-x)(dydz'-dzdy')+(y'-y)(dzdx'-dxdz')+(z'-z)(dxdy'-dydx')}{[(x'-x)^2+(y'-y)^2+(z'-z)]^{\frac{3}{2}}}$$

which denotes (x, y, z) spreading over C and (x', y', z') ranging over W.

CHAPTER 3

GEOMETRICAL APPROACH TO DNA

3.1. Frenet Framing

This is how we form the Frenet framing when a regular closed curve γ(s) is taken into account [74]. Let's hypothetically say that we have a simple closed curve γ(s), having been restricted by arc width s. Later, the tangent vector, denoted as T, has length of

1: ||T|| = ||γ̇ (s)|| = 1, which makes

γ̇ (s).γ̇(s) = 1,

so γ̈(s) is perpendicular to γ̇ (s), such as. γ̈(s) which is a normal vector. Here, γ̇ (s) ≠ 0 has to be by definition, since γ(s) is regular. For clearness, this and the regular vector, N, will be standardized as the unit vector:

$$N = \frac{\gamma\ddot{(}s)}{\|\gamma\ddot{(}s)\|}$$

and for all s as such $\ddot{\gamma}(s) \neq 0$. Then, the binormal

B := T × N

is taken into consideration.

Since T, N and B are orthogonal unit vectors, a ground is created by them thereby, being construed as a straight combining by any vector, especially each of the vectors $\dot{T}$, $\dot{N}$ and $\dot{B}$ can be construed like this. Considering

$\dot{T}$ = aT + κN + cB,

for some a, κ, c, later a = 0, for by over $\dot{T}$ is perpendicular to T. Moreover, because

$N = \frac{\dot{T}}{\|\dot{T}\|}$, later c = 0 and so $\kappa = \|\dot{T}\|$. (κ is frequently appointed the curve; at a defined point on γ(s))

Correspondingly,

$\dot{N} = -\kappa T + \tau B$,

in which τ is often named as the twist – or torsion –, an assessment of how nonplanar the curvature is. Lastly,

$\dot{B} = \dot{T} \times N + T \times \dot{N} = 0 + T \times (-\kappa T + \tau B) = -\kappa\tau T + \tau^2 B = \tau \dot{N}$.

Below are 3 connections which make up Serret-Frenet equations.

$dT/ds = \dot{T} = \kappa N$,

$dN/ds = \dot{N} = -\kappa T + \tau B$,

$dB/ds = \dot{B} = -\tau N$.

There are 2 types of crossings of the 2 backbone curvatures: local and nonlocal. Local crossings take place when considering a section of the ribbon, while nonlocal crossings happen upon the complete traversing of one section of the ribbon over another.

The orbit of the ribbon is crossing over itself, as well. This is how the depiction of Twist can be made considering these local crossings.[10]

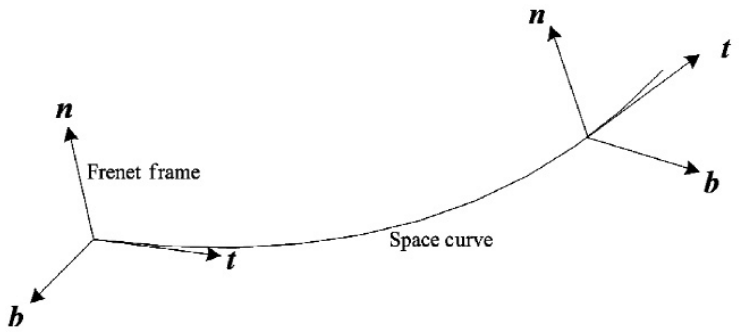

Figure 23. Space curve

3.2. Twist

The ribbon twists around its central orbit and this demonstrates how rigidly the ribbon twists around its orbit. With regards to DNA double helix, the helical pitch, which is the number of base pairs when an entire circumvolution occurs, is demonstrated by these local crossings. Instinctively, a hypothetical way of thinking might lead to the stringing of a penny on a bend. Supposing that the northern pole is pointed, at that instant Tw will gauge how this very pointed alters as the penny is glided along the curvature. More specifically, this is gauged as Twist, Tw, adopting the Frenet framing, since the total twist of the curve γ(s):

$$\text{Tw} = \frac{1}{2\pi}\int_{\gamma(s)}(T \times N).Nds = \frac{\dot{1}}{2\pi}\int_{\gamma(s)}\tau ds.$$

Afterwards, within the Frenet framing, T × N, $\dot{N}$ is the estimation of $\dot{N}$ in the steering of B, so by the Serret-Frenet equations, = τ .

3.3 Write

A writhe can be described as the way how the orbit of the ribbon is twisted in the space. It is possible to give the total of the marked crossing of the orbit with itself on condition that the orbit which obtains the direction of the two ribbon edges. This amount is entitled as writhe, that is, Wr. This being the case, Wr, which is within particular a Reidemeister Type I move, will subsidize a ±1 to Wr. Sure enough then, Wr is an invariant which is not topological which heavily rests on the specifically provided geometry of a projection. It should be noted here that the (Note that the jargon differs varies in the literature. For instance, Weber names the marked total of the crossings of a projection as the Tait number of the projection.) [11].

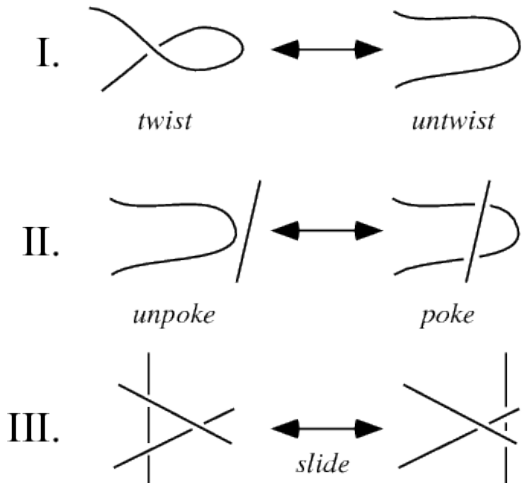

Figure 24. Reidemeister move, which in the field of mathematics of the knot theory, consists of each and every three local moves existing on a link diagram. Each move performs on a small zone of the diagram and is described as one of three types.

More precisely, writhe can well be taken into consideration with regards to Gauss' integral as in the formula provided in the Serret-Frenet section in which the orbit depicts C and W curve. Wr can be considered with regards to Gauss integral as Lk provided above, with both combinations along the same the ribbon orbit.

The self-intersections of the DNA orbit subsidize to writhe in terms of a DNA molecule. Metaphorically, the cord which adhered to the handset of the macroscopic analogue, the archaic telephone receiver, writhes by drooping from the handset when used. As this example, the connection between the topological and geometric features of a ribbon was first discovered by Călugăreanu [4] and White [5]. Its relation to DNA was first marked by Fuller [6].

This resemblance is frequently envisaged by biologists by making use of a length of clear plastic tubing, when two parallel coils are drawn along the length by depicting one edge of the ribbon. By bending and/or twisting the tube before the closure of the ends, we can get the depiction of the connection between Twist and Writhe.

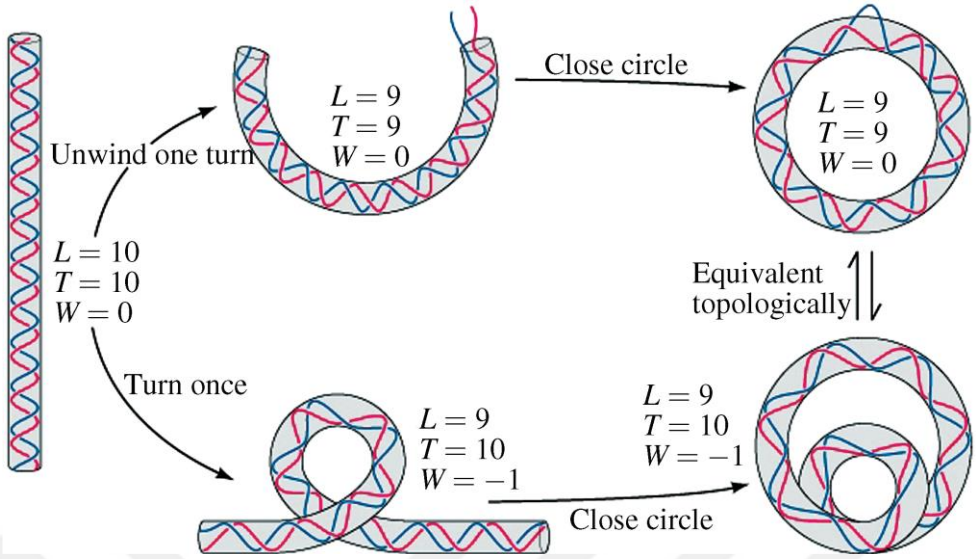

Figure 25. Writhe-2 Copyrigth 1999 John Wiley and Sons,Inc.All rights reserved.

For a sophisticated approach of this, see Dennis and Hannay [25]. For a more comprehensive analysis of these differential geometric concepts with regards to DNA, see Weber [11] or Pohl [12].

CHAPTER 4

THE COMPACTIFICATION OF DNA

4.1. Biological significance of DNA supercoiling

As stated before, nontrivial writhe takes place most usually when the orbit of DNA encloses itself in a circular way, as observed Vinograd and his colleagues [3]. More specifically, this is attributed as 'plectonemic supercoiling' (hereafter supercoiling), so as to reverses the coiled (writhed) orbit of the already coiled (helical) ribbon. DNA can also enclose proteins, as explained below, and this type of nontrivial writhe is attributed as 'solenoidal supercoiling'. This 'supercoiled' DNA was found to be the original condition in nearly all cells [13]. There are diverse organisms from cauliflower to humans, mice to monkeys with DNA extracted from their cells and it was discovered that the DNA was supercoiled with almost same denseness. We can gauge this denseness thus and thus. If N, the number of base pairs of DNA and h, the number of base pairs per helical repeat in given experimental circumstances, then, the precise linking disparity is σ:

$$\sigma = \frac{Lk - \frac{N}{h}}{\frac{N}{h}}$$

Supercoiling of DNA is unique since it is significant for three main reasons. Most importantly, it compactifies the DNA molecule. Secondly, the chirality state of the supercoiling lets the major/minor grooves be accessible to the base pairs. Thirdly, supercoiled DNA supplies a significant source of free energy for cellular reactions.

4.2. Compactification

Definiton 4.2.1 The process of compactifying involves turning a topological space into a compact space. Every open cover of a space is said to involve a restricted underlay in a compact space.

All creatures have genomic DNA, which is physically considerably taller than the region it resides in. The DNA hovers without restraint in prokaryotes, while it is enclosed within the cell nucleus in eukaryotes. Eukaryotic genomic DNA requires a bigger length of capacity, so though the ordinary compactification required to supply DNA in prokaryotes is important, it is even more significant in eukaryotes. For instance, the most typical bacteria *E. coli* has genomic DNA of length about 1.5 mm, and cell diameter less than 1 μ meter, which requires compactification approximately10^3. Human genomic DNA has a dimension of about a meter in length and a common cell nucleus has diameter less than 1μm, therefore the compactification which is required approximately 10^6. South American lungfish, with a DNA of 35 meters contains one of the most complex storing problems. Compactification has the stages of its own and supercoiling is the first step of a these stages the rest of which is not comprehended clearly. Later, the supercoiled – plectonemically – DNA molecule encloses giant proteins which are called as histones. Next, the histones are chained together like pearls on a necklace. This string of histones is later compacted more, likely as offered in Figure below.

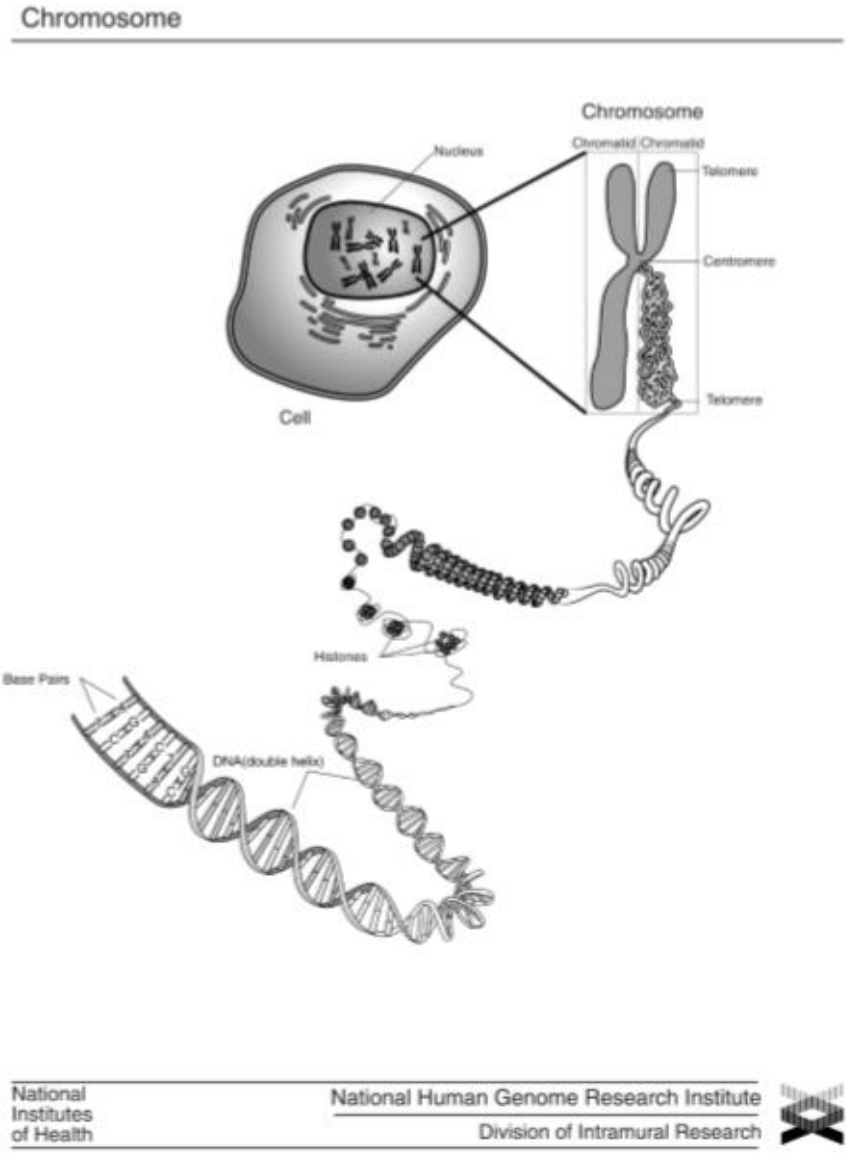

Figure 26. Compactification of DNA

Hence, supercoiling lets DNA to be stored in a rigidly and profoundly structured, arranged way. The organization of this operation is significant for productive information recuperation — since it is a must for the portions of DNA to be accessed briskly and continuously by proteins.

Let the circle with center at (0.1) and radius 1 in R^2 be denoted by *S*. Let the (0.2) vertex on the circle be denoted by *p*. Since the circle *S* is closed and limited, it is a compact. Let the point where circle *S* intersects the line joining each points of *x* of R to the point *p* be denoted as x'. In this way, a $f: R \to S, x \to x'$ function is obtained. Here, since R and S

\{p} are homeomorphic and the subclause S\{p} is dense with S, the circle S is a compactification of R [79].

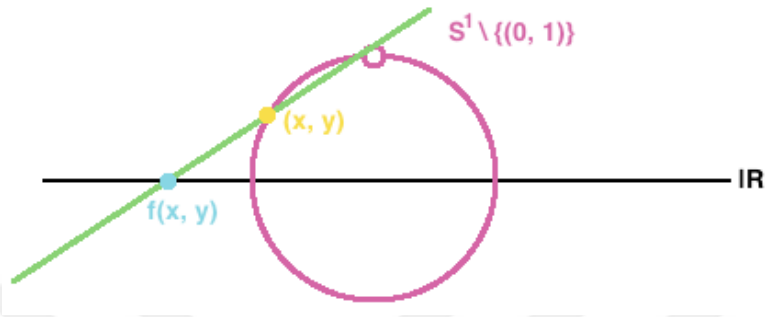

Figure 27. Compactification of a Topological Space.

4.3. Entrance to the base pairs

The chirality state of DNA supercoils happening or existing inside a living body is negative. In other words, the original double helix is not like the helix made up by the DNA orbit in chirality. For linking number, Lk, to remain steady, the circular molecule is supposed to remain covalently closed while supercoiling. Therefore, Lk = Tw+ Wr hints that there must be a repaying shift in *Tw*. The chirality – handedness – of the supercoils means that the molecule doesn't twist enough (becoming undertwisted). Therefore, the primary and secondary grooves, as depicted in Figure 14, open a small quantity. This permits the reciprocal base pairs to be accessible, thereby expediting varied reactions. It also aids in cleaving the DNA strands, which is fundamental for example for duplicating DNA.

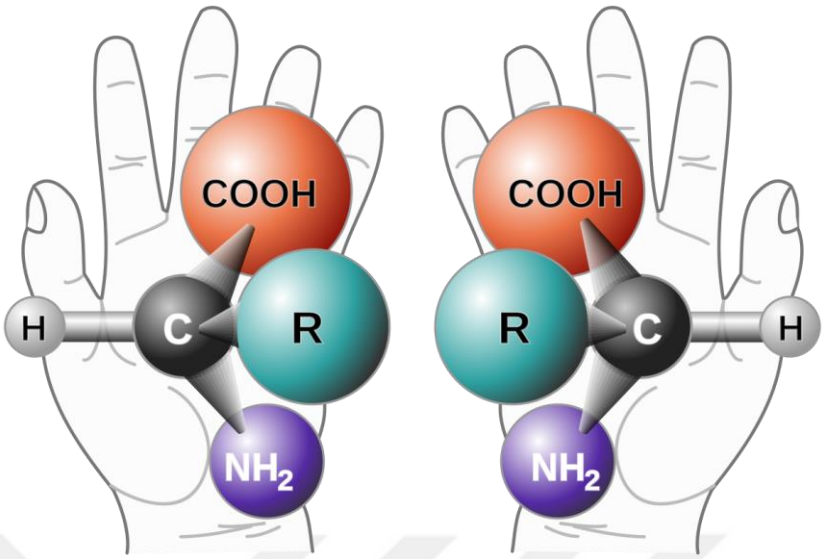

Figure 28. Chirality (handedness)

4.4. Free energy

As specified earlier, cleaving the DNA double helix might be hard. However, most of the fundamental reactions, encompassing the duplication of DNA, require precisely this. Luckily, Călugăreanu-White-Fuller theorem in which Lk = Tw+Wr. states that the linking number of loosened DNA is higher than that of the negatively supercoiled DNA, which is the innate condition of DNA in cells. For the cleaving of local strands to take place the energy to arise from the linking number 'deficiency' of supercoiled DNA is needed. Moreover, the orbit of DNA must be coiled or twisted for most for many biological progresses which comprise coiled enclosing around histones and Phage DNA packaging [14]. In what way the free energy relies on linking disparity in order to comprehend the dynamics of these progresses is concisely explained.

For DNA circles which are bigger than 2000bp, it has been empirically indicated in [15], [16], that the free energy of supercoiling, which is ΔG, follows the quadratic equation.

$$\Delta G = K(a - a_0)^2$$

in which K, relying on N, the number of base pairs, is comparatively perpetual, α is denominated as the linking number, and a_0 (for large DNA) is the mediocre linking number.

For big N, the output NK has been empirically indicated to be free of breadth [17].

4.5. Adjusting DNA Topology

Definition 4.5.1. Topoisomerases are a family of enzyme which generally brings DNA to its relaxed condition. This means that topoisomerases disentangle overwound DNA strands, curtailing the linking number back to the selected 10.4 base pairs per turn, or rewind underwound DNA strands, increasing the linking number back to 10.4 base pairs per turn. Topoisomerases function by cracking one or two DNA strands and going through the same number of DNA strands through the crack. This alteration brings about an escalation or diminishment in the linking number of our topologically restricted circular DNA [78]. Topoisomerase is an isomerase enzyme seen in the topology of DNA and function in order to solve the topological problems related to DNA replication, transcription, recombination, and chromatin remodeling.

Definition 4.5.2. 'Boltzmann's constant' is a physical coefficient that gives the relationship between energy and temperature. It can be found by dividing the gas constant by the Avogadro number.

DNA molecules that differ only in Lk (not including sequence or length) are entitled as topoisomers. According to the standart thermodynamics, the concentration of DNA with linking number αi, [αi], is

$$[a_i] = \frac{1}{Z}\exp\left(\frac{-\Delta F}{RT}\right)$$

in which Z is a normalizing constant, T is the temperature in Kelvin, and R is Boltzmann's stable. Thus, the concentration of a provided topoisomer with linking number a_i, is

$$[a_i] = \frac{1}{Z}\exp\left(\frac{-K(a_i - a_0)^2}{RT}\right)$$

Here standard deviation is $\sqrt{\frac{RT}{2K}}$ and the chart of αi v. [αi] for stable populations of topoisomers of big DNA circles is a common dispersion by taking a_0 to the center.

In order for this counterbalance to continue, two groups of proteins have expanded to topoisomerases. Topoisomerases are entitled in this way, since the main objective is to remodel within among topoisomers. Topoisomerases are divided into two groups, which is based on whether they split 1 DNA backbone thus altering the linking number by steps

of 1 (Type I topoisomerases) or both backbones hereby altering linking number in steps of 2 (Type II topoisomerases) [18].

The concise discussion of these two families is given below, but it should be noted that most of the obscurities of topoisomerases—especially, small disparities between family members of a given type are only touched on.

4.6. Regulating DNA supercoiling

4.6.1. Type I topoisomerases

The way in which Type I topoisomerase functions is by cracking one of the two DNA strands and going through the other strands through the hole. This escalates or diminishes the linking number by 1. As shown in figure 25, before the topoisomerase affixes a twist the top DNA double helix depicts the structure, and the bottom DNA double helix exists, after the enzyme carries out the reaction. Type I topoisomerase reactions do not require additional energy.

DNA supercoiling is apparent.There are proteins function of which is to adjust the quantity of supercoiling within DNA.These proteins are called as Topoisomerase -1 enzymes.These enzymes loosen up supercoiling [78].

Assuring the essential significance of DNA supercoiling, the existence of proteins the objective of which is to adjust the quantity of supercoiling, belonging to Topoisomerase-1 family should not be shocking. Members of this family loosen up supercoils.

Type I topoisomerases efficaciously control the theorem of Lk = Tw+Wr. so as to transform an alteration in twist to an alteration in writhe which is also called supercoiling. They are attached to the DNA molecule, make a temporal fracture in one of the DNA backbones and then transcend the other backbone before letting the DNA out. This alters the twist, thereby altering the writhe for a round or topologically strained molecule. The cases where type I topoisomerases get involved is when the DNA coil (helix) is unfastened like at the time of DNA duplication (replication). If the molecule is strained at the edge towards where the unfastening divergence is moving, then supercoils are added instantly to the unfastened zone. The eclipsing (torsional) strain gets too constrained and molecule of DNA fractures on condition that a type I topoisomerase does not free the supercoils.

All organisms which have been studied so far contain Type I topoisomerases the absence of which will be fatal for the cell.

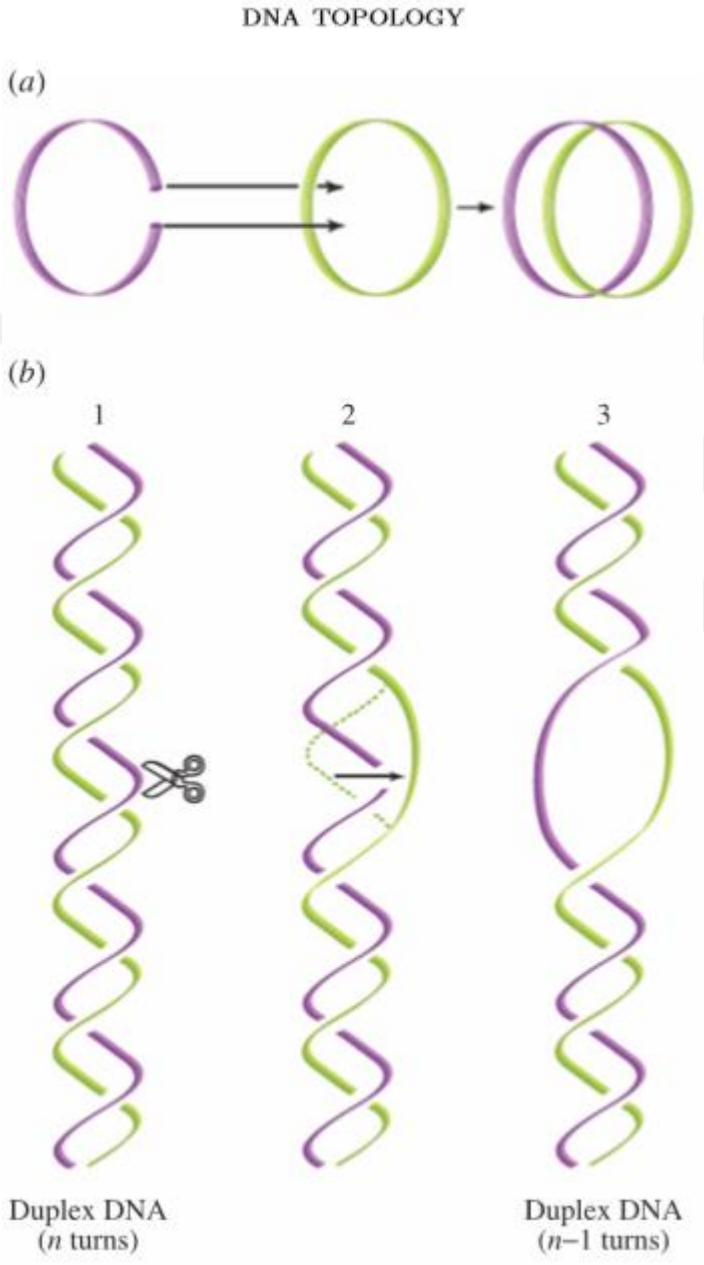

Figure 29. The operation of Topoisomerase I. From Biochemistry by Donald Voet and Judith G. Voet (2004).

4.7. The adjustment of the knotting and linking of DNA

4.7.1. Type II topoisomerases

Type II topoisomerase cracks both of the two DNA strands and goes through the whole double helix through the hole. This escalates or diminishes the linking number by 2. Here, the top DNA double helix depicts the structure before the topoisomerase functions, and the bottom DNA double helix exists after the enzyme fulfills the reaction. Under this circumstance, the outcome is that the writhe of the molecule escalates by 2. Type II topoisomerase reactions necessiate some form of energy, like ATP or NADH. Bacteria also own a particular kind of Type II topoisomerases known as ‘gyrase’ [78].

Type II topoisomerases are not only pervasive but also fundamental like type I topoisomerases. They can take DNA supertwists (supercoils) out (the instance of DNA gyrase) or add them though their main function is to alter DNA knot or link kind.

When round DNA like bacterial genomic DNA is duplicated type II topoisomerases operate by their nature. The final product of this action is two related circling molecules. The bacterial cell is unable to split perfectly and self-destructs if a type II topoisomerase connects two daughter molecules.

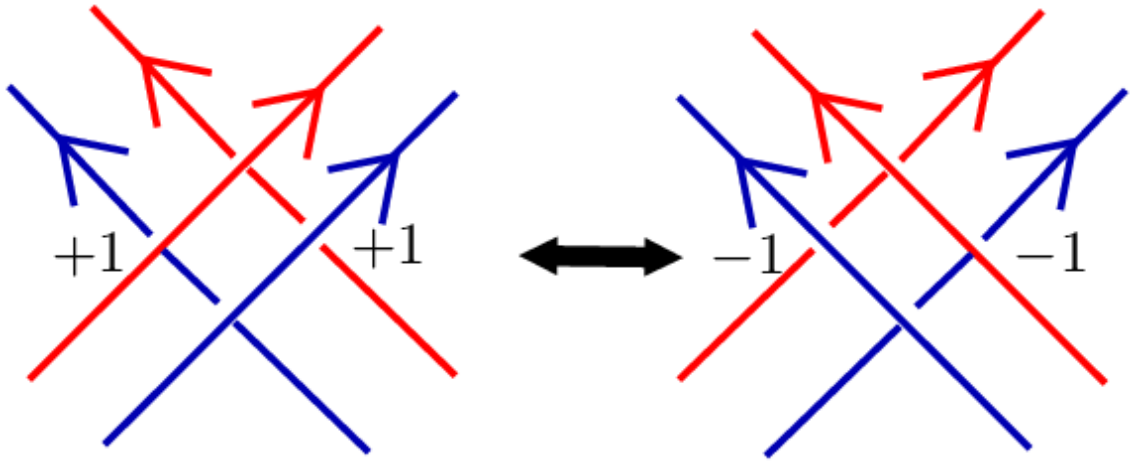

Figure 30. The unknotting of Topoisomerase II

Type II topoisomerases have been the objective of drug targets for human cancer and contagious illnesses due to their primary function in most of the cellular procedures [19, 20]. For instance, because of the fact that bacterial type II topoisomerases are distinct from human, many antibiotics operate by hindering these proteins, which destroy the

bacterial infection by hampering cleavage of daughter DNA as depicted. Therefore, comprehending the system through which type II topoisomerases operate has been very significant in terms of from both a primary biochemical and pharmaceutical aspect.

4.7.2. The system of type II topoisomerases

It may be presumed that type II topoisomerases, which are tiny molecules, operate freely of the world-wide DNA topology. Particularly, nevertheless, they operate in a way which preferably does not knot and link DNA, the process which is his is acknowledged as topological simplification. Moreover, DNA topology is simplified by type II topoisomerases in an extremely effectively way [21]. Precisely how this is accomplished has been comprehensively discussed.

That a generic type II topoisomerase ties helix-helix proximities (juxtapositions) like a supercoil or twist/link crossing as shown in figure 19 is frequently accepted. The first knot, which is entitled as the G (standing for Gate) section, is fractured to let the T (standing for Transported) section go beyond in a one direction way [22]. Then, the G section is resealed, causing in a crossing alteration (from +1 to −1 or vice versa). This procedure depends on ATP hydrolysis [22]. An appealing new crystal system of a type II topoisomerase tied to the G section gracefully shows that a type II topoisomerase firmly curves the DNA (roughly 150◦) [23]. This backs types that blend the curving of DNA [24, 25]. The detailed discussion of the two is provided below:

G. Buck, and L. Zeichiedrich [24] a topologist and a biologist consecutively, who suggested the first model together explains how the global topology can be determined by utilizing local data retrieved at a crossing. They divide crossings with regards to many vector parameters, and indicate that particular crossing kinds such as ‘hooked junctions’, are linked to nonsupercoiled crossings. Crossing which alters at these hooked junctions is probably to cause to topological simplification later.

The second model [25] by Vologodskii and co-workers covers numeric (Monte Carlo simulations) and indirect empirical proof that a type II topoisomerase, in fact, twists the G section into a hairpin conformation while tying. There is some empirical proof that DNA knots pinpoint, which would cause the T-segment in the hairpin to exist frequently for knots than compared to the unknots.

4.7.3. Topological and numeric approaches to comprehend the unknotting of type II topoisomerase

In topologic sense, comprehending the unknotting of type II topoisomerase is component of a bigger move to comprehend what kind of knots can be retrieved from a provided one through a sole crossing alteration, that is to say, by a sole round of topoisomerase loosening.

In a broad sense, this is a very big query—for instance, one field of knot theory exploration has been to (up to now not entirely) categorize all knots with ‘unknotting number 1’—in other words, all knots which turn out to be the unknot after a sole crossing alter. More broadly, there is a strand passage distance which can also be called as the Gordian passage between 2 links or knots (K and L) and this d(K,L) is explained to be the slightest number of crossing changes (+1 to −1). These changes are then required to transform K into L. It is a measurement on the space of kinds of knots—like d(K,L) is not more than or equivalent to the total of their unknotting numbers.

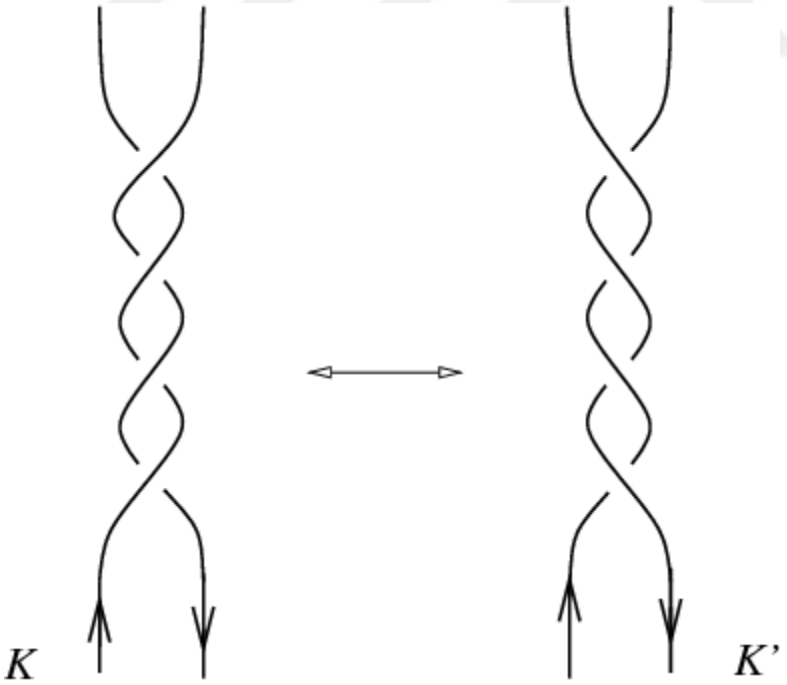

Figure 31. Knot adjacency

In the scheme below, we can think about which knots L are beside to a provided knot K—i.e. L, s.t. *d(K,L)* = 1 or equivalently which knots L are single Topoisomerase-II go from K. Lately, there has been an abundance of action in this field of 'knot adjacency'.

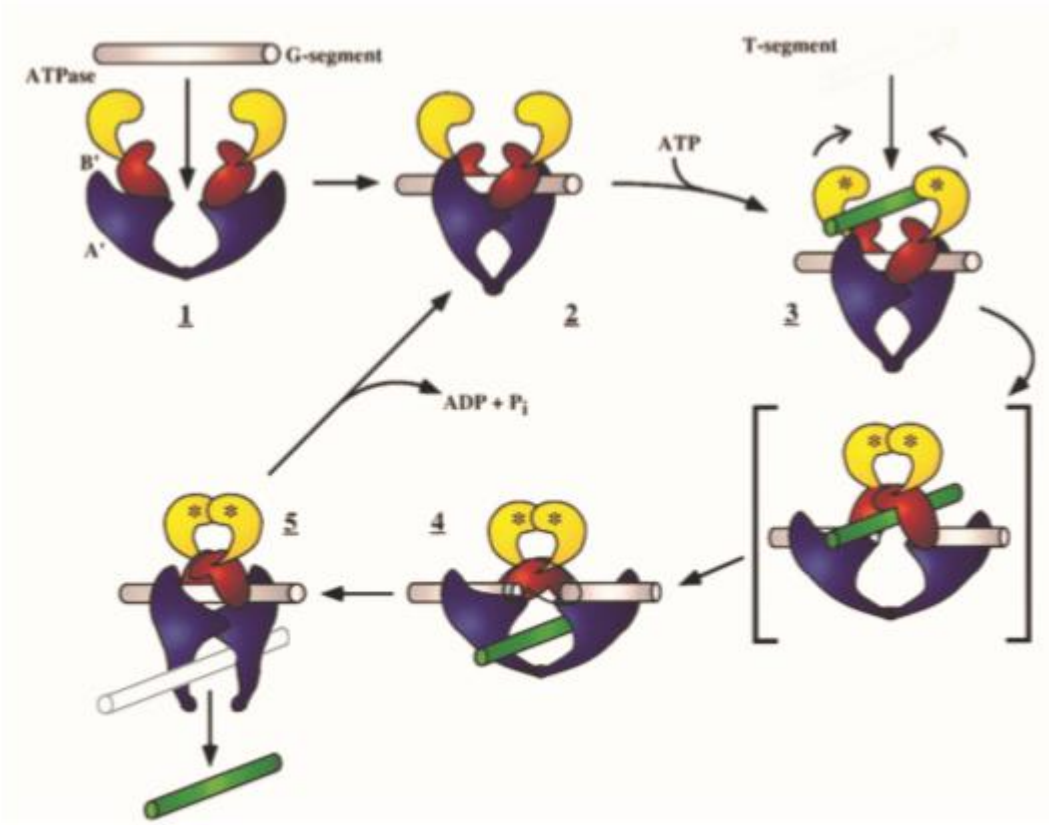

Figure 32. Approximate system of a type II topoisomerase. (J. M. Berger, S. J. Gamblin, S. C. Harrison, J. C. Wang (1996).

Numerical simulations by Stasiak and colleagues have examined the possibility of K being transformed into L (where probably K, L) through a sole type II topoisomerase [34]. Approximately, they commence with a closed knotted polymer chain, randomize it, and then let 1 more move that could (or could not) include a sole strand passage. They then compute the new knot type. Some impressive things turn up—such as in 95% of the conditions for the unknot, a sole intersegmental passage recoiled the unknot anew.

CHAPTER 5

SITE-SPECIFIC RECOMBINATIONS

5.1. Alteration to DNA topology

As mentioned before, the primary and only function of type II topoisomerases is to alter DNA knot or link type. Now the focus and attention of my study is towards knots and links which derive from the site-specific recombinases, which is another group (family) of proteins. Site-specific recombination, rearranging of the genetic sequence such as altering GATTACA to ACATTAG are all carried out thanks to these proteins.

A wide array of biological processes contains site-specific recombination, which makes it quite important. The outcome of site-specific recombination is the removal, infusion or transposition of a DNA segment. (See Figure below) This conforms to diverse physiological processes, including significant strides in viral infections.

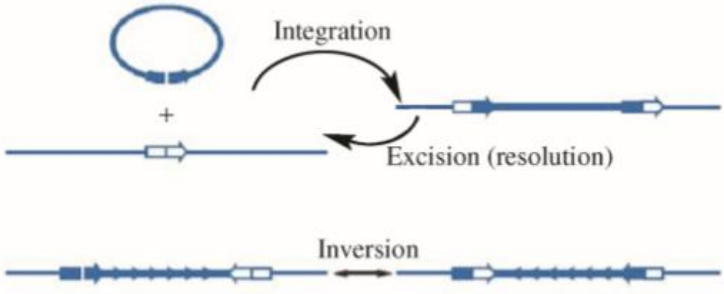

Figure 33. Results of site-specific recombination.

Besides their innate biochemical nature, genetically altering organisms or testing as to whether a mutation in a specific gene causes an illness have become an area of research and attention for pharmaceutical and agricultural industries. Consequently, site-specific recombinases as means for correctly manipulating DNA intrigue these industries [56].

Although altering DNA topology is not the fundamental function of these site-specific recombinases, it can be an offshoot of the response. The supercoils can be transformed into knot or link crossings during the course of recombination on condition that the authentic round DNA is supercoiled [57], [58].

5.2. The nuts and bolts of site-specific recombination

Basically, both the site-specific recombinase and two short (30–50bp) DNA segments, the crossover sites which are embedded into 1 or 2 tiny round DNA molecules is necessitated by site-specific recombination. Direct or indirect (inverted) orientations of site-specific recombinases are possible on condition that there are 2 crossover sites on a sole molecule of round DNA (head-to-tail, e.g., ...ATGC...ATGC) or (for instance, head-to-head ...ATGC...CGTA). (See Figure below) Extra proteins and DNA sites could also be necessitated by bigger site-specific recombination systems as well [59].

While site-specific recombination occurs, two recombinase molecules first tie each crossover site. The two crossover sites are next aggregate within a *recombinase complex*, which is B (abridged of *Ball*). This is the tiniest curve zone which includes the four tied recombinase molecules and the two crossover sites [88]. Therefore *B* is a *topological ball* (i.e., it can be distorted to a circle ball). The crossover sites can be pinpointed either on the exterior or interior of the 4 recombinase subunits [61, 62, 39].

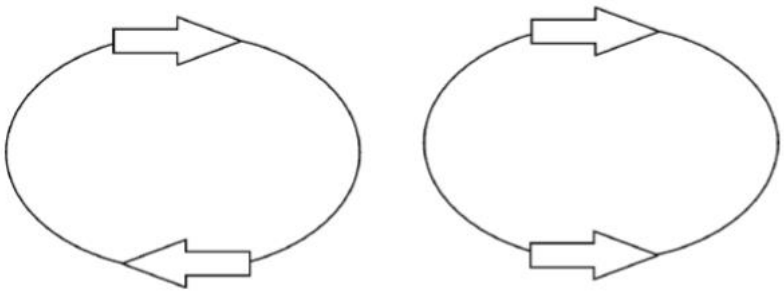

Figure 34. On a round DNA molecule, the two crossover sites can be in direct (left image) or inverted orientation

There are three approximate phases of site-specific recombination. First, two recombinase molecules tie to each of two specific sites on one or two molecules of round DNA (which is known as the substrate) and then band them close together. The sites together with proteins which are tied are called the recombinase complex. Next, the sites are separated,

alter with each other and resealed. The exact nature of this intermediary step is ascertained by two recombinase subfamilies as which specific protein it belongs to. In the end, the readjusted DNA, the output, is discharged.

Prior to the discharge of DNA numerous rounds of strand alteration can take place, which is the process called as processive recombination. In distributive recombination, however, numerous rounds of the whole procedure of recombination containing the discharge and releasing and rebinding takes place.

5.3. The two families of site-specific recombinases

There are two families of site-specific recombinases; the first one being the—the serine (also known as the resolvase) and the second one being tyrosine (also known as the integrase) recombinases—which are all hinged on the specific amino acid in the protein polymer which catalyzes the reaction of dividing **[**59]. The serine and tyrosine recombinases are distinct in terms of their means of incision and assembling DNA at the crossover sites. A developmental study has been carried out on 72 serine recombinases [63] together with a current sequence inquiry documents roughly 1000 associated sequences of presumed tyrosine recombinases, [64] resulting in the fact that both families are considerable.

5.4. DNA knots and links in terms of recombination outputs

As mentioned earlier, supercoiled round DNA can be converted into a knot or link through site-specific recombination. Considering a substrate having crossover sites in a specific orientation, a single circle of recombination conciliated by a serine recombinase produces a unique topological output. For instance, Tn3 operating on an unknot with straight sites produces entirely the (2,2)-torus link (also called as the Hopf link).

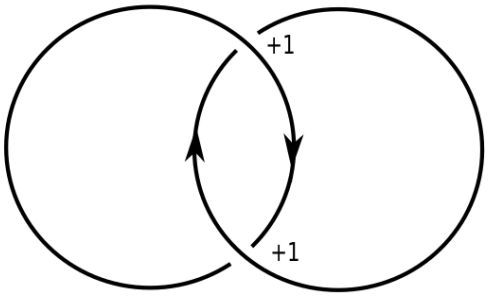

Figure 35. The Hopf Link

Comprehending exactly which knots and links emerge while site-specific recombination is taking place makes it possible for us to understand the specific aspects of the process [68]. The identification of the knotted and linked outputs of site-specific recombination is made through topological techniques. For instance, several means have been established to find out a specific DNA knot or catenane type, containing the usage of the node number for knots [35], the Jones polynomial for catenanes [69], Schubert's classification of 4-plats [70] and the HOMFLY polynomial [71]. Now, two huge topological contributions for us to comprehend the site-specific recombination are given below.

5.5. The envision of DNA knot and link outputs

Lately, Buck and Flapan [72, 57] have advanced an auguring model for DNA knots and links that appear as outputs of site-specific recombination. More particularly, instead of concetrating on a particular recombinase as have been carried out in several of the former studies, the writers introduced a topological model which envisions which knots and links can take place as outputs of site-specific recombination commonly. They manage this by identifying the topology of how DNA knots and links are assembled as a consequence of a single—or multiple circular of advancing—recombination event(s), considering a plectonemically supercoiled unknot, unlink, or T(2,m) torus knot or catenane substrate. (See Figure below for demonstrations of the kind each substrate.)

The common concept is to regard a ball represented as *B* as including the convex hull of the four recombinase molecules, and a enclosing surface D (including a disc for the unknot, two disjoint discs for the unlink, or twisted band for a torus knot or link). Its

borderline is the unknotted DNA orbit. Later D ∩ B pre-recombination and post-recombination is to be taken into consideration. Likewise, D ∩ *cl* (S^3 \ B) for each substrate is to be typified. If topological propositions are taken into account, we can confine these possibilities. Then stick each of the post-recombinant forms of D ∩ B to each form of D ∩ *cl* (S^3 \ B) to categorize likely output knots and links.

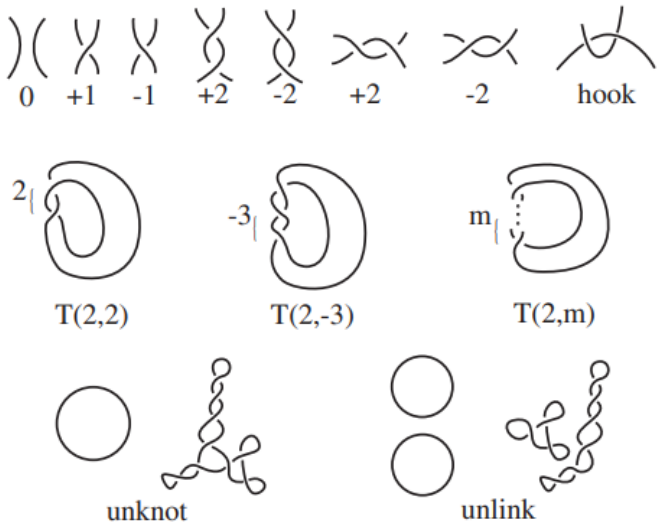

Figure 36. Likely arrangements of the recombinase complex.

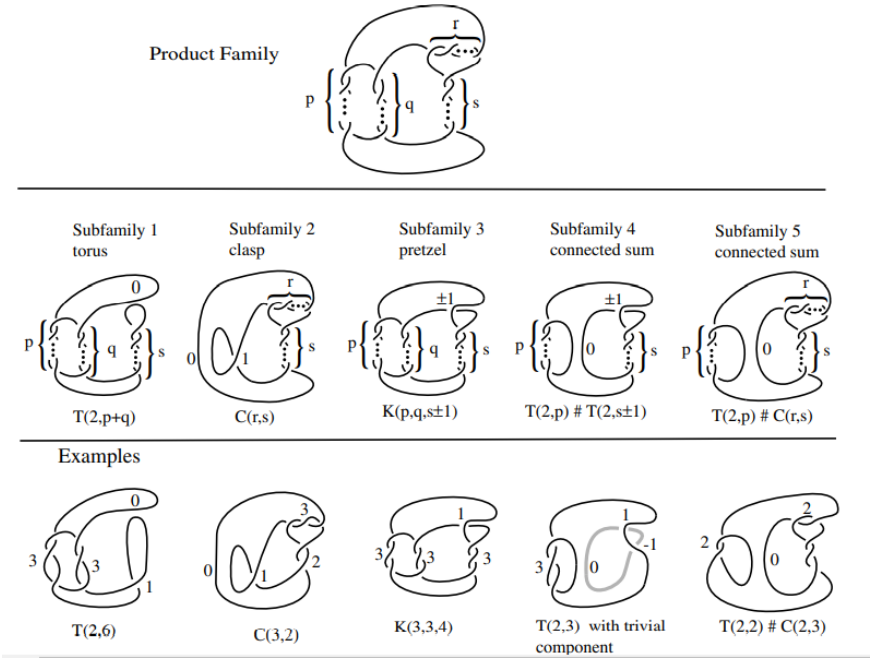

Figure 37. Possible substrates: the unknot, unlink, or torus knot or link.

The model is free of the magnitude of the substrate, the indication of the supercoils and the site direction. Biological proofs back up three of the hypotheses on which the model depends. It is envisioned that the outputs emerging from site-specific recombination must be members of an individual group of outputs when these hypotheses are taken into consideration. (As demonstrated in Figure above). Though it shall be explained more elaborately this is a particular family of Montesinos knots and links, presented like (1/p , 1/q , r/rs+1), with p, q, r, s ∈ Z.

This family is advantegous in terms of magnitude which is small enough. Empirical approach can be used to detect the slightest crossing number (MCN) of a DNA knot as mentioned earlier [49]. Several knots or links do not exist with a provided value for small values of the MCN. Yet, the number of knots and links with MCN = n becomes increasingly larger as a function of n [73], and available number of knots is 1,701,936 with MCN ≤ 16 [74]. So to detect the knot knowing the MCN is not enough.

There are several utilizations of this study. One of them is it can envision knot and link outputs for formerly undefined data. It can also detect way of recombination: processive versus distributive. It especially anticipates that the knot or link outputs which is not within the family mentioned above must rise in a scattered way. And lastly, sequence of outputs of processive recombination can be detected through this work.

This novel endeavors mentioned below used the tangle model and many biologically plausible hypotheses to resolve the equations to arise from tangle. Later, which knots and links ensue as a result of site-specific recombination on the unknot for the serine subfamily of recombinases have been ascertained.

CHAPTER 6

MODELLING OF DNA

6.1. The tangle model

Let B be a 3 sphere and t be a non-directional pair of arcs embedded in B. The four end points of these pairs of arcs are at the equator points of the sphere (NW, NE, SW, SE). A tangle is a (B, t) pair. A tangle diagram is obtained by projecting the tangle onto the equatorial plate. The endpoints on the diagram will be marked as NW, NE, SW, SE. Rational tangles are tangles that can be transformed into an apparent tangle by shifting the end points. Rational tangles resemble the structure of DNAs and will therefore be our prime focus. Also, there are also tangles that cannot be transformed in this way which are the primary tangle and locally knotted tangle structures.

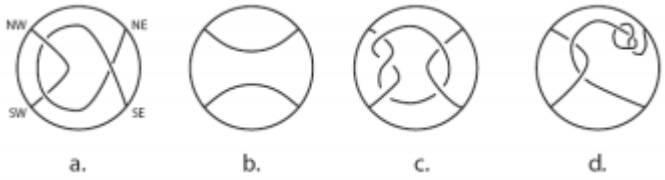

Figure 38. Tangle Samples a.Rational b.Apparent c.Rational d.Locally knotted

Each rational tangle (a_1, a_2,... a_n) is formed by the vector $a_i \in Z$, ∀i. Additionally, the tangle diagram can be drawn with the help of this vector: First, the points NW, NE, SW, SE are to be connected with a circle marked and draw the arcs as shown in the figure. If n is even, binding has to be carried out at the bottom (SW and SE) and half shift as much as the number a_1 (if a_1 is positive, we do a right-hand shift, if a_1 is negative, a left-hand shift is done.) Then, a shift as much as a - 2 in the NW-NE part of the diagram is done. Then, heading to the bottom and continue to do the same is necessary. If n is odd, the binding is done from the right side and the operation continues as before. For example, the rational diagram of (2, 1, 2) is drawn as follows.

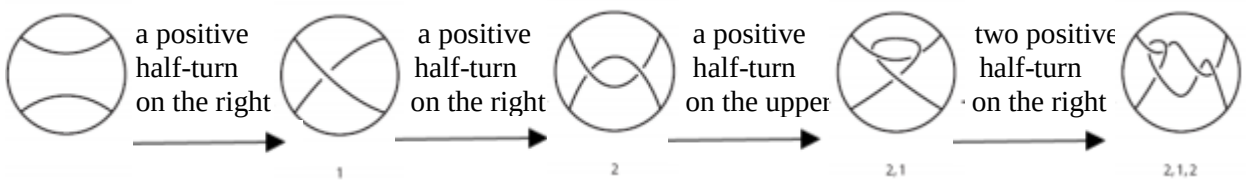

Figure 39. Rational tangle

Each integer vector can be expressed as a continuous fraction equal to the rational number $\frac{\beta}{\alpha}$. If T is expressed by tangle $(a_1, a_2, \dots a_n)$ vectors, then continuous fraction is found in the form of;

$$\frac{\beta}{\alpha} = a_n + \frac{1}{a_{n-1} + \frac{1}{a_{n-2} + \cdots + \frac{1}{a_1}}}$$

The rational number $\frac{\beta}{\alpha}$ is called the fraction of the tangle T.

Theorem 6.1.1.: The two tangles are isotopic if and only if they have the same fractions.

If one of the two tangles can be transformed into the other without any strand breaking or one strand running over the other without moving the end points, they are equivalent. And thanks to this theorem, it can actually be understood how important the fraction of tangle is in determining the properties of tangle.

Above, how a fraction tangle is obtained from a vector has been seen, similarly, the vector can be obtained thanks to the tangle fraction:

$$\frac{\beta}{\alpha} = a_n + \frac{1}{a_{n-1} + \frac{1}{a_{n-2} + \cdots + \frac{1}{a_1}}}$$

Of course, different vectors can be obtained in this way.

For example; The vectors (3, −2, 2) and (2, 2, 1) both give the fraction of 7/5. However, here the two vectors denote the same tangle. Also, the entire rational tangle can be expressed by a single canonical vector called the Conway symbol. If a vector $(a_1, a_2, \dots a_n)$ is in canonical form, for every $1 \leq i \leq n - 1 \mid a_1 \mid > 1$, $a_i \neq 0$, and all

nonzero expressions have the same sign. The Conway symbol for our example above is (2, 2, 1).

Theorem 6.1.2. There is a 1 - 1 mapping between the fraction $\frac{\beta}{\alpha} \in Q \cup \frac{1}{0} = \infty$ ($\alpha \in N \cup \{0\}$, $\beta \in Z$ and gcd (α, β) = 1) and the set of rational tangles. tangles can also be expressed as matrices. These tangles are expressed in terms of 2 × 2 matrices, like this:

$$\begin{bmatrix} u & v' \\ v & u' \end{bmatrix} = \begin{bmatrix} 1 & a_{2k} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a_{2k-1} & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 \\ a_1 & 1 \end{bmatrix}$$

Example 6.1: The matrix expression of

$$\frac{\beta}{\alpha} = \frac{23}{17} = 1 + \cfrac{1}{2 + \cfrac{1}{1 + \frac{1}{5}}} = (5,1,2,1)$$

is as follows:

$$\begin{bmatrix} u & v' \\ v & u' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 23 & 4 \\ 17 & 3 \end{bmatrix}$$

6.2.Tangle Operations

Definition 6.2.1. Given the tangles A and B, the sum of these two tangles A + B is obtained by adding the NE and SE endpoints of one tangle to the NW and SW endpoints of the other tangle, respectively.

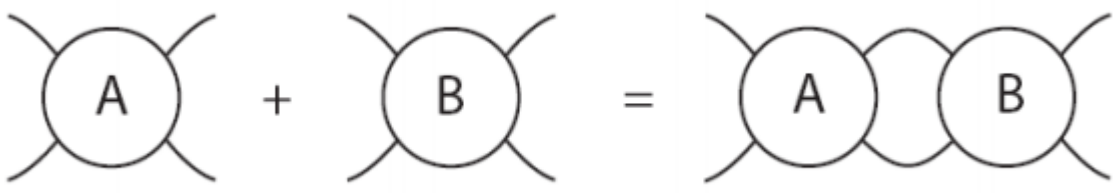

Figure 40. Tangle Operation-1

Definition 6.2.2. The process of N (T), known as the numerator closure of a T tangle, is obtained by joining the NW and NE endpoints and combining the SW and GD endpoints

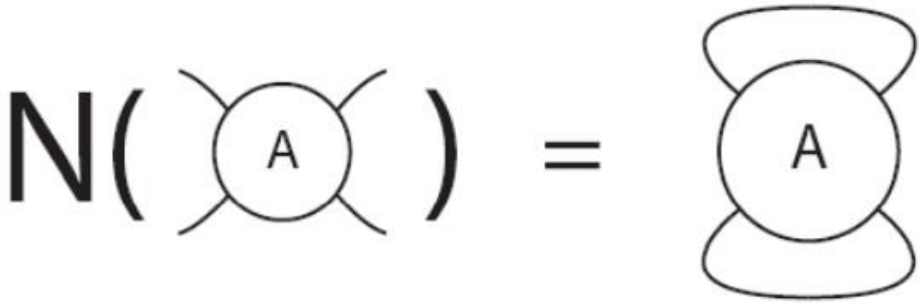

Figure 41. Tangle Operation-2

Definition 6.2.3. The process of D (T), known as the numerator closure of a T tangle, is obtained by joining the NW and SW endpoints and combining the NE and SE endpoints.

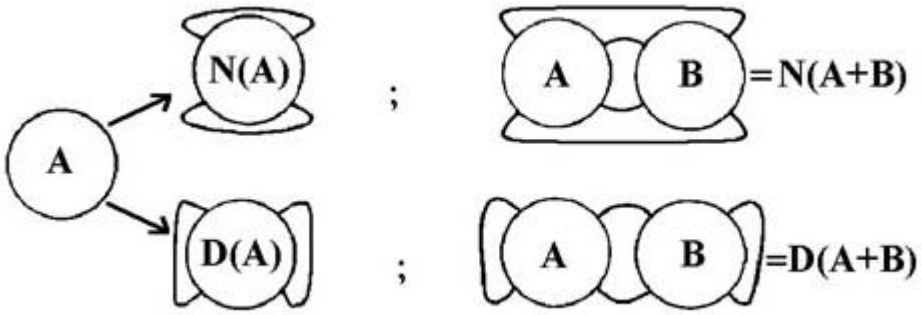

Figure 42. Tangle Operation-3

Example 6.2.1: Tangle operations can also be used together. The knot obtained as a result of N ((2, 0) + (1)) = ⟨3⟩ operation is trefoil, that is 3_1. Here (2, 0) is the Hopf chain.

So this a K knot which is obtained as the result of N (A + B) = K operation. Also, the sum of two rational tangles may not always give a rational tangle. This is exactly how we can think about how the process will work; but the numerator closure of the sum of two rational tangles yields a knot called 4-plat, and this is an important issue in DNA modeling.

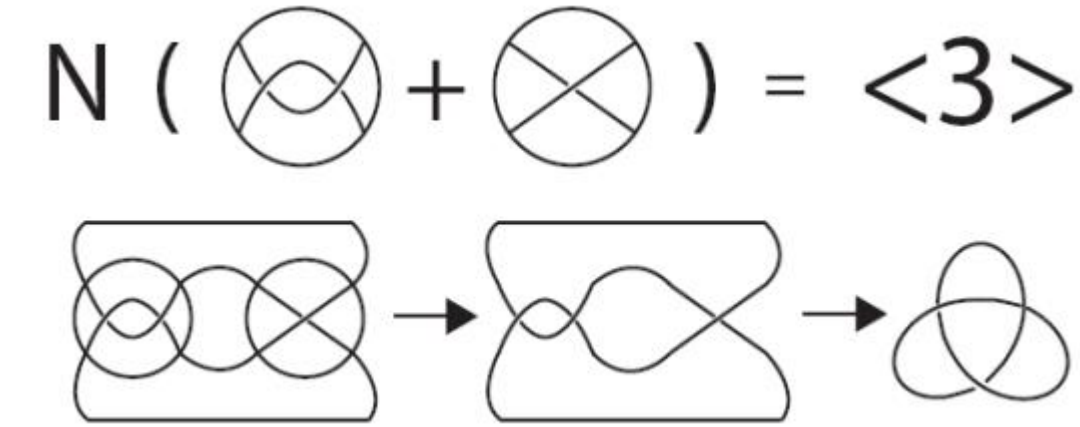

Figure 43. Tangle Operation-4

6.3. 4-Plat

A 4-plat is a knot obtained by knitting four strands and tying the ends as shown in the figure below. 4-plats are generally known as 2-bridged rational knots. All prime knots with less than eight crossovers and two-component principal chains with less than seven crossovers are 4-plat. 4-plat knots can be expressed by integer vectors like typical rational tangles. 4-plat vectors are vectors with an odd number of components. Namely;$\langle c_1, \ldots, c_{2k+1} \rangle$ i for each i is $c_i \geq 1$, and where each integer component strings represent one half shift. So, as with rational tangles, vectors can also be used to draw a 4-plat diagram. In doing so, the following way is taken: starting with four strands, half-shift is done as much as c_1 on the middle fiber, then half-shift as much as c_2 on the top two strands, then the same on the middle two strands is done and this process continues until all integer components in the vector are finished. Finally, the endpoints as shown in the figüre are to be connected. This expression of 4-plat is called the Conway symbol and corresponds to the minimal diagram of the 4-plat.

Theorem 6.3.1. Two 4-plat is equal if and only if they have the same conway symbols, or if one has a Conway symbol that is exactly the opposite of the other, i.e. while one of them has the Conway symbol $\langle c_1, \ldots, c_{2k+1} \rangle$ the other one has a Conway symbol of $\langle c_1, \ldots, c_{2k+1} \rangle$

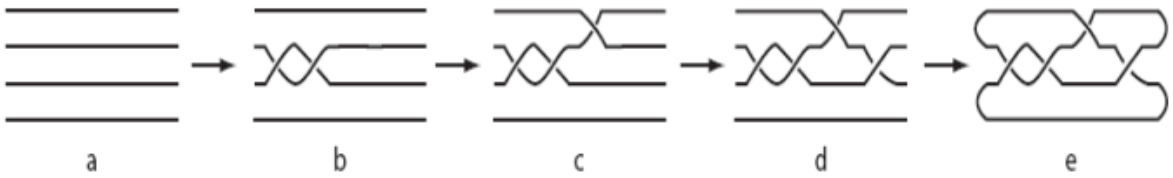

Figure 44. Plat drawing

Note: The Conway symbol is used to calculate the rational number β α with 0 <β <α.

$\frac{\beta}{\alpha} = \frac{1}{c_1 + \frac{1}{c_{2+\cdots}}}$

This is how it is calculated. 4-plat β The number α is denoted as b (α, β).

Theorem 6.3.2. The two 4-plat b(α, β) and b(α',β′) are equivalent if and only if α = $\alpha'$and $\beta^{\pm 1} \equiv \beta'$ (modα)

For example; If b is examined (17, 5), b (17, 7) 4-plat, b (17, 5) (3, 2, 2) e corresponds to b (17, 7) to (2, 2,3). As a result, it is seen that these two 4-plates are identical. It can be easily seen that they are equivalent because they are already 17 = 17 and $5^{-1} \equiv 7$ (mod17).

Rational tangle and 4-platers are quite similar in terms of the use of rational numbers. If the given rational number $\frac{\beta}{\alpha}$is in the range 0 < $\frac{\beta}{\alpha}$ <1, the denominator closure of the rational tangel $\frac{\beta}{\alpha}$ gives b (α, β) 4-plat, and if the given number$\frac{\beta}{\alpha}$ is in the range $\frac{\beta}{\alpha} \geq 1$, then the numerator closure of $b(\beta, -\alpha)$ rational tangle gives the 4-plat. For any integer x, D (($d_1, \ldots\ d_{2k+1}, x$)) = $\langle d_1, \ldots, d_{2k+1} \rangle$ and N (($d_1, \ldots\ d_{2k+1}, x$,0)) = $\langle -d_1, \ldots, d_{2k+1} \rangle$. As mentioned earlier, the numerator closure of the sum of two rational tangles was a 4-plat. The next theorem will give information about the equivalence of rational knots obtained by the numerator closure of rational tangles.

Theorem 6.3.3. Let's take two rational tangles with reduced fractions $\frac{p}{q}$and $\frac{p^\iota}{q^\iota}$.If N ($\frac{p}{q}$). and N ($\frac{p^\iota}{q^\iota}$) , which are the rational knots obtained as a result of the numerator closure of tangles, correspond to each other, N ($\frac{p}{q}$) and N ($\frac{p^\iota}{q^\iota}$) are topologically equivalent if and only if $p = p^\iota$and q ± 1 ≡ q' (modp).

6.4. Solving Tangle Equations

As we have seen in previous chapters, the tangle equations were N (A + B) = K, with K knot or chain, A and B tangle. Solving these equations will help us better understand enzyme mechanisms.

Lemma 6.4.1. Let two rational tangles $A_{1=}\ \frac{\beta_1}{\alpha_1}$ and $A_2\ = \frac{\beta_2}{\alpha_2}$ The operation N $(A_1 + A_2)$ defines a 4-plat in the form b (α, v β) and α = | $\alpha_1\beta_2 + \alpha_2\beta_1$| and β is defined as follows:

(a) If α = 0, then β = 1;

(b) If α = 1, then β= 1;

(c) If α> 1, β is obtained as follows: where 0 <β <α and σ = sign $(\alpha_1\beta_2 + \alpha_2\beta_1)$ and $\alpha_2' and\ \beta_2'$ tangle of α 2 and $\frac{\beta_2}{\alpha_2}$ are components of column $\beta \equiv \sigma(\alpha_1\alpha_2' + \beta_1\beta_2')$(modα).

Example 6.4.1: Let's take $A_1 =$ 2 and A_2 = 23/17. α = | 1 × 23 + 17 × 2 | = 57. In the example above we found the matrix of tangle β / α = 23/17 and it looks like this:

$$\begin{bmatrix} u & v' \\ v & u' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 23 & 4 \\ 17 & 3 \end{bmatrix}$$

We can find the values of α_2' and β_2' from here.$\beta \equiv (1 \times 4 + 2 \times 3)\ (\text{mod}57) = 10$. As a result, the result of the operation is N $(A_1 + A_2)$ = b (57, 10).

Theorem 6.4.2. Let A = $\frac{\beta}{\alpha}$= $(\alpha_1, \ldots\ \alpha_{2n})$ be a rational tangle and K = $(\langle c_1, c_2, \ldots c_{2k+1}\rangle$ a 4-plat. The rational tangle solution of the equation N (X + A) = K ≠ ⟨0⟩, X = $(c_1, \ldots, c_{2k+1}, r, -a_1, \ldots, -a_{2n})$ where r is any integer or X = $(c_{2k+1}, \ldots, c_1, r, -a_1, \ldots, -a_{2n})$. If K = ⟨0⟩ then X = $(-a_1, -a_2, \ldots -a_{2n})$ is the only solution.

Theorem 6.4.3. A1 and A2 be two different rational tangles and $K_1$and K_2 be 4-plat. The equations N (X +A_1) = K_1 and N (X + A_2) = K_2 have at most two different rational tangle solutions.

Proof ; Let ; X = $\frac{U}{V}$, $A_1 = \frac{\beta_1}{\alpha_1}$, $A_2 = \frac{\beta_2}{\alpha_2}$, $K_1 =$ b (α, β) and $K_2 =$ b (α', β'). From the lemma7.4.1 above it is found that α = | $v\beta_1 + \alpha_1 u$| and α '= | $v\beta_2 + \alpha_2 u$ |. In the (u, v) -plane, these equations denote two parallel pairs. Since u / v = −u / −v, these four points represent at most two different rational tangles for this system of equations.

Example 6.4.2. Let $A_1 = \frac{1}{3}$, $A_2 = \frac{5}{17}$, $K_1 = b(5,3)$ and $K_2 = b(29,17)$

$$| v + 3u | = 5$$

$$| 5v + 17u | = 29$$

The solution of this system of equation:

$$v + 3u = 5$$

$$5v + 17u = -29$$

From here we get the solution X = - $\frac{27}{86}$. Another solution is:

$$v + 3u = -5$$

$$5v + 17u = 29$$

From here, X = - $\frac{27}{86}$.

6.5. Tangle Model

Introduced by DeWitt Sumners in 1980, the purpose of the tangle model is to express the events occurring during recombination mathematically. In this way, we can express what the enzyme does topologically and geometrically of the DNA product and its substrate. It can be seen in electron micrographs that the DNA lipheres are wrapped around each other. Since 4-plat and rational tangles consist of twisted strands, they are very suitable candidates for DNA modeling. If we recall the definition of Tangle, it was the non-directional arc pair of t and the (B, t) pair embedded in B where B is the 3-sphere. A tangle can be used in the modeling of the enzyme-DNA complex as follows; Enzyme 3-sphere and two recombination sites with two strands. The most observed product of recombination is 4-plat, which is quite plausible because we can model the enzyme-DNA complex with 4-plat and express the changes with tangle equations.

However, before expressing the enzyme mechanism with the tangle model, we must make a few assumptions. Our first assumption is that we will express the enzyme-DNA complex as the sum of tangles. Let E be the enzyme, the part of O_b DNA that binds to the enzyme, and P the part that changes during the reaction. Therefore, we can express the enzyme-DNA complex as $E = O_b$ +P. Of course, we will also need a DNA that is not bound to an enzyme. Here, the expression of this form of DNA with tangle will also be Of. Now we get the equation $N(O_f + O_p + P) = K_0$ tangle, and this gives us the substrate molecule. Our second assumption is that we express the form of the recombination P region tangle after being rotated by recombination with the R recombinant tangle. With this assumption, a recombination transforms the P region tangle into the R recombinant tangle. Let us express the model after the recombination spin below:

$$N(O_f + O_p + P) = K_0 \text{ (substrat)}$$

$$N(O_f + O_p + P) = K_0 \text{ (product)}$$

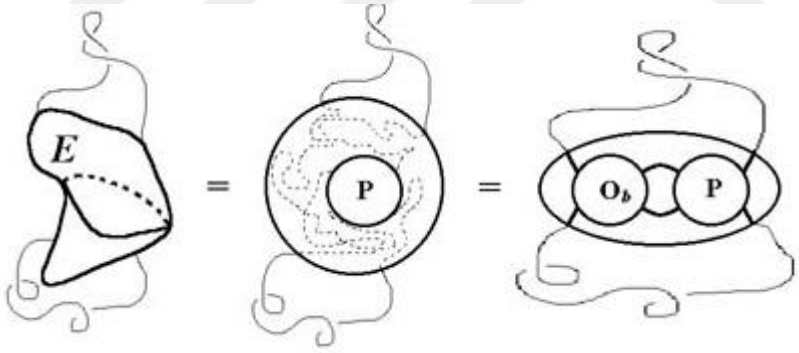

Figure 45.

Also, we should not forget that; The recombination mechanism is fixed, free of substrate geometry and topology. It means; So long as all substrate molecules are in the same knot kind, the tangles O_f, O_b, P, and R do not transform from one action to the next. In the event that the substrate molecules are distinct kinds of knots, only the O_f tangle changes. However, in this case, the only exception to consider is site orientation. Our final assumption is the processive recombination model given by the tangle system of

equations:

If O =$O_f + O_b$ and O, P and R are unknown

N (O + P) = K_0 (substrate)

N (O + R) = K_1 (product resulting from the first spin)

N (O + nR) = K_n (the product resulting from n. spin) emerges

6.6. Sample

The tangle model was employed in 2002 by Mariel Vazquez and De Witt Sumners Gin to examine specific-site recombination. We shall discuss their inventions in this part. This is only one instance of specific-site recombination using the tangle model. A bacteriophage known as Mu encodes the specific-site recombination procedure known as Gin. Infecting bacteria are viruses called bacteriophages. Gix L and Gix R are two recombination sites found in the phage genome. Gin splits both sides after one binds to the DNA, guiding the terminals and joining them. Gin works by creating several recombinations during double bonding, a mechanism known as processive recombination. The tangle study of the gin recombination on the unknot substrate molecule yielded the following results, which are inversely reproduced in the gix regions:

$$K_0 = \langle 1 \rangle \text{(knotless)}$$

$$K_1 = \langle 1 \rangle \text{ (knotless)}$$

$$K_2 = \langle 3 \rangle = 3_1 \text{ (trefoil knot)}$$

$$K_3 = \langle 2,1,1 \rangle = 4_1 \text{ (8-figure knot)}$$

$$K_4 = \langle 2,2,1 \rangle = (5 - twist\ knot)$$

In 2004, De Witt Sumners and Mariel Vazquez discovered a result that gives the solution to the above four equations and predicted exactly the fifth equation.

Theorem 6.6.1. (O,R) which is he solution of the following system of equations for tangels O, P, R

(1) $N(O + P) = \langle 1 \rangle$ =knotless

(2) $N(O + R)$=⟨1⟩ = without knots

(3) N$(O + R + R)$=⟨3⟩ = trefoil node

It is either ((-2, 0), 1) or ((4,1), (- 1)). Also if

(4) N$(O + R + R + R)$= ⟨2,1,1⟩ = 8-figure knot

There is only one solution in the form of (O, R) = ((−2, 0), (1)). We have not checked the rationality of O and R here, but instead examined how the tangle equation is solved. In a lemma given earlier, when two rational tangles like , $A_1 = \frac{\beta_1}{\alpha_1}$ ve $A_2 = \frac{\beta_2}{\alpha_2}$ are given and α = $|\alpha_1\beta_2 + \alpha_2\beta_1|$ is considered, we found that N $(A_1 + A_2)$ and b (α, β) 4-plat is equal. From the equations (2) and (3) given above, we can obtain the following system:

| u + rv | = 1

|yu + 2rv | = 3

u, r, v are unknown values. We can obtain ten different solutions for the ordered pair (u / v, r). Thus, we get ten different solutions for the (O, R) tangle pair. These solutions are ((−2, 0), (1)), ((1), (- 2)), ((5), (- 4)), ((- 2, −2), (2)), ((4, 1), (- 1)) and their mirror reflections. With the help of the next theorem, we can eliminate some of these results.

Theorem 6.6.2. The tangles given in the equations (1), (2), (3) and given in the next theorem come from Gin recombination with inversely repeated regions. And these provide the features we give below:

O ≈ (0, 0), R ≈ (1), P ≈ (0).

Since O ≈ (0, 0) and integral tangle (0) has parity with (1), we can ignore the results obtained for the integral tangle O. Additionally, if R ≈ (1), we can also get rid of the solution R = (2) since integral tangles have (0) parity. We can also ignore mirror reflections since the knot product of equation (3) is chiral. Thus, we only have two solutions left, and only one of them satisfies equation (4).

In light of this tangle analysis, when Sumners and Vazquez act on a substrate with the Gin gix regions, the enzyme mechanism adds a positive cross-over to the substrate at the turn corresponding to each recombination.

Theorem 6.6.3. (O,R) which is he solution of the following system of equations for tangels O, P, R

(1) N (O + P) = ⟨1⟩ = knotless

(2) N (O + R) =⟨3⟩ = trefoil knot

(3) N (O + R + R) = (⟨1,2,1,2⟩ = (- 5) twist knot

Either ((−2, 0), (2)) or ((2, 1, 1, 2), (- 2)). In addition if

(4) If N (O + R + R + R) = (⟨1,4,2⟩)= (- 7) is a twist knot, then (O, R) = ((−2, 0), (2)) and

(5) for every $n \geq 4$, N (O + nR) = - (2n + 1) becomes twist knots.

In this example, the tangle model is used to represent the structure of the Gin mechanism mathematically. And as a result, this shows us; inversely repeating regions of recombination adds tangle (1), in other words R = (+1) In the regions that are directly repeated, R = (+2) [80].

REFERENCES

[1] Goldblum, J. 1953 the race for the double helix. (http://www2.nau.edu/lrm22/lessons/dna_notes/dna_notes.html)

[2] Watson, J.D.Crick F.H.C., 1953. A Structure for Deoxyribose Nucleic acid, 737/738.

[3] Shure, M.Vinograd, J., 1976. The number of Superhelical Turns in Native Virion SV40 and Minicol DNA Determined by the Band Counting Method, 215/226.

[4] C˘alug˘areanu, G. 1961. Sur las Classes D'isotopie des Noeuds Tridimensionnels et Leurs Invariants, Czech Math. 588/625.

[5] White, J.H., 1969. Self-linking and the Gauss Integral in Higher Dimensions, Amer. Math. 693/728.

[6] Fuller, F.B. 1971. The Writhing Number of a Space Curve, Proc. Natl. Acad. Sci. USA, 815/819.

[7] Adams, C. 2008. Introduction to Knot Theory, Applications of Knot Theory: Proceedings of Symposia in Applied Mathematics.

[8] Rolfsen, D. 1971. Knots and Links, Publish or Perish Press, Berkeley, CA.

[9] Porteous, I.R.2001. Geometric Differentiation for the Intelligence of Curves and Surfaces, Cambridge University Press, Cambridge, England,

[10] Dennis, M.R. , Hannay J.H.2005. Geometry of C˘alug˘areanu's theorem, Proc. Roy. Soc. A 461, 3245/3254.

[11] Weber, C.2002. O`u en sont les math´ematiques?, ch. Questions de Topologie en Biologie Mol´eculaire. Appendix by Andrzej Stasiak and Jacques Dubochet, pp. 328/347, Vuibert.

[12] William, F.1981. DNA and Differential Geometry, Math. Intelligencer 3 no. 1, 20/27. MR617886 (83d:92120)

[13] Cozzarelli ,N.R, Christian-Boles, T., White J.H.,1990. Primer on the Topology and Geometry of DNA Supercoiling, DNA Topology and its Biological Effects, 139/184.

[14] Shore, D. Baldwin, R.L.1983. Energetics of DNA Twisting: I. Relation Between Twist and Cyclization Probability, J. Mol. Biol. 170, 957/981.

[15] Depew ,R.E. , Wang J.C.1975. Conformal Fuctuations of the DNA Helix, Proc. Natl. Acad. Sci. USA 72, no. 11, 4275/4279.

[16] Pulleyblank, D.E. Shure, M. Tang, D. Vinograd, J. Vosberg, H.P. 1975.Action of Nickingclosing Enzyme on Supercoiled and Nonsupercoiled Closed Circular DNA:

Formation of a Boltzmann distribution of topological isomers, Proc. Natl. Acad. Sci. USA 72 no. 11, 4280/4284.

[17] Horowit , D.S. Wang, J.C 1984.Torsional Rigidity of DNA and Length Dependence of the Free Energy of DNA Supercoiling, J. Mol. Biol. 173 , 75/91.

[18] Maxwell,A. Bates, A. D.2005.DNA topology, Oxford University Press,

[19] Froelich-Ammon , S.J. Osheroff, N. 1990.Topoisomerases Poisons: Harnessing the Dark Side of Enzyme Mechanism, J. Biol. Chem. 270, 21429/21432.

[20] Levine, C. , Hiasa, H. , Marians, K.J. 1998.DNA Gyrase and Topoisomerase IV: Biochemical Activities, Physiological Roles during Chromosome Replication and Drug Sensitivities. Biochem. Biophys. Acta 1400 29/43.

[21] Rybenkov, V.V. Ullsperger, C. Vologodskii, A.V. , Cozzarelli, N.R.1997. Simplification of DNA Topology Below Equilibrium Values by type ii Topoisomerases, Science 277 ,no. 5326, 690/3.

[22] Berger, J.M.1999.Type II topoisomerases, Curr. Opin. Struct. Biol. 8 ,26/32.

[23] Dong, K.C. , Berger, J.M.2007. Structural Basis for gate-DNA Recognition and Bending by type II A Topoisomerases, Nature450 1201/1206.

[24] Buck G. Zechiedrich E.L.2004. DNA Disentangling by Type-2 Topoisomerases, J. Mol. Biol. 340 .

[25] Vologodskii, A.V.Zhang,W. Rybenkov,V.V. Podtelezhnikov,A.A. D. Subramanian, J.D. Griffith, Cozzarelli, N.R.2001. Mechanism of topology simplification by type ii DNA topoisomerases, Proc. Natl. Acad. Sci. USA 98 , 3045/3049.

[26] Rasmussen, J.A.2004 Khovanov homology and the Slice Genus, arXiv:math/0402131v1 [math.GT]

[27] Lobb, A.A.2007.slice genus lower bound from sl(n) khovanov-rozansky homology, arXiv:math/0702393 .

[28] Ozsv´ath ,P. , Szab´ o, Z.2005. Knots with unknotting number one and heegaard floer homology, Topology 44 , 705/745.

[29] Kanenobu ,T. Murakami, H.1986. Two-bridge Knots with Unknotting Number One., Proc. Am. Math. Soc. 98 , 499/502.

[30] Kohn, P. Two-bridge Links with Unlinking Number One , 1991.Proc. Am. Math. Soc. 113 . 1135/1147.

[31] Torisu,I.1998. The Determination of the Pairs of TwoBbridge Knots or Links with Gordian Distance One, Proc. Am. Math. Soc. 126 , 1565/1571.

[32] Isabel K. 2001.Darcy, Biological Distances on DNA Knots and Links: Applications to XER Recombination, J. Knot Theory Ramifications 10 (2001), no. 2, 269/294, Knots in Hellas ’98, Vol. 2 (Delphi).

[33] Khovanov, R.2007. Homology of Two-Bridge Knots and Links., Duke Math. J. 136 551/583.

[34] Stasiak, A., Flamini, A. 2004.Simulation of Action of DNA tTopoisomerases to Investigate Boundaries and Shapes of Spaces of Knots, Biophys. J.87 , 2968/2975.

[35] Cozzarelli, N.R. ,Krasnow,M.A Gerrard,S.P. , White,J.H.1984. A Topological Treatment of Recombination and Topoisomerases, Cold Spring Harb. Symp. Quant. Biol. 49 , 383/400.

[36] Pathania,S. Jayaram,M. Harshey,R.M.2002. Path of DNA Within the Mu Transpososome. Transposase Interactions Bridging Two Mu Ends and the Enhancer Trap Five DNA Supercoils, Cell 109 , no. 4, 425/36.

[37] Wasserman, S.A. Dungan,J.M. Cozzarelli, N.R.1985. Discovery of a Predicted DNA Knot Substantiates a Model for Site-Specific Recombination, Science 229 , no. 4709, 171/4.

[38] Heichman, K.A. Moskowitz, I.P, . Johnson, R.C.1991.Configuration of DNA strands and mechanism of strand exchange in the Hin invertasome as revealed by analysis of recombinant knots, Genes Dev 5 , no. 9, 1622/1634.

[39] Merickel,S.K. Johnson, R.C .2004.Topological Analysis of Hin-catalysed DNA Recombination in Vivo and in Vitro, Mol. Microbiology 51 , no. 4, 1143/1154.

[40] Kanaar, R. Klippel, A. Shekhtman, E. J. Dungan,M. Kahmann,R. , Cozzarelli,N. R. 1990.Processive Recombination by the Phage Mu Gin system: Implications for the Mechanisms of DNA Strand Exchange, DNA Site Alignment, and Enhancer Action, Cell62 . no. 2, 353/66.

[41]Crisona, N.J. Kanaar, R.T. Gonzalez,N. Zechiedrich,E.L. Klippel, A. Cozzarelli,N.R.1994. Processive Recombination by Wild-Type Gin and an Enhancer-Independent Mutant. Insight into the Mechanisms of Recombination Selectivity and Strand Exchange, J. Mol. Biol.243 , no. 3, 437/57.

[42] Hoess, R. Wierzbicki,H. A. , Abremski,K. 1986.The Role of the loxP Spacer Region in P1 Site-Specific Recombination, Nucleic Acids Res. 14 , no. 5, 2287/300.

[43] Liu, L.F. Davis ,J.L.1981. Novel topologically knotted DNA from bacteriophage p4 capsids: studies with topoisomerases, Nuc. Acids Res. 9 , 3979/3989.

[44] Krasnow,M. A. Stasiak, A. S. Spengler, J. F. Dean, T. Koller, Cozzarelli,N. R. 1983.Determination of the Absolute Handedness of Knots and Catenanes of DNA, Nature304 (1983), no. 5926, 559/60.

[45] Trigueros,S. Arsuaga, J.M.E. Vazquez, D.W. Sumners, Roca,J. 2001.Novel Display of Knotted DNA Molecules by Two-Dimensional Gel Electrophoresis, Nucleic Acids Res. 29 , no. 13, E67/7.

[46] Zechiedrich E.L.Crisona,N.J.1989. Methods in Molecular Biology: DNA Topoisomerase Protocols, M. Bjornsti and N. Osheroff (eds), vol. 1, ch. Coating DNA with RecA Protein to Distinguish DNA Path by Electron Microscopy, pp. 99/108, Humana Press, Totowa, NJ,

[47] Katritch, V. Bednar, J. Michoud, D. Dubochet, J. Stasiak, A. 1996.Electrophoretic Mobility of DNA Knots, Nature 384 , no. 6605/122.

[48] Sundin O. Varshavsky A.1980. Terminal Stages of SV40 DNA Replication Proceed via Multiply Intertwined Catenated Dimers, Cell21 (1980), no. 1, 103/14.

[49] Levene, S.Tsen,H. 1999.Analysis of DNA Knots and Catenanes by Agarose-gel Electrophoresis, Methods Mol. Biol. 94 , 75/85.

[50] Hoste,J. Thistlethwaite, M.Weeks,J.1998. The first 1,701,936 knots, Math. Intelligencer 20 , no. 4, 33/48.

[51] Grosberg,A. Feigel, Y. Rabin,A. Y. 1996. Flory-Type Theory of a Knotted Ring Polymer, Phys. Rev. E Sta.t Phys. Plasmas Fluids Relat. Interdiscip. Topics 54 , no. 6, 6618/6622.

[52] Buck, D.2009. Advanced School and Conference on Knot Theory and its Applications to Physics and Biology

[53] Vologodskii,A.V. Crisona,N.J. B. Laurie, Pieranski, P. Katritch, V. J. Dubochet, A. Stasiak,1998. Sedimentation and Electrophoretic Migration of DNA Knots and Catenanes, J. Mol. Biol. 278 , no. 1, 1/3.

[54] Cantarella,J. Robert B. Kusner, J. Sullivan,M.2002. On the minimum Ropelength of Knots and Links, Invent. Math. 150 no. 2, 257/286. MR1933586 (2003h:58014)

[55] Wasserman S.A. Cozzarelli,N.R. 1991.Supercoiled DNA-Directed Knotting by T4 Topoisomerase, J. Biol. Chem. 266 , no. 30, 20567/73.

[56] Feil,R.2007. Conditional Somatic Mutagenesis in the Mouse using Site-specific Recombinases, Handbk. Exp. Pharmacol. 178 , 3/28.

[57] Buck, D. Flapan, E.2007. A Topological Characterization of Knots and Links Arising from Sitespecific Recombination, J. Mol. Biol. 374 , 1186/1199.

[58] Craig, N. Craigie, R. Gellert, M. Lambowitz A.2002. (eds.), Mobile DNA, ASM Press, Washington, DC,

[59] Grindley,N. D. Whiteson, K. L. Rice, P. A.2006. Mechanisms of Site-specific Recombination, Ann. Rev. Biochem.no 75 , 567/605.

[60]Stark,W. M.Boocock,M.R. Sherratt,D.J.1992. Catalysis by Site-specific Recombinases, Trends Genet 8 , 432/439.

[61] Leschziner A.E. Grindley, N.F. 2003.The architecture of the γδ Resolvase Crossover Site Synaptic Complex Revealed by using Constrained DNA Substrates., Mol. Cell12 , 775/ 781.

[62] Nollmann, M. J. He, O. Byron, W.M. Stark,2004. Solution Structure of the Tn3 Resolvasecrossover Site Synaptic complex, Mol. Cell16 , no. 1, 127/137.

[63] Smith M.C. Thorpe, H.M.2002. Diversity in the Serine Recombinases, Mol Microbiol 44 , no. 2, 299/307.

[64] Altschul, S.F. Madden,T.L. Schaffer, A.A. Zhang,J. Zhang, Z. W. Miller, D.J. Lipman,1997 Gapped Blast and Psi-blast: A New Generation of Protein Database Search Programs., Nucleic Acids Res 25 , 3389/3402.

[65] Leschziner A.E. Grindley, N.D.F. 2003.The Architecture of the γδ Resolvase Crossover Site Synaptic Complex Revealed by using Constrained DNA Substrates., Mol. Cell12 , 775– 781.

[66] Grainge I. Jayaram,M. 1999.The Integrase Family of Recombinase: Organization and Function of the Active site, Mol. Microbiol. 33 , no. 3, 449/56.

[67] Biswas,T. Aihara, H. Radman-Livaja, M. D. Filman, A. Landy, T. Ellenberger,2005. A Structural Basis for Allosteric Control of DNA Recombination by λ integrase, Nature435 , 1059/1066.

[68] Grainge, I. D. Buck,M. Jayaram,2000. Geometry of Site Alignment during Int family Recombination: Antiparallel synapsis by the Flp Recombinase, J. Mol. Bio.298 . 749/764.

[69] Boles,T.C. White, J.H. Cozzarelli,N.R. 1990.Structure of Plectonemically Supercoiled DNA, J. Mol. Biol. 213 , no. 4, 931/51.

[70] White, J.H. Cozzarelli,N.R.1984. A simple Topological Method for Describing Stereoisomers of DNA Catenanes and Knots, Proc. Natl. Acad. Sci. USA 81 , no. 11, 3322/6.

[71] White, J.H. Millett,K.C. Cozzarelli, N.R.1987.Description of the Topological Entanglement of DNA Catenanes and Knots by a Powerful Method Involving Passage and Recombination, J. Mol. Biol. 197 , no. 3, 585/603.

[72] Tommy K.2007.A topological characterization of knots and links arising from site-specific recombination, J. Phys. A40 12377/12395.

[73] Ernst, C. D. Sumners, W.1987. The Growth of the Number of Prime Knots, Math. Proc. Cambridge Philos. Soc. 102 (1987), no. 2, 303/315.

[74] Hoste,J. Thistlethwaite, M. Weeks,J.1998.The First 1,701,936 Knots, Math. Intelligencer 20 , no. 4, 33/48.

[75] Silver, D.S.2006. Knot Theory's Odd Origins, American Scientist 94 , 158/165.

[76] Murasuhi K.1996. Knot Theory and its Applications , 7/14

[77] Summens, W.2000. DNA Topology: Experiments and Analysis , 5/8

[78] M. Bio.2016.dna topology.(https://www.youtube.com/watch?v=HyP0cEbqKTc) (May 2016)

[79] Mucuk O. Topoloji ve Kategori (2011), 251

[80] Karaca İ. Geometrik Topoloji Ders Notları 153-164

CURRICULUM VITAE

1. Adı Soyadı: Havva Nur KARABAY

Lisans	Erciyes Üniversitesi Fen Fakültesi	Matematik	2010-2013
Lise	Seydişehir Enis Şanlıoğlu Anadolu Lisesi		2006-2010

5.Sertifikalar:

Erciyes üniversitesi eğitim fakültesi pedogojik formasyon 2014

6.İş Tecrübesi:

Verimli Temel Lisesi-Matematik Öğretmeni 2016-2018

Artı Matematik Eğitim Kurumları Matematik Öğretmenii 2018-2019

İdeal Eğitim Kurumları Matematik-Geometri Öğretmeni 2020-