



UNIVERSIDAD DE MURCIA
FACULTAD DE MATEMÁTICAS

Trabajo de fin de grado

GEOMETRÍA Y TOPOLOGÍA DEL ADN Y SU APLICACIÓN AL DISEÑO DE FÁRMACOS

Silvia Díaz Belchí

Supervisado por:
José Antonio Pastor

23 de septiembre de 2015

Índice general

Abstract	V
Resumen	IX
1. Biología del ADN	1
1.1. Estructura primaria, secundaria y terciaria del ADN	2
1.2. Replicación, transcripción, y recombinación	5
1.3. Modelización en laboratorio	6
2. Geometría y topología del ADN y sus invariantes	7
2.1. Preliminares	7
2.2. Invariantes del ADN	11
2.3. Fórmula de White	20
2.4. Interpretación de la fórmula de White	32
3. Cambios en la geometría y topología por la acción de enzimas, y diseño de fármacos.	35
3.1. Topoisomerasas	36
3.2. Recombinación. Recombinación específica de sitio	37
3.3. Recombinasas	40
3.4. Enfermedades y diseño de fármacos	40
A. Teoría de nudos	43
A.0.1. Introducción	43
A.0.2. Nudos y primeras definiciones	44
A.0.3. Equivalencia de nudos	46
A.0.4. Nudos en biología	49
A.0.5. Software TangleSolve	53
Bibliografía	53

Abstract

As you can guess by the title, this paper is about to study the topology and geometry of the DNA's molecule. Nowadays there are many people who know what is DNA, who know that it is *something* used to identify and differentiate ourselves from others. Scientists (biologists, chemists, doctors and nurses, etc.), meanwhile, know more about this molecule, they know it's not a simple box in which the genetic material is stored, and that the genetic material is encoded. Moreover they know that the DNA's molecule carries out important functions within the cell that harbours it. However, many of them are unaware of other things, such as why a result is obtained, or another, after one of the above functions. Then the researchers are who study the changes undergone by the DNA after a certain reaction to try to predict the behaviour of the DNA's molecule in similar situations.

These changes that are studied by the researchers, although they unknown it, vary the geometry and the topology of the DNA's molecule. Therefore, it is at this point where maths, and the topic of this work, come into play.

The text is divided into three chapters and an appendix, which are preceded by a summary in Spanish and an abstract in English. This summary will provide us an overview of the work that will help us to understand better the general ideas of work allowing us to have a comprehensive understanding about it.

The first chapter is called *Biology of the DNA*. This chapter, in turn, is divided into four sections. In the first section of the chapter there is a brief definition of what is DNA and what its basic functions. After this definition of just over four lines, we find a historical introduction about the DNA's discovery and how it was unraveling its structure.

In this section appear names as Friedrich Miescher, who was the first person who isolated the DNA's molecule; Levene, who basically decoded that the DNA was a solenoid (or spring)-shaped; Chargraff, which provided important results after their experiments; the universally well-known, Watson and Crick, who clarified that the DNA has a double helix structure, winning a Nobel for that reason; and the character that remains in the shade on most occasions, but without their contributions the work of Watson and Crick had perhaps been difficult, Rosalind Franklin. However, the highlight of this section are the dates, which show that the discovery of the DNA's molecule is too early.

In the next section we talk about the structure of DNA, so three levels of complexity are distinguished. The first structure of DNA, which refers to the lineal arrangement

of the nucleotides, which are the basic units forming the molecule; secondary structure, which refers to the double helix model presented by Watson and Crick's model; and the tertiary structure, which refers to the spatial arrangement can have this double helix, appearing the term *supercoiling*, which will play a key role in this work.

It is followed by section that discusses the basic functions performed by the DNA within the molecule in which it is located. Thus appearing, the replication, the transcription, and the recombination. It is in these processes that DNA undergoes changes, changes that accept a precise mathematical description.

Finally, to conduct any type of investigation, it is necessary to work in a laboratory. To study the changes that have been mentioned, is essential to work in the lab with circular DNA's molecules, since circular DNA is restricted from the topological point of view, facilitating the study and observation.

The second chapter is entitled *Geometry and topology of DNA, and its invariants*, and the main purpose of this chapter is to prove the formula of White. The proof of this formula is not easy, and it requires handling some previous concepts and results. It is for this reason that in the first section of this chapter are introduced the differential forms, the pull-back of a function, and the concept of exterior derivative, as well as some propositions that set properties and relationships between these concepts.

In the following section are defined the number of link, denoted by Lk defined; the number of twists and turns, denoted by Tw ; and the number of winding, denoted by Wr , mathematically, that is to say, in terms of differential geometry and by an integral. For this one of the strands of the double helix, denoted by C , is parameterized by the arc length by the closed curve α . Considering then V as the field vector perpendicular to the curve α , the other strand of the double helix is denoted by $C_V := C + V$, and parameterized by the arc length by the closed curve β . In this context, then it considers the trihedral of Frenet, and the above concepts are defined as follows:

$$Lk(C, C_V) := \frac{1}{4\pi} \iint_{C \times C_V} e^*(d\Sigma)$$

$$Tw(C, V) := \frac{1}{2\pi} \int_{[0, \ell]} \langle V^\perp, V' \rangle ds$$

$$Wr(C) := \frac{1}{4\pi} \iint_{C \times C} e^*(d\Sigma)$$

where e is the application in the sphere given by: $e(x, y) := \frac{y - x}{|y - x|}$, and $*$ denote the pull-back.

Given these concepts, it is then in the third section where the proof of the formula of White appears, which establishes a relationship between the number of link, a topological number; and the number of twisting and winding, which are geometrical quantities. This relationship is as follows:

$$Lk(C, C_v) = Tw(C, V) + Wr(C)$$

To prove this formula has made use the article [11], that uses others definitions of these parameters. It is for this reason that in this section the equivalence between the above definitions and the definitions provided by the author of the article is studied.

Finally, to finish the chapter, the following section contains some result that can be obtained from the formula of White, studying how to vary the parameters Tw and Wr , depending how does Lk . In this way, the main consequence that we can extract of that theorem is the study of the phenomenon of supercoiling.

In the last chapter, whose title is *Changes in the geometry and the topology by the action of the enzymes and design of drugs*, some enzymes, whose action on DNA is important, are described, and we talk about the other part of the title of work, that is, on drug's design.

The main protagonist of this chapter is supercoiling, because the effect of all the enzymes that have been described influences directly the state of supercoiling that may have the DNA's molecule. Because of this, the chapter begins with an introduction on the phenomenon.

The first enzyme that is explained is topoisomerase. These enzymes alter the topology of DNA by cutting and pasting strands, either to further entangle order to achieve greater packing, either to unroll to carry out other DNA's functions. This group of enzymes is divided into two types, topoisomerase type I and type II, depending on the number of strands in each round of cutting action.

The next section of the chapter does not speak of any enzyme. In this case appear the concept of site-specific recombination. This type of recombination is carried out on a specific site or area of the DNA's molecule, therefore is the method used in the laboratory for the study of the products obtained after the reaction of an enzyme on DNA. To describe it, previous concepts as synaptosome recombination, synapses, substrate and product, and recombinase, which is the type of enzyme that acts in this type of recombination, are required. Due to the emergence of the recombinases, the next section is dedicated to it for completeness.

The chapter closes with a section in which some known diseases, such as HIV, cancer or tuberculosis, can be treated with a specific enzyme that acts on the DNA of the microorganism (virus or bacteria) responsible for the disease, paralyzing its reproduction.

This is the point where all concepts that had been studied previously are applied, because in the lab working with a circular DNA's molecule, to which it attaches a specific enzyme in a specific area of it. A synaptic complex is created, which can not be determined by direct observational methods, and the product obtained is studied by observing the topological and geometrical changes. Thus, we can predict what happened during the site-specific recombination, predicting the mode of action of the enzyme. In this way, when we want to design a potential drug, we can know what kind of enzyme we add in order to obtain the desired result.

In the laboratory the study of the action of an enzyme on a DNA's molecule is based in the study of the product of the reaction that occurs, and the substrate (or part of the DNA that is exposed to the reaction) initial. But to study this it is necessary to rely on what is known as *Knot theory*, a relatively young branch of topology.

It is this theory which is develops, briefly, in the only Appendix of the work. Knot theory is a long and difficult theory, which is still developing, therefore, to learn more about it, you can see [6], [7], [8], and [9].

The first paragraph of the appendix provides us a brief historical introduction about of the *birth* of this theory, highlighting the fact that the possible *father* of it is Gauss, working with induced magnetic fields. However, others important names like Alexander, Reidemeister, and Conway appear. It should also be said that the main objective of this theory is to determine when two nodes or links are equivalent.

The following section seeks to define mathematically the concept of knot and link. To try to reach the aim of this theory we should define the equivalence of knots, but they require concepts as homotopy, isotopy, etc.

On the other hand, pursuing the main aim again in the next section the knot invariants are described (mainly emphasizing the Reidemeister moves) for deciding, in some cases, if two knots are not equivalent. That is, the invariants of knots are unidirectional:

If two knots are equivalent \Rightarrow Its invariants are equal.

Finally, ending the appendix and thus this work, we apply all this theory to the study of enzyme-DNA's reactions, as we have said, in the *Knots in biology* section. This adaptation crosses over into the *tangles model*, in order to model mathematically what happens in the synaptosome during the site-specific recombination.

Resumen

Tal y como se puede intuir por el título, este trabajo trata de estudiar la topología y geometría de la molécula de ADN. Hoy en día son muchas las personas que conocen lo que es el ADN, saben que es *algo* que sirve para identificarnos y diferenciarnos de otras personas. Los científicos (biólogos, químicos, médicos y enfermeros, etc), por su parte, conocen algo más sobre esta molécula, saben que no es una simple caja en la que se almacena el material genético, y que dicho material genético se encuentra codificado. Además saben que la molécula de ADN lleva a cabo importantísimas funciones, dentro de la célula que la alberga. Sin embargo, muchos de ellos desconocen otras cosas, como por ejemplo, por qué se obtiene un resultado, u otro, después de que se produzca una de las funciones anteriores. Son, entonces, los investigadores los que estudian los cambios que sufre el ADN después de una determinada reacción para intentar predecir el comportamiento de la molécula de ADN en circunstancias similares.

Estos cambios que estudian los investigadores, aunque ellos lo desconozcan, varían la topología y la geometría de la molécula de ADN. Por tanto, es en este punto en el que entra en juego la matemática, y el tema de este trabajo.

El texto se estructura en tres capítulos y un apéndice, los cuáles son precedidos de un resumen en español, y un *abstract* (resumen) en inglés. Este resumen nos proporcionará una visión general del trabajo, que nos ayudará a comprender mejor las ideas generales del trabajo permitiéndonos quedarnos con un conocimiento global del mismo.

El primer capítulo se llama *Biología del ADN*. Este capítulo, a su vez, se divide en cuatro secciones. En la primera sección del capítulo aparece una breve definición de lo que es el ADN y de cuál es su función básica. Después de esta definición de poco más de cuatro líneas, encontramos una introducción histórica sobre el descubrimiento del ADN y cómo se fue desentrañando su estructura.

En esta sección aparecen nombres como Friedrich Miescher, quien aisló por primera vez la molécula de ADN; Levene, que básicamente descifró que el ADN se encontraba con forma de solenoide (o muelle); Chargraff, que aportó importantes resultados tras sus experimentos; los mundialmente conocidos, Watson y Crick, quienes esclarecieron que el ADN tiene una estructura de doble hélice, ganando un Nobel por ello; y el personaje que permanece a la sombra en la mayoría de las ocasiones, pero que sin sus aportaciones la tarea de Watson y Crick quizás hubiese sido ardua, Rosalind Franklin. Sin embargo, lo más destacable de esta sección son las fechas, que dejan ver que el descubrimiento de la

molécula de ADN es muy temprano.

En la siguiente sección se habla sobre la estructura del ADN, y se distinguen tres niveles de complejidad. La estructura primaria del ADN, que hace referencia a la disposición lineal de los nucleótidos, que son las unidades básicas que conforman la molécula; la estructura secundaria, la cual se refiere al modelo de doble hélice presentado por Watson y Crick; y la estructura terciaria, que se refiere a la disposición espacial que puede tener esta doble hélice, apareciendo el término *superenrollamiento*, que jugará un papel fundamental en este trabajo.

Le sigue la sección que habla sobre las funciones elementales que realiza el ADN dentro de la molécula en la que se encuentra. Apareciendo así la replicación, transcripción, y recombinación. Será en estos procesos en los que el ADN sufra cambios, cambios que admiten una descripción matemática precisa.

Por último, para llevar a cabo cualquier tipo de investigación, es necesario trabajar en un laboratorio. Para estudiar los cambios que ya se han mencionado, se trabaja en el laboratorio con moléculas de ADN circular, pues el ADN circular está restringido desde el punto de vista topológico, lo cuál facilita el estudio y la observación.

El segundo capítulo se titula *Geometría y topología del ADN, y sus invariantes*, y el objetivo principal de este capítulo es demostrar la fórmula de White. La demostración de dicha fórmula no es sencilla, y requiere manejar varios conceptos y resultados previos. Es por eso por lo que en la primera sección de este capítulo se introducen las formas diferenciales, el pullback de una función, y el concepto de derivada exterior, así como algunas proposiciones que establecen propiedades y relaciones entre dichos conceptos.

En la sección siguiente se define el número de enlace, denotado por Lk ; el número de enrollamiento o giros, denotado por Tw ; y el número de retorcimiento, denotado por Wr , matemáticamente, es decir, en términos de la geometría diferencial y mediante una integral. Para ello se parametriza por la longitud de arco una de las hebras que componen la doble hélice, denotada por C , mediante la curva cerrada α . Considerando entonces V como el campo de vectores perpendicular a la curva α , la otra hebra de la doble hélice se denota por $C_V := C + V$, y se parametriza por la longitud de arco mediante la curva cerrada β . En este contexto, se considera entonces el Triedro de Frenet, y se definen los conceptos anteriores como sigue:

$$Lk(C, C_V) := \frac{1}{4\pi} \iint_{C \times C_V} e^*(d\Sigma)$$

$$Tw(C, V) := \frac{1}{2\pi} \int_{[0, \ell]} \langle V^\perp, V' \rangle ds$$

$$Wr(C) := \frac{1}{4\pi} \iint_{C \times C} e^*(d\Sigma)$$

donde e es la aplicación en la esfera dada por: $e(x, y) := \frac{y - x}{|y - x|}$, y donde $*$ denota el pullback.

Dados estos conceptos, es entonces en la sección tercera donde aparece la demostración de la fórmula de White, que establece una relación entre el número de enlace, una cantidad topológica; y el número de enrollamiento y de retorcimiento, que son cantidades geométricas. Dicha relación es la siguiente:

$$Lk(C, C_V) = Tw(C, V) + Wr(C)$$

Para demostrar dicha fórmula se ha recurrido al artículo [11], que usa otras definiciones de estos parámetros. Es por ello por lo que en esta misma sección se estudia la equivalencia de las definiciones anteriores, con las definiciones aportadas por el autor de dicho artículo.

Por último, y para terminar el capítulo, la siguiente sección recoge algunas consecuencias que se pueden obtener de la fórmula de White, estudiando cómo varían los parámetros Tw y Wr dependiendo de como lo hace Lk . De esta forma, la principal consecuencia que podemos extraer del teorema es el estudio del fenómeno del superenrollamiento.

En el último capítulo, cuyo título es *Cambios en la geometría y topología por la acción de enzimas, y diseño de fármacos*, se describen algunas enzimas, cuya acción sobre el ADN es importante, y se habla sobre la otra parte del título del trabajo, es decir, sobre el diseño de fármacos.

El gran protagonista de este capítulo es el superenrollamiento, pues el efecto de todas las enzimas descritas influye directamente sobre el estado de superenrollamiento que pueda tener la molécula de ADN. Debido esto, el capítulo empieza con una introducción sobre este fenómeno.

La primera enzima que se explica es la topoisomerasa. Estas enzimas modifican la topología del ADN cortando y pegando hebras, bien para enredarlo aún más con el fin de conseguir un mayor empaquetado, bien para desenrollarlo para que se lleven a cabo otras funciones del ADN. Este grupo de enzimas se divide en dos tipos, las topoisomerasas de tipo I, y las de tipo II, dependiendo de la cantidad de hebras que cortan en cada ronda de acción.

El siguiente apartado del capítulo no habla de ninguna enzima. En este caso aparece el concepto de recombinación específica de sitio. Este tipo de recombinación se lleva a cabo en un sitio o área concreta de la molécula de ADN, por tanto es el método empleado en el laboratorio para el estudio de los productos obtenidos tras la reacción de una enzima sobre el ADN. Para describir este tipo de recombinación se precisan conceptos como sinaptosoma, sinapsis, sustrato y producto, así como recombinasa, que es el tipo de enzima que actúa en este tipo de recombinación. Debido a la aparición de las recombinasas, la siguiente sección se dedica en su completitud a ellas.

El capítulo se cierra con un apartado en el que se tratan algunas enfermedades conocidas, como el VIH, el cáncer, o tuberculosis, que pueden tratarse con una enzima específica que actúe sobre el ADN del microorganismo (virus o bacteria) responsable de la enfermedad, paralizando su reproducción.

Aquí es donde se aplican todos los conceptos estudiados previamente, pues en el labo-

ratorio se trabaja con una molécula de ADN circular, a la que se le adhiere una enzima específica, en una zona concreta de ésta. Se crea un complejo sináptico, del cual no se puede determinar nada por métodos observacionales directos, y se estudia el producto obtenido, observando los cambios topológicos y geométricos. Así, se puede predecir qué ha ocurrido durante la recombinación específica de sitio, prediciendo el modo de actuación de la enzima. De este modo, a la hora de diseñar un posible fármaco, ya se podría saber qué enzima añadir con el fin de obtener el resultado deseado.

El estudio en el laboratorio de la acción de una enzima sobre una molécula de ADN se basa en el estudio del producto obtenido de la reacción que se produce, y del sustrato (o parte del ADN que se expone a la reacción) inicial. Pero para estudiar esto es necesario apoyarse en lo que se conoce como la *Teoría de nudos*, una rama de la topología relativamente joven.

Es esta teoría la que se desarrolla, de manera breve, en el único apéndice del trabajo. La teoría de nudos es una teoría extensa y difícil, que aún está en desarrollo, por tanto, para saber más sobre ella, se puede consultar [6], [7], [8], y [9].

El primer apartado del apéndice proporciona una breve introducción histórica sobre el *nacimiento* de esta teoría, destacando el hecho de que el posible *padre* de ella sea Gauss, al trabajar con campos magnéticos inducidos. Sin embargo, aparecen otros nombres importantes como Alexander, Reidemeister y Conway. Cabe decir también que el principal objetivo de esta teoría es determinar cuándo dos nudos o enlaces son equivalentes.

En el apartado siguiente se busca definir matemáticamente el concepto de nudo y enlace. Para intentar alcanzar el fin de esta teoría se debe definir la equivalencia de nudos, pero ello requiere de conceptos como homotopía, isotopía, etc.

Por otro lado, persiguiendo de nuevo el fin principal, en la siguiente sección se describen los invariantes de los nudos (destacando principalmente los movimientos de Reidemeister) que permiten decidir, en algunos casos, si dos nudos no son equivalentes. Es decir, los invariantes de los nudos son unidireccionales:

Si dos nudos son equivalentes \Rightarrow Sus invariantes son iguales.

Por último, finalizando el apéndice y con ello este trabajo, aplicamos toda esta teoría de nudos al estudio de las reacciones enzima-ADN, como ya hemos comentado, en el apartado *Nudos en biología*. Esta adaptación pasa por lo que se conoce el *modelo de ovillos*, que trata de modelizar matemáticamente lo que ocurre en el sinaptosoma durante la recombinación específica de sitio.

Capítulo 1

Biología del ADN

¿Qué es el ADN o ácido desoxirribonucleico? El ácido nucleico es la molécula que contiene la información de la vida. El ADN contiene instrucciones genéticas usadas en el desarrollo y funcionamiento de todos los organismos vivos conocidos y algunos virus, y es responsable de su transmisión hereditaria. Por tanto, su función principal es el almacenamiento a largo plazo de esta información genética.

El ADN fue aislado por primera vez, durante el invierno de 1869, por el médico suizo Friedrich Miescher. Esto nos lleva a darnos cuenta de que esta molécula se descubrió hace relativamente poco tiempo.

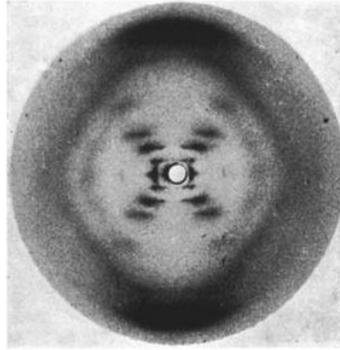
En 1919 fue Phoebus Levene quien observó que cada nucleótido (unidad básica que conforma el ADN) está formado por una base nitrogenada, un azúcar y un fosfato (en la siguiente sección hablaremos más de ello). Levene sugirió que el ADN generaba una estructura con forma de solenoide (muelle) con unidades de nucleótidos unidos a través de los grupos fosfato, pero pensaba que la cadena era corta y que las bases se repetían en un orden fijo.

En 1940 Chargaff realizó algunos experimentos que le sirvieron para establecer las proporciones de las bases nitrogenadas en el ADN. Estas aportaciones de Chargaff y los datos de difracción de rayos X fueron los que permitieron a Rosalind Franklin [14] aceptar la idea de que el ADN es una doble hélice.

No obstante, fueron James Watson y Francis Crick quienes, tomando como base los trabajos realizados en el laboratorio por el propio Crick y el biofísico Wilkins, y de Franklin, desentrañaron la estructura de doble hélice de la molécula de ADN. Estas investigaciones proporcionaron los medios para comprender cómo se copia y transmite, de una generación a otra, la información hereditaria del ser humano. Fue por este descubrimiento por el que ambos recibieron en 1962 el Premio Nobel de Fisiología y Medicina.



(a) Rosalind Franklin



(b) Foto 51



(c) Watson (izq) y Crick(dcha), y su modelo de doble hélice

1.1. Estructura primaria, secundaria y terciaria del ADN

Una vez introducido el ADN, vamos a estudiar la estructura de esta molécula, usando, como no podía ser de otra forma, los estudios realizados por los autores anteriormente mencionados.

Estructura primaria y secundaria. La estructura primaria hace referencia a la disposición lineal de los nucleótidos, que son las unidades fundamentales de la molécula.

Estos nucleótidos están compuestos por una base nitrogenada, una pentosa (un azúcar), y un ácido fosfórico. Dependiendo entonces de la base nitrogenada que los compongan, adenina (A), guanina (G), citosina (C) o timina (T), podemos distinguir un nucleótido de otro.

Esta disposición lineal hace que aparezcan las llamadas cadenas de ADN¹, y es en ellas donde se encuentra la información genética. Las diferencias de esta información radica en la distinta secuencia de bases nitrogenadas que la formen, pues el esqueleto es el mismo para todos.

Por convenio, la lectura de las cadenas polinucleótidas del ADN se realiza desde el extremo 5' (es el último fosfato libre, es decir, es aquel extremo al que está unido el último nucleótido) al extremo 3' (que es la última pentosa libre, es decir, aquel extremo que ya no va unido a ningún otro nucleótido) nombrando la secuencia de las bases, y fijando una dirección de lectura.

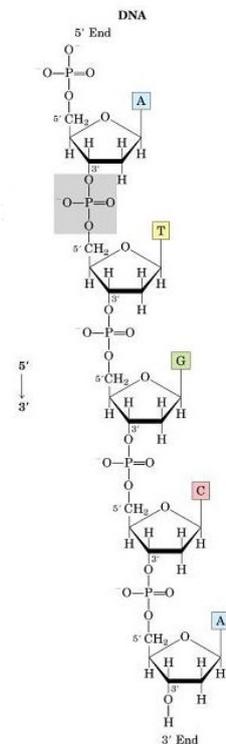


Figura 1.1: Cadena ADN

¹En ocasiones, nos referiremos a esta cadena de ADN con el nombre de *hebra*.

El siguiente nivel de complejidad es la estructura secundaria, que hace referencia al modelo de doble hélice presentado por Watson y Crick ². En este modelo se concluyó que el ADN estaba formado por dos cadenas *antiparalelas*, y complementarias, unidas por las bases nitrogenadas (A,G,C,T) a través de puentes de hidrógeno (que es un tipo específico de enlace [29]).

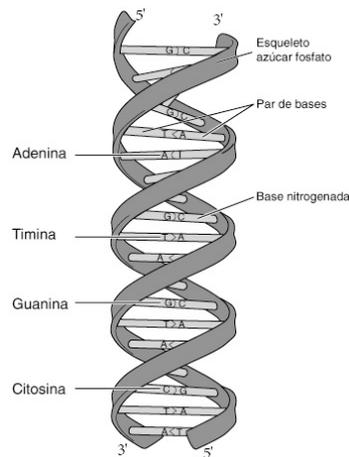


Figura 1.2: Estructura de doble hélice del ADN

Pero ahora bien, ¿qué significa que las cadenas sean *antiparalelas* y complementarias?. Que las cadenas sean *antiparalelas* significa que la dirección de los nucleótidos en una de las cadenas ($5' \rightarrow 3'$) es opuesta a la dirección de la otra cadena ($3' \rightarrow 5'$). Es decir, las cadenas tienen direcciones opuestas. Por tanto, fijada la orientación de una de las cadenas, tendremos determinada la orientación de la otra.

Por otro lado, cada tipo de base, en una cadena, forma un enlace únicamente con un tipo de base de la otra cadena, lo que se denomina complementariedad de las bases. De esta forma, la timina (T) sólo se enlaza con la adenina (A), mediante dos puentes de hidrógeno; y la citosina (C) sólo lo hace con la guanina (G), mediante tres puentes de hidrógeno. Gracias a esta complementariedad, conocida la secuencia de una cadena, conocemos la secuencia de la otra.

Por último, cabe destacar que la longitud de la molécula de ADN se mide en función del número de pares de bases que la formen. Por ejemplo, el ADN del genoma humano tiene tres billones de pares de bases, midiendo de ancho, cada hélice doble 25 Angstroms ($3 \times 10^{-10}m$).

²Conviene decir que existen también moléculas de ADN formadas por una sola hebra, configuración que se conoce como *ADN monocatenario*, en contraposición al *ADN bicatenario*, que es el que estudiamos aquí.

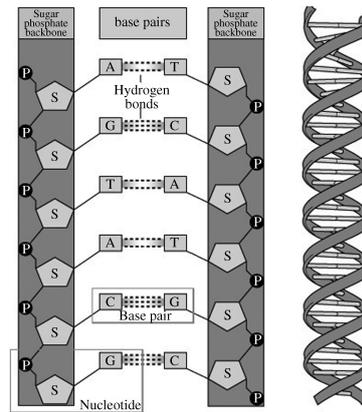
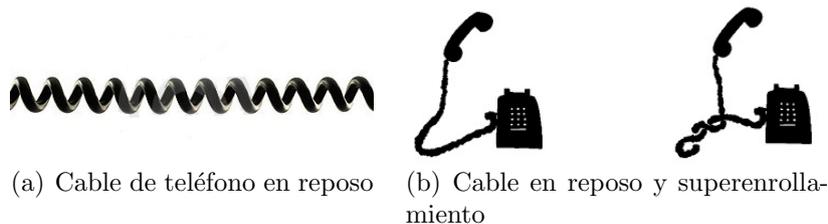


Figura 1.3: Cadenas complementarias del ADN

Estructura terciaria. Una década después del descubrimiento de la doble hélice, Vinograd encontró la característica fundamental de la estructura terciaria: el eje del ADN también puede estar enrollado en el espacio, lo que conduce a un ADN *superenrollado*.

El eje de la molécula de ADN del genoma humano es lineal, pero, no obstante, la molécula de ADN adopta generalmente configuraciones espaciales más complejas encaminadas a lograr un empaquetamiento que sea lo más efectivo posible, pues la longitud de la molécula es, en ocasiones, del orden de un millón de veces el tamaño del núcleo que las alberga.

Estamos habituados a los *superenrollamientos*. Por ejemplo, fijémonos en un teléfono con base y cable. En su situación de reposo, el cable del teléfono está enrollado en espiral. Si descolgamos el teléfono y empezamos a añadir vueltas o giros a la cuerda, ésta se mantiene en su situación de reposo sólo mientras la mantenemos estirada o tensa. Tan pronto como la dejemos coger holgura, el cable se retorcerá sobre sí mismo, obteniendo el *superenrollamiento*.



Por otro lado, en algunas células, el eje de la molécula de ADN adquiere forma circular. A principios de 1960 se descubrió que la molécula de ADN, en una cierta bacteria, aparecía como un solo anillo trenzado. Desde entonces, se ha descubierto que el ADN cíclico no aparece solo en bacterias y virus, sino que también aparece en las mitocondrias de las células humanas.

Nosotros, por nuestra parte, centraremos toda nuestra atención al ADN circular bicatenario (el eje de la molécula de ADN es circular, y está formado por una doble hélice), de forma que, cuando el ADN presente esta estructura (sin superenrollamientos), diremos que la molécula de ADN está relajada. El motivo de esto se explicará a continuación.

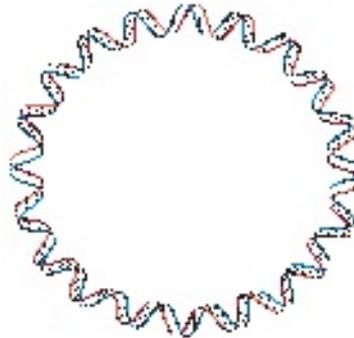


Figura 1.4: Molécula de ADN circular bicatenario, en su forma relajada.

1.2. Replicación, transcripción, y recombinación

Aunque la información genética se encuentra codificada en la secuencia de los nucleótidos (en la estructura primaria), investigaciones recientes apuntan que la *forma de anudarse* la molécula de ADN provoca importantes y decisivos efectos en las funciones que desempeña el ADN en la célula. Estas funciones son la replicación, la transcripción y la recombinación. Veamos los aspectos más determinantes de cada una de ellas.

La **replicación** es el proceso a través del cual se copia el ADN paterno dando lugar a dos moléculas hijas de ADN con idéntica información entre ellas y el ADN parental, permitiendo el paso de información genética entre generaciones.

La **transcripción** es el proceso mediante el cual parte del ADN se copia en forma de ARN, concretamente en ARN mensajero, para formar, posteriormente, proteínas.

Con respecto a estos procesos, ciertas enzimas (por ejemplo, las topoisomerasas, de las que también hablaremos más adelante) actúan sobre una estructura molecular dada para promover o impedir los citados procesos.

Por ejemplo, para que puedan tener lugar los procesos de replicación y transcripción es necesario que se separen las dos hebras complementarias, lo cual supone un desenrollamiento local de la doble hélice. Inversamente, algunos antibióticos están diseñados para interactuar con ciertas enzimas y aumentar la *cantidad de enrollamiento* del ADN al que agrede, dificultando así los procesos vitales anteriores.

De esto, podemos deducir que ciertas enzimas provocan cambios en la geometría de la molécula de ADN y, mediante estos cambios, se puede concluir qué efecto tiene esta enzima sobre el ADN, o incluso, qué enzima ha actuado sobre la molécula de ADN.

En cuanto a la **recombinación**, éste es el proceso por el cual una hebra de un material genético dado es cortada para luego ser unida a una molécula de otro material genético diferente. En biología molecular, “recombinación” también se refiere a la recombinación artificial y deliberada de piezas de ADN distintas, a menudo de diferentes organismos.

Es en este proceso donde aparecerán las enzimas recombinasas, de las que hablaremos posteriormente. De nuevo, cabe destacar, que este proceso involucra interesantes cambios en la geometría de la molécula.

Todos estos cambios geométricos que han sido comentados admiten una descripción matemática precisa, y de esto es de lo que nos vamos a encargar en el resto del documento.

1.3. Modelización en laboratorio

Para detectar los cambios promovidos por una enzima en el ADN, se crea en el laboratorio (*in vitro*) un sustrato de ADN circular construido vía clonación, de tal manera que las moléculas del sustrato tienen la estructura deseada para que la enzima actúe y produzca la reacción buscada.

De este modo, una vez que actúa la enzima sobre la molécula de ADN, pueden ser observados los cambios topológicos y/o geométricos mediante *electroforesis en gel de agarosa*³ y, posteriormente, mediante el microscopio electrónico, en este caso, para determinar con mayor exactitud el resultado obtenido.

Resaltar también un aspecto importante de la modelización en el laboratorio, y es que se asume que el mecanismo de la acción de las enzimas es constante (la enzima siempre hace lo mismo), y que éstas actúan independientemente de la geometría y la topología del sustrato.

En este capítulo (en la sección 2.1) ya se ha comentado que en el laboratorio se clona ADN circular con el fin de estudiar el resultado final obtenido tras la acción de una enzima sobre una molécula de ADN. El motivo de esto es porque los extremos de las hebras de ADN están fijos (o *sujetos*) en esta situación. De no ser así, puede ocurrir que el efecto producido por la enzima tras la reacción no pueda ser apreciado.

Un ejemplo de esto es cuando una enzima provoca que las hebras de ADN se enrollen entre sí aún más. En este contexto, si los extremos de la molécula de ADN estuvieran libres (sin sujeción a nada) el cruce producido por la acción de la enzima sería *invisible*. Sin embargo, si los extremos permaneciesen fijos (ya sean sujetos a algo, o unidos entre ellos), apreciaríamos que la molécula de ADN tiene un *giro* más que cuando aún no había actuado la enzima.

³La electroforesis en gel es un método que se emplea para separar macromoléculas en función del tamaño, la carga eléctrica y otras propiedades físicas. La electroforesis en gel es una técnica consistente en aplicar corriente eléctrica a las moléculas para que atraviesen una placa de gel.

Capítulo 2

Geometría y topología del ADN y sus invariantes

2.1. Preliminares

El objetivo de este capítulo es demostrar la fórmula de White, que relaciona el número de enlace, con el número de enrollamiento y el número de retorcimiento (conceptos que definiremos en breve). Por tanto, para ello, debemos definir algunos conceptos previos, y demostrar algunas de sus propiedades, relacionados con formas diferenciales en \mathbb{R}^3 , el espacio euclídeo.¹

Ya que la demostración de todas las propiedades que presentaremos a continuación excede el fin de este trabajo, y alargaría considerablemente el número de páginas del mismo, omitiremos la mayoría de ellas. No obstante, dichas demostraciones pueden consultarse en el libro de Manfredo Do Carmo, titulado *Differential forms and applications* [24].

Sea $p \in \mathbb{R}^3$, y denotemos al espacio tangente de p en \mathbb{R}^3 , generado por los vectores $q \in \mathbb{R}^3$ y $q-p$, por \mathbb{R}_p^3 , cuya base está generada por el conjunto $\{(e_i)_p\}$ (donde cada $(e_i)_p$ es la traslación al punto p del vector canónico (e_i)).

A cada espacio tangente, podemos asociarle su espacio dual, denotado por $(\mathbb{R}_p^3)^*$, que no es más que el conjunto de las aplicaciones lineales $\varphi : \mathbb{R}_p^3 \rightarrow \mathbb{R}$.

Una base para este espacio es el conjunto $\{(dx_i)_p$ tal que $i = 1, 2, 3\}$, tal que

$$(dx_i)_p(e_j) = \frac{\partial x_i}{\partial x_j} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

donde $x_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ asigna a cada punto p su i -ésima coordenada.

¹Todos los conceptos se pueden extender a \mathbb{R}^n , pero, puesto que nuestra molécula vive en un espacio de tres dimensiones, nos centraremos en el caso particular en que $n = 3$.

Definición 2.1.1. Un campo de formas lineales, o una **forma exterior diferencial de grado 1** en \mathbb{R}^3 es una aplicación $\omega : \mathbb{R}^3 \rightarrow (\mathbb{R}_p^3)^*$ definida mediante

$$\omega(p) := a_1(p)(dx_1)_p + a_2(p)(dx_2)_p + a_3(p)(dx_3)_p = \sum_{i=1}^3 a_i dx_i$$

donde las a_i son funciones diferenciables reales en \mathbb{R}^3 .

Denotemos ahora $\Lambda^2(\mathbb{R}_p^3)^*$ como el conjunto de aplicaciones $\varphi : \mathbb{R}_p^3 \times \mathbb{R}_p^3 \rightarrow \mathbb{R}$ que son bilineales y antisimétricas ($\varphi(v_1, v_2) = -\varphi(v_2, v_1)$). En este contexto, definimos la siguiente *operación*:

$$(\varphi_1 \wedge \varphi_2)(v_1, v_2) := \det(\varphi_i(v_j))$$

De este modo, si denotamos $(dx_i)_p \wedge (dx_j)_p \in \Lambda^2(\mathbb{R}_p^3)^*$ por $(dx_i \wedge dx_j)_p$, se verifica que:

1. $(dx_i \wedge dx_i)_p = 0$
2. $(dx_i \wedge dx_j)_p = -(dx_j \wedge dx_i)_p \quad i \neq j$
3. $\{(dx_i \wedge dx_j)_p \text{ tal que } i < j\}$ es una base de $\Lambda^2(\mathbb{R}_p^3)^*$

siendo las propiedades 1. y 2. triviales, y pudiéndose ver la prueba de 3. en [24]

Definición 2.1.2. Un campo de formas bilineales y antisimétricas, o una **forma exterior diferencial de grado 2** en \mathbb{R}^3 es una aplicación $\omega : \mathbb{R}^3 \rightarrow \Lambda^2(\mathbb{R}_p^3)^*$ definida mediante

$$\begin{aligned} \omega(p) &:= a_{12}(p)(dx_1 \wedge dx_2)_p + a_{13}(p)(dx_1 \wedge dx_3)_p + a_{23}(p)(dx_2 \wedge dx_3)_p = \\ &= \sum_{i < j} a_{ij}(p)(dx_i \wedge dx_j)_p \quad i, j = 1, 2, 3 \end{aligned}$$

donde las a_{ij} son funciones diferenciables reales en \mathbb{R}^3 .

Si generalizamos este concepto, nos queda la siguiente definición:

Definición 2.1.3. Una **k -forma exterior diferencial** en \mathbb{R}^n es una aplicación $\omega : \mathbb{R}^n \rightarrow \Lambda^k(\mathbb{R}_p^n)^*$ definida mediante ²

$$\omega(p) := \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k}(p)(dx_{i_1} \wedge \dots \wedge dx_{i_k})_p \quad i_j \in \{1, \dots, n\}$$

donde las $a_{i_1 \dots i_k}$ son funciones diferenciables reales en \mathbb{R}^n .

²La definición es una buena definición porque se puede demostrar que el conjunto dado por $\{(dx_{i_1} \wedge \dots \wedge dx_{i_k})_p \text{ tal que } i_1 < \dots < i_k, \quad i_j \in \{1, \dots, n\}\}$ es una base de $\Lambda^k(\mathbb{R}_p^n)^*$.

Notación: Consideraremos \mathbf{I} como la k -upla $(i_1 \dots i_k)$, con $i_1 < \dots < i_k$; $i_j \in \{1, \dots, n\}$. Así, con esta notación, la k -forma ω queda definida como sigue:

$$\omega(p) = \sum_I a_I dx_I.$$

También aceptaremos por convenio que una 0-forma diferenciable es una función diferencial $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Veamos ahora un nuevo concepto:

Definición 2.1.4. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación diferenciable. Esta f induce una aplicación $f^* : \mathbb{R}^m \rightarrow \mathbb{R}^n$, llamada **pull-back** de \mathbf{f} , y que está definida por:

$$(f^*\omega)(p)(v_1, \dots, v_k) := \omega(f(p))(df_p(v_1), \dots, df_p(v_k))$$

donde ω es una k -forma en \mathbb{R}^m , $p \in \mathbb{R}^n$, $v_1, \dots, v_k \in \mathbb{R}^n$, y donde $df_p : \mathbb{R}^n \rightarrow \mathbb{R}_{f(p)}^m$ es la diferencial de f .

Notación: Si ω es una 0-forma, admitimos por convenio:

$$f^*(g) = g \circ f$$

Veamos algunas propiedades, cuyas pruebas pueden verse en [24]:

Proposición 2.1.5. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación diferenciable, sean ω y φ k -formas en \mathbb{R}^m , y sea $g : \mathbb{R}^m \rightarrow \mathbb{R}$ una 0-forma en \mathbb{R}^m . Entonces:

1. $f^*(\omega + \varphi) = f^*\omega + f^*\varphi$
2. $f^*(g\omega) = f^*(g)f^*(\omega)$
3. Si $\varphi_1, \dots, \varphi_k$ son 1-formas en \mathbb{R}^m , $f^*(\varphi_1 \wedge \dots \wedge \varphi_k) = f^*(\varphi_1) \wedge \dots \wedge f^*(\varphi_k)$

Algo más general...

Proposición 2.1.6. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación diferenciable. Entonces:

1. $f^*(\omega \wedge \varphi) = f^*(\omega) \wedge f^*(\varphi)$ donde ω y φ son k -formas en \mathbb{R}^m .
2. $(f \circ g)^*\omega = g^*(f^*(\omega))$, donde $g : \mathbb{R}^p \rightarrow \mathbb{R}^n$ es una carta diferenciable.

A continuación definimos una operación para las formas diferenciales que generaliza el concepto de la diferencial de una función, con algunas propiedades:

Definición 2.1.7. Sea $\omega(p) = \sum_I a_I dx_I$ una k -forma en \mathbb{R}^n . Se define la **diferencial exterior**, denotada por $d\omega$, como:

$$d\omega := \sum_I da_I \wedge dx_I$$

que lleva k -formas a $(k + 1)$ -formas.

Con esta definición, si consideramos $g : \mathbb{R}^n \rightarrow \mathbb{R}$ una 0-forma (es decir, una función diferenciable), entonces su diferencial exterior es simplemente la diferencial de g :

$$dg = \sum_{i=1}^n \frac{\partial g}{\partial x_i} dx_i$$

Proposición 2.1.8. *Se verifican las siguientes propiedades:*

1. $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$, donde ω_1 y ω_2 son k -formas.
2. $d(\omega \wedge \varphi) = d\omega \wedge \varphi + (-1)^k \omega \wedge d\varphi$, donde ω es una k -forma, y φ es una s -forma
3. $d(d\omega) = d^2\omega = 0$

El siguiente resultado se usará en la demostración de la fórmula de White, por tanto desarrollaremos su demostración con detalle.

Proposición 2.1.9. *Si ω es una k -forma en \mathbb{R}^m y $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una aplicación diferenciable, entonces*

$$d(f^*\omega) = f^*(d\omega)$$

Demostración. Supongamos, en primer lugar, que $k = 0$. Sea $g : \mathbb{R}^m \rightarrow \mathbb{R}$ una función diferenciable con coordenadas y_1, \dots, y_m . De este modo, se tiene:

$$f^*(dg) = f^*\left(\sum_{i=1}^m \frac{\partial g}{\partial y_i} dy_i\right) = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial g}{\partial y_i} \frac{\partial f_i}{\partial x_j} dx_j = \sum_{j=1}^n \frac{\partial(g \circ f)}{\partial x_j} dx_j = d(g \circ f) = d(f^*g)$$

Veamos ahora el caso en que $k \geq 1$. En este caso, consideramos $\omega = \sum_I a_I dx_I$ una k -forma. Como los a_I son 0-formas o funciones diferenciables, por la proposición 2.1.5.2 tenemos:

$$f^*\omega = f^*\left(\sum_I a_I dx_I\right) = \sum_I f^*(a_I dx_I) = \sum_I f^*(a_I) f^*(dx_I)$$

Con esto, usando ahora la proposición 2.1.8.1, y la 2.1.5.3 (pues los dx_I son 1-formas, nos queda:

$$d(f^*\omega) = d\left(\sum_I f^*(a_I) f^*(dx_I)\right) = \sum_I d(f^*(a_I) f^*(dx_I)) = \sum_I d(f^*(a_I)) (f^*(dx_{i_1}) \wedge \dots \wedge f^*(dx_{i_k}))$$

Usando ahora la definición 2.1.7 y el caso $k = 0$, llegamos a lo siguiente:

$$\sum_I d(f^*(a_I)) (f^*(dx_{i_1}) \wedge \dots \wedge f^*(dx_{i_k})) = \sum_I (f^*(da_I) \wedge (f^*(dx_{i_1}) \wedge \dots \wedge f^*(dx_{i_k})))$$

Con los apartados 3 y 1, respectivamente, de la proposición 2.1.5, concluimos finalmente que:

$$\begin{aligned} d(f^*\omega) &= \sum_I (f^*(da_I) \wedge (f^*(dx_{i_1}) \wedge \dots \wedge f^*(dx_{i_k}))) = \sum_I f^*(da_I \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}) = \\ &= f^*\left(\sum_I da_I \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}\right) = f^*(d\omega) \end{aligned}$$

□

2.2. Invariantes del ADN

En la sección anterior hemos comentado que la fórmula de White establece una relación entre ciertos parámetros. Bien, es el momento entonces de definirlos. En primer lugar, para entender su relación, vamos a considerar la molécula de ADN circular como una cinta larga enrollada sobre sí misma. Matemáticamente, esta banda B puede corresponderse con un subconjunto del espacio homeomorfo al producto $\mathbb{S}^1 \times [0, 1]$. Las dos curvas que constituyen el borde la banda B representan las hebras que se van entrelazando alrededor del eje de B . A estas hebras las llamaremos, indistintamente, W y C (por Watson y Crick).

Lo importante de las hebras es que poseen orientaciones distintas (como ya vimos en el capítulo anterior), lo que nos permite concluir que no puede existir una molécula de ADN circular con la *forma* o *aspecto* de una banda de Möbius, pues la frontera de la cinta de Möbius \mathbb{M} , $\partial\mathbb{M}$, posee una sola orientación.

Para dar una interpretación matemática de estos conceptos necesitamos establecer un contexto. Consideremos, en primer lugar, una curva C , cerrada y simple en \mathbb{R}^3 , cuya parametrización es $\alpha : [0, \ell] \rightarrow C \subseteq \mathbb{R}^3$. No es restrictivo suponer que ℓ es el parámetro arco, pues de lo contrario podríamos reparametrizar por la longitud de arco y obtener curvas con la misma traza.

Consideramos ahora $V : [0, \ell] \rightarrow \mathbb{R}^3$ un campo de vectores a lo largo de la curva α , unitario y perpendicular a ésta. Esto es:

$$\begin{cases} V(s) \in \mathbb{R}^3 & \forall s \\ |V(s)| = 1 \\ \langle V(s), \alpha'(s) \rangle = 0 \end{cases}$$

Además, supondremos que $V(0) = V(\ell)$. Con esto se ha formado una banda B dada por (C, V) , la cual supondremos que no tiene autointersecciones, es decir, que es una superficie reglada regular.

Definimos ahora la curva C_V generada por los puntos finales del campo V tal que, abusando de la notación, podemos considerar que $C_V = C + V$. La parametrización de esta curva C_V viene dada por $\beta : [0, \ell] \rightarrow C_V \subseteq \mathbb{R}^3$ y definida por $\beta(s) := \alpha(s) + V(s)$. Gracias a la condición de que $V(0) = V(\ell)$, podemos asegurar que C_V también es una curva cerrada y simple.

Con esta notación, el par formado por la curva C y el campo de vectores V representan, respectivamente, una de las dos cadenas de la molécula de ADN (digamos la cadena C), y la dirección de los puentes de hidrógeno. Con esto, la curva C_V denotará la otra cadena de la doble hélice de ADN (en este caso W). Así el *nucleótido* $\alpha(s)$ está unido por el *puente de hidrógeno* $V(s)$ al *nucleótido* $\beta(s)$.

Si tomamos ahora el vector $t(s)$ como el vector unitario y tangente a C en el punto $\alpha(s)$ (y positivamente orientado) tal que $t(s) = \alpha'(s)$, podemos considerar la siguiente tríada

$$\{t(s) = \alpha'(s), V(s), V^\perp(s) = \alpha'(s) \wedge V(s)\}$$

que es una base ortonormal de \mathbb{R}^3 (positivamente orientada).

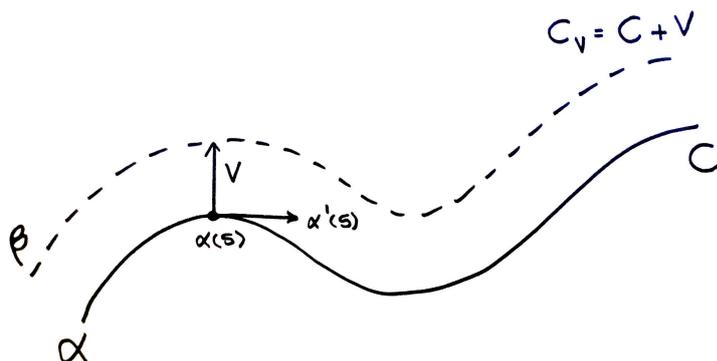


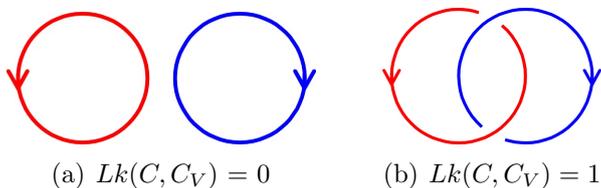
Figura 2.1: Esquema de la situación del contexto.

Una vez hecha esta introducción, pasamos a dar las definiciones.

El número de enlace

El primer parámetro del que hablaremos es el **número de enlace**, denotado por $Lk(C, C_V)$. Este parámetro es un invariante topológico que describe la conexión de dos curvas cerradas en el espacio tridimensional.

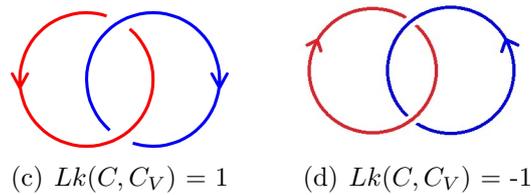
Intuitivamente podríamos definirlo como la cantidad de veces que una hebra está enrollada con la otra, o como la cantidad de veces que deberíamos cortar una hebra (digamos la C) para separarla de la otra hebra (C_V).



(a) $Lk(C, C_V) = 0$

(b) $Lk(C, C_V) = 1$

Que el parámetro sea un invariante topológico es porque se conserva bajo homeomorfismos que preservan la orientación, de hecho, al invertir la orientación de una las hebras, el valor de $Lk(C, C_V)$ cambia de signo:



Una forma de contabilizar el número de enlace de dos curvas es siguiendo un algoritmo de etiquetas positivas y negativas que se rige por las siguientes reglas:

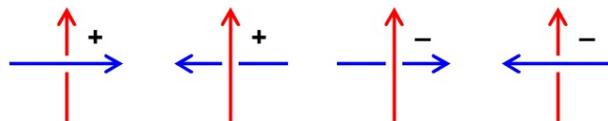


Figura 2.2: Cruces y sus signos

De esta manera, se puede contabilizar el número de enlace como

$$Lk(C, C_V) = \frac{1}{2} \sum_{\text{cruces}} \pm 1$$

Así, a modo de ejemplo y aplicación, el número de enlace de la siguiente figura es 2:

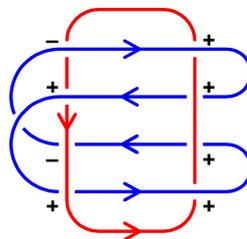


Figura 2.3: $Lk(C, C_V) = 2$

La definición que hemos dado del número de enlace podríamos considerarla *numérica*, pero podemos dar una definición *geométrica* en términos de una integral, que será la definición que utilizaremos para demostrar la fórmula de White.

Heredando la notación de la introducción del inicio de la sección, podemos considerar la variedad 2-dimensional $C \times C_V \in \mathbb{R}^6$, donde ambas curvas están parametrizada por el parámetro arco. Tomamos ahora la siguiente aplicación en la esfera \mathbb{S}^2 :

$$e : C \times C_V \longrightarrow \mathbb{S}^2$$

que viene dada por

$$e(x, y) := \frac{y - x}{|y - x|}$$

Con esto, se define el número de enlace mediante la integral siguiente:

$$Lk(C, C_V) := \frac{1}{4\pi} \iint_{C \times C_V} e^*(d\Sigma)$$

donde $e^*(d\Sigma)$ es el pullback del elemento de área $d\Sigma$ de la esfera \mathbb{S}^2 bajo la aplicación en la esfera e anterior.

Vamos a ver otra definición del número de enlace, pero posteriormente demostraremos que ambas definiciones son equivalentes. Esta *nueva* definición es la siguiente:

$$Lk(C, C_V) = -\frac{1}{4\pi} \iint_{C \times C_V} g_1^*(\omega)$$

donde $g_1 : C \times C_V \rightarrow \mathbb{R}^3 \setminus \{0\}$ viene dada por

$$g_1(s, t) := \alpha(s) - \alpha(t) - V(t)$$

y donde ω es la 2-forma en $\mathbb{R}^3 \setminus \{0\}$ definida por

$$\frac{1}{(x^2 + y^2 + z^2)^{3/2}} (x dy \wedge dz - y dx \wedge dz + z dx \wedge dy)$$

Interpretemos esta última integral, más conocida como integral de Gauss, apoyándonos en [22] y sirviéndonos del electromagnetismo. Sea C una curva plana, y sea C' otra curva que rodea a C m veces. Supongamos además que ambas curvas son cerradas y simples. Entonces:

$$\iint_{C \times C'} g_1^*(\omega) = 4\pi m$$

siendo m el número de enlace de las curvas C y C' .

La demostración de esta igualdad puede verse en [22], no obstante, veamos las ideas de la misma. Para ello utilizaremos lo siguiente:

Teorema 2.2.1 (Gauss, 1839). *El potencial que crea una membrana magnética Σ , en un punto arbitrario P , viene dado por el ángulo sólido³ subtendido por la membrana desde dicho punto.*

Por otra parte, usaremos también el siguiente resultado clásico de electromagnetismo:

³Para calcular el ángulo sólido bajo el cual se ve un objeto desde un punto, se proyecta el objeto sobre una esfera de radio R conocido, centrada en el punto de vista. Si la superficie de la proyección del objeto sobre la esfera es S , el ángulo sólido bajo el cual se ve el objeto es $\frac{S}{R^2}$, y mide el tamaño aparente del objeto.

Teorema 2.2.2 (Ampère, 1831). *Un círculo cerrado C produce el mismo campo magnético que cualquier membrana magnética Σ con $\partial\Sigma = C$.*

Usando estos dos resultados para C una curva plana cerrada y simple, y Σ la región plana acotada tal que $\partial\Sigma = C$ se demuestra que la diferencia del potencial V entre dos puntos muy próximos a cada lado de la membrana viene dado por:

$$V(P_+) - V(P_-) = 4\pi$$

tal que

$$\frac{\Delta V}{4\pi} = m$$

Por otra parte, calculando el potencial a lo largo de una curva C' se demuestra que éste se expresa como:

$$V = \iint_{C \times C'} = g_1^*(\omega)$$

Cada vez que la curva C' pasa por dentro de C y atraviesa la membrana Σ se suma (o se resta) 4π en virtud de la expresión anterior.

Además, este razonamiento es válido para cualquier curva C no necesariamente plana, y toda membrana Σ . El motivo es que si modificamos de forma continua la curva C la integral debe variar de modo continuo. Ahora bien, como la expresión de V anterior solo toma valores discretos, entonces la integral permanece invariante.

Por tanto, queda justificada la definición del número de enlace, y el hecho de que el número de enlace sea una cantidad entera.

El número de enrollamiento

El segundo parámetro que desarrollaremos es el **número de enrollamiento** o número de giros o vueltas, que se denota mediante $Tw(C, V)$, y mide el número de vueltas total que da V a lo largo de la curva α , desde $\alpha(0)$ a $\alpha(\ell)$ (usando, de nuevo, la notación del contexto en el que estamos).

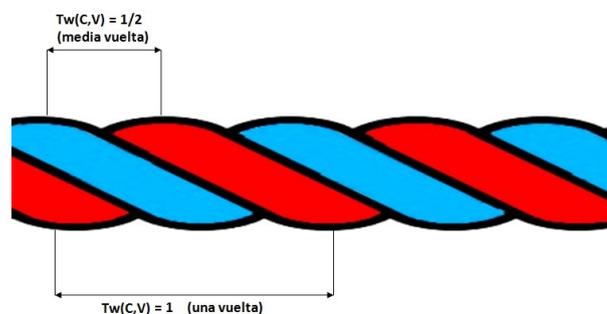


Figura 2.4: Número de enrollamiento

Como la curva C es cerrada y hemos establecido la condición de que $V(0) = V(\ell)$, el número de retorcimiento es una cantidad entera, al igual que el parámetro anterior. Si no fuese una curva cerrada, el número de retorcimiento no tiene por qué ser entero. Por otro lado, esta cantidad no es un invariante topológico, sino que, en este caso, es una cantidad geométrica.

Dada esta definición o interpretación, se define el número de enrollamiento, en términos de una integral, como sigue:

$$Tw(C, V) := \frac{1}{2\pi} \int_{[0, \ell]} \langle V^\perp, V' \rangle ds$$

Justifiquemos esta definición. Para ello vamos a considerar $k(s)$ la curvatura de la curva α . Esta curvatura puede ser mayor o igual que cero. El caso en que $k(s)$ sea igual a cero lo vemos a continuación. Por tanto, si suponemos que $k(s) = |\alpha''(s)|$ es mayor que cero ⁴ podemos construir el Triedo de Frenet (que es un referencial ortonormal), determinado por:

$$\{t(s), n(s), b(s)\}$$

donde $t(s) = \alpha'(s)$, es el vector tangente con el que ya hemos trabajado; $n(s) = \frac{t'(s)}{k(s)}$, es el vector normal a la curva; y $b(s) = t(s) \wedge n(s)$, es el vector binormal de α .

Operando con los vectores anteriores, podemos conseguir las siguientes relaciones:

$$\begin{cases} t'(s) = k(s)n(s) \\ n'(s) = -k(s)t(s) - \tau(s)b(s) \\ b'(s) = \tau(s)n(s) \end{cases}$$

donde $\tau(s) := \langle t(s) \wedge n'(s), n(s) \rangle = \langle b'(s), n(s) \rangle$ es lo que se conoce como la torsión de la curva α .

Con esto, como la tríada anterior es una base, podemos expresar el campo V en función de dicha base, de manera que $V(s) = p(s)t(s) + q(s)n(s) + r(s)b(s)$, siendo $p(s)$, $q(s)$ y $r(s)$ diferenciables. Sin embargo, teniendo en cuenta que $|V(s)| = 1$, y que $\langle V(s), \alpha'(s) \rangle = 0$, la expresión anterior queda reducida a $V(s) = q(s)n(s) + r(s)b(s)$. Pero, usando de nuevo la condición de que V es unitario, existe una función $\theta(s)$ diferenciable tal que

$$q(s) = \cos(\theta(s)) \quad \text{y} \quad r(s) = \sin(\theta(s))$$

De este modo, $V(s) = \cos(\theta(s))n(s) + \sin(\theta(s))b(s)$, y derivando nos queda:

$$V'(s) = -\theta'(s)\sin(\theta(s))n(s) + \cos(\theta(s))n'(s) + \theta'(s)\cos(\theta(s))b(s) + \sin(\theta(s))b'(s)$$

⁴Puede ocurrir que $k(s_0) = 0$ para algún $s_0 \in [0, \ell]$, pero como los puntos son conjuntos de medida nula, el valor de la integral de la definición no se vería alterado. Así podemos suponer, sin pérdida de generalidad, que trabajamos con una curvatura positiva.

Así:

$$\begin{aligned} V^\perp(s) &= t(s) \wedge V(s) = t(s) \wedge (\cos(\theta(s))n(s) + \operatorname{sen}(\theta(s))b(s)) = \\ &= \cos(\theta(s))(t(s) \wedge n(s)) + \operatorname{sen}(\theta(s))(t(s) \wedge b(s)) = \cos(\theta(s))b(s) + \operatorname{sen}(\theta(s))n(s) \end{aligned}$$

donde la última igualdad se tiene debido a que $t(s) \wedge n(s) = b(s)$, por el triedro de Frenet; y a que $\langle (t \wedge b), n \rangle = \det(t, b, n) = -\det(t, n, b) = -1$ y, por tanto, $t(s) \wedge b(s) = -n(s)$.

Con estas operaciones, y teniendo en cuenta que el Triedro de Frenet es un referencial ortonormal, junto con la propiedad $\langle b'(s), n(s) \rangle = -\langle b(s), n'(s) \rangle$, nos queda:

$$\begin{aligned} \langle V'(s), V^\perp(s) \rangle &= \cos^2(\theta(s))\langle n'(s), b(s) \rangle - \operatorname{sen}^2(\theta(s))\langle b'(s), n(s) \rangle + \\ &\quad + \theta'(s)\operatorname{sen}^2(\theta(s))\langle n(s), n(s) \rangle + \theta'(s)\cos^2(\theta(s))\langle b(s), b(s) \rangle = \\ &= \theta'(s) + \langle b(s), n'(s) \rangle = \theta'(s) - \tau(s) \end{aligned}$$

Así, con esta última expresión, la definición del número de enrollamiento puede escribirse de la siguiente forma:

$$Tw(C, V) := \frac{1}{2\pi} \int_0^\ell (\theta'(s) - \tau(s)) ds$$

De esta última definición, podemos obtener la siguiente consecuencia: si la torsión $\tau(s) = 0$, entonces

$$Tw(C, V) := \frac{1}{2\pi} \int_0^\ell \theta'(s) ds = \frac{1}{2\pi} (\theta(\ell) - \theta(0)) = \frac{2m\pi}{2\pi} = m$$

donde $m \in \mathbb{Z}^+$. Este hecho pone de manifiesto que el número de enrollamiento es una cantidad entera.

Si, en este mismo contexto, consideramos el caso particular en que $V(s) = n(s)$, entonces $\theta'(s) = 0$, de forma que el número de enrollamiento estaría determinado solamente por la torsión de la curva:

$$Tw(C, V) := \frac{1}{2\pi} \int_0^\ell \tau(s) ds$$

Pero claro, puesto que la torsión vale 0, podemos observar claramente que el número de enrollamiento es 0 (como era de esperar).

Número de retorcimiento

Para finalizar, hablaremos del último parámetro, el **número de retorcimiento**, denotado por $Wr(C)$. Se define en función de una curva, en este caso la curva C , porque este parámetro sólo depende de una sola curva.

El número de retorcimiento no admite una interpretación intuitiva que sea sencilla, pero no obstante, la fórmula de White nos va a permitir comprender qué es este número y lo que representa en relación a la banda B .

En cualquier representación de C , habrá un número de cruces. Para cada uno de esos cruces se adjunta un signo, como en el caso del número de enlace. Sin embargo, a diferencia del número de enlace, este número puede variar dependiendo de los diferentes puntos de vista de la misma curva.

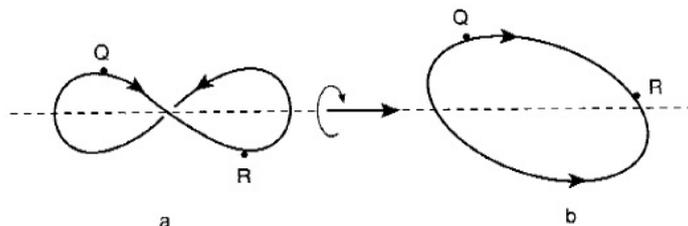


Figura 2.5: En (a), el número de retorcimiento es -1; y en (b), es 0.

Si el ADN se encuentra en un plano a excepción de unos pocos lugares en los que ésta se cruza sobre sí misma, entonces el número de retorcimiento es el total del número de signos correspondiente a cada auto-cruce.

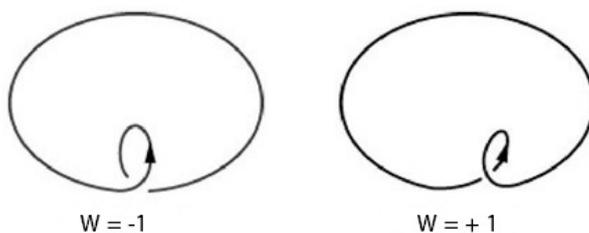


Figura 2.6: Número de retorcimiento.

Destacaremos el hecho de que cuanto mayor sea el valor absoluto del número de retorcimiento, más compacta estará la molécula de ADN. Cabe decir también, que este valor no es un invariante topológico, pues depende de la geometría que tenga la representación que se haga de la curva C , y por tanto, es una cantidad geométrica.

Pasemos ahora a dar una definición geométrica, en términos de una integral, del número de retorcimiento. Para ello, tal y como hemos hecho con los parámetros anteriores, vamos a considerar la notación empleada en la introducción del principio de la sección.

Consideremos entonces la variedad 2-dimensional $C \times C \in \mathbb{R}^6$, donde la curva C está parametrizada por el parámetro arco s (para distinguir el parámetro arco de cada curva C que forma la 2-variedad, usaremos los parámetros s_1 y s_2 , respectivamente). Tomamos ahora la siguiente aplicación en la esfera \mathbb{S}^2 :

$$e : C \times C \longrightarrow \mathbb{S}^2$$

que viene dada por

$$e(x, y) := \frac{y - x}{|y - x|}$$

Con esto, se define el número de retorcimiento, en términos de la geometría diferencial, como:

$$Wr(C) := \frac{1}{4\pi} \iint_{C \times C} e^*(d\Sigma)$$

donde $e^*(d\Sigma)$ es el pullback del elemento de área $d\Sigma$ de la esfera \mathbb{S}^2 bajo la aplicación en la esfera e anterior.

Justifiquemos la definición. En primer lugar, $d\Sigma$ es un tensor (0,2) en la esfera, es decir, una 2-forma. Por tanto, $e^*(d\Sigma)$ es un tensor (0,2) en la variedad $C \times C$, equivalentemente, una 2-forma. Recordemos de la sección 2.1 que

$$\left\{ (dx_{i_1} \wedge \dots \wedge dx_{i_k})_p \quad \text{tal que} \quad i_1 < \dots < i_k, \quad i_j \in \{1, \dots, n\} \right\}$$

es una base de $\Lambda^k(\mathbb{R}_p^n)^*$. Por tanto, como consecuencia de ello, existe una función real diferenciable $\lambda : [0, \ell] \times [0, \ell] \rightarrow \mathbb{R}$ de forma que podemos escribir

$$e^*(d\Sigma) = \lambda(s_1, s_2)(ds_1 \wedge ds_2)$$

Pasemos ahora a determinar el valor de $\lambda(s_1, s_2)$. Para ello, tengamos en cuenta que $ds_1 \wedge ds_2 = ds_1 \otimes ds_2 - ds_2 \otimes ds_1$. Así:

$$\lambda(s_1, s_2) = e^*(d\Sigma)(\partial s_1, \partial s_2) = d\Sigma_{e(s_1, s_2)}(de(\partial s_1), de(\partial s_2)) = \det \left(\frac{\partial e}{\partial s_1}, \frac{\partial e}{\partial s_2}, e(s_1, s_2) \right)$$

donde la última igualdad se tiene al usar que e es el normal unitario a la esfera.

Con esto, podemos decir que $Wr(C)$ mide el área de $e(C \times C)$ en \mathbb{S}^2 ; y considerar esta otra definición del número de retorcimiento:

$$Wr(C) := \frac{1}{4\pi} \iint_{[0, \ell] \times [0, \ell]} \det \left(\frac{\partial e}{\partial s_1}, \frac{\partial e}{\partial s_2}, e(s_1, s_2) \right) ds_1 ds_2$$

De esta última definición podemos extraer un caso particular con una interpretación intuitiva: supongamos que la curva C es plana, entonces $e(C \times C)$ estará contenido en un círculo máximo de la esfera \mathbb{S}^2 , tal que $\left\{ \frac{\partial e}{\partial s_1}, \frac{\partial e}{\partial s_2}, e(s_1, s_2) \right\}$ son colineales. Por tanto, $\det \left(\frac{\partial e}{\partial s_1}, \frac{\partial e}{\partial s_2}, e(s_1, s_2) \right) = 0$ y, en consecuencia, $Wr(C) = 0$.

Para acabar este apartado, observemos la siguiente imagen, que recoge las tres definiciones anteriores, y en donde se puede ver cierta relación entre cada uno de estos conceptos.

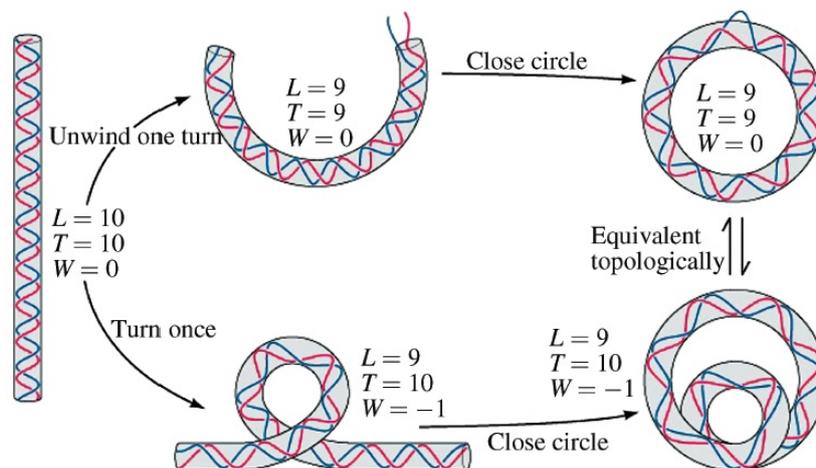


Figura 2.7: Número de enlace, de giros, y de retorcimiento.

2.3. Fórmula de White

El número de enlace, a pesar de ser una cantidad topológica, puede ser descompuesta en el número de enrollamiento y en el número de retorcimiento, que son dos cantidades geométricas, es decir, existe una relación entre un invariante topológico y dos cantidades geométricas. Ésta es la relación que viene dada en el siguiente resultado, conocido como la **Fórmula de White**, y que sirve para explicar el fenómeno del superenrollamiento.

Esto se debe a que, en cierto modo, el número de enlace es una medida de los cruces de las cadenas. Estos cruces se pueden dividir en cruces locales, que se producen a causa del enrollamiento helicoidal de una cadena de la molécula en torno al eje, que se mide con el número de enrollamiento o giros; y cruces distantes, que se producen debido al enrollamiento del eje sobre sí mismo, el cual se mide con el número de retorcimiento.

Para demostrar la fórmula de White nos vamos a basar en la demostración que aparece en [11]. El autor de esta demostración emplea distintas definiciones de Lk , y Wr de las que hemos visto, por lo que vamos a ver que ambas definiciones son equivalentes. Para ello vamos a heredar la notación que aparece en la introducción de la sección anterior.

Comenzamos con el número de enlace Lk . La definición que se da en dicha demostración ya ha sido considerada cuando se ha definido este parámetro, y es la siguiente:

$$Lk(C, C_V) = -\frac{1}{4\pi} \iint_{C \times C_V} g_1^*(\omega)$$

donde $g_1 : C \times C_V \rightarrow \mathbb{R}^3 \setminus \{0\}$ está definida por

$$g_1(s_1, s_2) := \alpha(s_1) - \alpha(s_2) - V(s_2)$$

y donde ω es la 2-forma en $\mathbb{R}^3 \setminus \{0\}$ definida mediante

$$\frac{1}{(x^2 + y^2 + z^2)^{3/2}}(xdy \wedge dz - ydx \wedge dz + zdx \wedge dy)$$

Como $g_1^*(\omega)$ es una 2-forma en la variedad $C \times C_V$, existe una función real diferenciable $\mu : [0, \ell] \times [0, \ell] \rightarrow \mathbb{R}$ tal que

$$g_1^*(\omega) = \mu(s_1, s_2)(ds_1 \wedge ds_2)$$

De este modo, considerando $\alpha(s_1) = (\alpha_1(s_1), \alpha_2(s_1), \alpha_3(s_1))$ y $\beta(s_2) = (\beta_1(s_2), \beta_2(s_2), \beta_3(s_2))$, tendríamos:

$$\begin{aligned} \mu(s_1, s_2) &= g_1^*(\omega)(\partial s_1, \partial s_2) = \omega_{g_1(s_1, s_2)}(dg_1(\partial s_1), dg_1(\partial s_2)) = \omega_{\alpha(s_1) - \beta(s_2)}(\alpha'(s_1), -\beta'(s_2)) = \\ &= \frac{-1}{|\alpha(s_1) - \beta(s_2)|^3} \begin{vmatrix} \alpha_1(s_1) - \beta_1(s_2) & \alpha_2(s_1) - \beta_2(s_2) & \alpha_3(s_1) - \beta_3(s_2) \\ \alpha'_1(s_1) & \alpha'_2(s_1) & \alpha'_3(s_1) \\ \beta'_1(s_2) & \beta'_2(s_2) & \beta'_3(s_2) \end{vmatrix} \end{aligned}$$

Por la analogía de la definición de Lk con la de Wr , razonando como se ha hecho en la justificación de la definición del número de retorcimiento Wr , podemos asegurar la existencia de una función real diferenciable $\lambda : [0, \ell] \times [0, \ell] \rightarrow \mathbb{R}$ de forma que podemos escribir

$$e^*(d\Sigma) = \lambda(s_1, s_2)(ds_1 \wedge ds_2)$$

donde el valor de λ viene determinado por:

$$\lambda(s_1, s_2) = \det \left(\frac{\partial e}{\partial s_1}, \frac{\partial e}{\partial s_2}, e(s_1, s_2) \right)$$

Para obtener la equivalencia entre las dos definiciones dadas del número de enlace Lk , vamos a operar con esta expresión de λ para llegar a expresarla en términos de μ . Para ello vamos a calcular con detalle $\frac{\partial e}{\partial s_1}$, dejando al cuidado del lector, aunque con cuentas análogas, el cálculo de $\frac{\partial e}{\partial s_2}$.

De este modo,

$$\begin{aligned} \frac{\partial e}{\partial s_1} &= \frac{\partial}{\partial s_1} \left(\frac{\beta(s_2) - \alpha(s_1)}{|\beta(s_2) - \alpha(s_1)|} \right) = \\ &= \frac{\partial}{\partial s_1} \left(\frac{1}{|\beta(s_2) - \alpha(s_1)|} \right) (\beta(s_2) - \alpha(s_1)) + \frac{1}{|\beta(s_2) - \alpha(s_1)|} (-\alpha'(s_1)) \end{aligned}$$

tal que,

$$\begin{aligned} \frac{\partial}{\partial s_1} \left(\frac{1}{|\beta(s_2) - \alpha(s_1)|} \right) &= \frac{\partial}{\partial s_1} \left(\frac{1}{\sqrt{(\beta_1(s_2) - \alpha_1(s_1))^2 + (\beta_2(s_2) - \alpha_2(s_1))^2 + (\beta_3(s_2) - \alpha_3(s_1))^2}} \right) = \\ &= \frac{\partial}{\partial s_1} \left((\beta_1(s_2) - \alpha_1(s_1))^2 + (\beta_2(s_2) - \alpha_2(s_1))^2 + (\beta_3(s_2) - \alpha_3(s_1))^2 \right)^{-1/2} = \\ &= -\frac{1}{2} |\beta(s_2) - \alpha(s_1)|^{-3} \left(2(\beta_1(s_2) - \alpha_1(s_1))(-\alpha'_1(s_1)) + \right. \\ &\quad \left. + 2(\beta_2(s_2) - \alpha_2(s_1))(-\alpha'_2(s_1)) + 2(\beta_3(s_2) - \alpha_3(s_1))(-\alpha'_3(s_1)) \right) = \\ &= -\frac{1}{2} |\beta(s_2) - \alpha(s_1)|^{-3} (-2) \left((\beta_1(s_2) - \alpha_1(s_1))(\alpha'_1(s_1)) + \right. \\ &\quad \left. + (\beta_2(s_2) - \alpha_2(s_1))(\alpha'_2(s_1)) + (\beta_3(s_2) - \alpha_3(s_1))(\alpha'_3(s_1)) \right) = \\ &= |\beta(s_2) - \alpha(s_1)|^{-3} \langle \beta(s_2) - \alpha(s_1), \alpha'(s_1) \rangle \end{aligned}$$

Por tanto, con todo

$$\frac{\partial e}{\partial s_1} = \frac{\langle \beta(s_2) - \alpha(s_1), \alpha'(s_1) \rangle}{|\beta(s_2) - \alpha(s_1)|^3} (\beta(s_2) - \alpha(s_1)) - \frac{\alpha'(s_1)}{|\beta(s_2) - \alpha(s_1)|}$$

Como hemos comentado, de forma análoga,

$$\frac{\partial e}{\partial s_2} = -\frac{\langle \beta(s_2) - \alpha(s_1), \beta'(s_2) \rangle}{|\beta(s_2) - \alpha(s_1)|^3} (\beta(s_2) - \alpha(s_1)) + \frac{\beta'(s_2)}{|\beta(s_2) - \alpha(s_1)|}$$

Así:

$$\begin{aligned}
\lambda(s_1, s_2) &= \det \left(\frac{\partial e}{\partial s_1}, \frac{\partial e}{\partial s_2}, e(s_1, s_2) \right) = \\
&= \det \left(\frac{\langle \beta(s_2) - \alpha(s_1), \alpha'(s_1) \rangle}{|\beta(s_2) - \alpha(s_1)|^3} (\beta(s_2) - \alpha(s_1)) - \frac{\alpha'(s_1)}{|\beta(s_2) - \alpha(s_1)|}, \right. \\
&\quad \left. - \frac{\langle \beta(s_2) - \alpha(s_1), \beta'(s_2) \rangle}{|\beta(s_2) - \alpha(s_1)|^3} (\beta(s_2) - \alpha(s_1)) + \frac{\beta'(s_2)}{|\beta(s_2) - \alpha(s_1)|}, \frac{\beta(s_2) - \alpha(s_1)}{|\beta(s_2) - \alpha(s_1)|} \right) = \\
&= \det \left(- \frac{\alpha'(s_1)}{|\beta(s_2) - \alpha(s_1)|}, \frac{\beta'(s_2)}{|\beta(s_2) - \alpha(s_1)|}, \frac{\beta(s_2) - \alpha(s_1)}{|\beta(s_2) - \alpha(s_1)|} \right) = \\
&= \frac{1}{|\alpha(s_1) - \beta(s_2)|^3} \begin{vmatrix} \alpha'_1(s_1) & \alpha'_2(s_1) & \alpha'_3(s_1) \\ \beta'_1(s_2) & \beta'_2(s_2) & \beta'_3(s_2) \\ \alpha_1(s_1) - \beta_1(s_2) & \alpha_2(s_1) - \beta_2(s_2) & \alpha_3(s_1) - \beta_3(s_2) \end{vmatrix} = -\mu(s_1, s_2)
\end{aligned}$$

quedando demostrado que las definiciones de Lk son equivalentes.

Ahora estudiemos el número de retorcimiento Wr . En este caso, la definición de M.H.Eggar en [11] del número de retorcimiento es la siguiente:

$$Wr(C) = -\frac{1}{4\pi} \iint_{[0, \ell] \times [0, \ell] \setminus \Delta} g_0^*(\omega)$$

donde $g_0 : C \times C \rightarrow \mathbb{R}^3 \setminus \{0\}$ viene dada por

$$g_0(s_1, s_2) := \alpha(s_1) - \alpha(s_2)$$

siendo ω la 2-forma ya vista en la definición del número de enlace, y donde Δ es el conjunto $\{(s, s) \text{ tal que } s \in [0, \ell]\}$.

Como $g_0^*(\omega)$ es una 2-forma en la variedad $C \times C$, existe una función real diferenciable $\mu : [0, \ell] \times [0, \ell] \rightarrow \mathbb{R}$ tal que

$$g_0^*(\omega) = \mu(s_1, s_2)(ds_1 \wedge ds_2)$$

Así, considerando otra vez $\alpha(s) = (\alpha_1(s), \alpha_2(s), \alpha_3(s))$, nos queda:

$$\begin{aligned}
\mu(s_1, s_2) &= g_0^*(\omega)(\partial s_1, \partial s_2) = \omega_{g_0(s_1, s_2)}(dg_0(\partial s_1), dg_0(\partial s_2)) = \omega_{\alpha(s_1) - \alpha(s_2)}(\alpha'(s_1), -\alpha'(s_2)) = \\
&= \frac{-1}{|\alpha(s_1) - \alpha(s_2)|^3} \begin{vmatrix} \alpha_1(s_1) - \alpha_1(s_2) & \alpha_2(s_1) - \alpha_2(s_2) & \alpha_3(s_1) - \alpha_3(s_2) \\ \alpha'_1(s_1) & \alpha'_2(s_1) & \alpha'_3(s_1) \\ \alpha'_1(s_2) & \alpha'_2(s_2) & \alpha'_3(s_2) \end{vmatrix}
\end{aligned}$$

De nuevo, para obtener la equivalencia de las definiciones del número de retorcimiento Wr , vamos a operar con la expresión de λ :

$$\det \left(\frac{\partial e}{\partial s_1}, \frac{\partial e}{\partial s_2}, e(s_1, s_2) \right)$$

con el fin de expresar λ en función de μ . Para ello vamos a seguir el mismo esquema que hemos seguido en el estudio de la equivalencia de las definiciones de Lk .

Por un lado,

$$\begin{aligned} \frac{\partial e}{\partial s_1} &= \frac{\partial}{\partial s_1} \left(\frac{\alpha(s_2) - \alpha(s_1)}{|\alpha(s_2) - \alpha(s_1)|} \right) = \\ &= \frac{\partial}{\partial s_1} \left(\frac{1}{|\alpha(s_2) - \alpha(s_1)|} \right) (\alpha(s_2) - \alpha(s_1)) + \frac{1}{|\alpha(s_2) - \alpha(s_1)|} (-\alpha'(s_1)) \end{aligned}$$

Por otro lado,

$$\begin{aligned} \frac{\partial}{\partial s_1} \left(\frac{1}{|\alpha(s_2) - \alpha(s_1)|} \right) &= \frac{\partial}{\partial s_1} \left(\frac{1}{\sqrt{(\alpha_1(s_2) - \alpha_1(s_1))^2 + (\alpha_2(s_2) - \alpha_2(s_1))^2 + (\alpha_3(s_2) - \alpha_3(s_1))^2}} \right) = \\ &= \frac{\partial}{\partial s_1} \left((\alpha_1(s_2) - \alpha_1(s_1))^2 + (\alpha_2(s_2) - \alpha_2(s_1))^2 + (\alpha_3(s_2) - \alpha_3(s_1))^2 \right)^{-1/2} = \\ &= -\frac{1}{2} |\alpha(s_2) - \alpha(s_1)|^{-3} \left(2(\alpha_1(s_2) - \alpha_1(s_1))(-\alpha'_1(s_1)) + \right. \\ &\quad \left. + 2(\alpha_2(s_2) - \alpha_2(s_1))(-\alpha'_2(s_1)) + 2(\alpha_3(s_2) - \alpha_3(s_1))(-\alpha'_3(s_1)) \right) = \\ &= -\frac{1}{2} |\alpha(s_2) - \alpha(s_1)|^{-3} (-2) \left((\alpha_1(s_2) - \alpha_1(s_1))(\alpha'_1(s_1)) + \right. \\ &\quad \left. + (\alpha_2(s_2) - \alpha_2(s_1))(\alpha'_2(s_1)) + (\alpha_3(s_2) - \alpha_3(s_1))(\alpha'_3(s_1)) \right) = \\ &= |\alpha(s_2) - \alpha(s_1)|^{-3} \langle \alpha(s_2) - \alpha(s_1), \alpha'(s_1) \rangle \end{aligned}$$

En definitiva, con todo

$$\frac{\partial e}{\partial s_1} = \frac{\langle \alpha(s_2) - \alpha(s_1), \alpha'(s_1) \rangle}{|\alpha(s_2) - \alpha(s_1)|^3} (\alpha(s_2) - \alpha(s_1)) - \frac{\alpha'(s_1)}{|\alpha(s_2) - \alpha(s_1)|}$$

Análogamente,

$$\frac{\partial e}{\partial s_2} = -\frac{\langle \alpha(s_2) - \alpha(s_1), \alpha'(s_2) \rangle}{|\alpha(s_2) - \alpha(s_1)|^3} (\alpha(s_2) - \alpha(s_1)) + \frac{\alpha'(s_2)}{|\alpha(s_2) - \alpha(s_1)|}$$

De modo que:

$$\begin{aligned}
\lambda(s_1, s_2) &= \det \left(\frac{\partial e}{\partial s_1}, \frac{\partial e}{\partial s_2}, e(s_1, s_2) \right) = \\
&= \det \left(\frac{\langle \alpha(s_2) - \alpha(s_1), \alpha'(s_1) \rangle}{|\alpha(s_2) - \alpha(s_1)|^3} (\alpha(s_2) - \alpha(s_1)) - \frac{\alpha'(s_1)}{|\alpha(s_2) - \alpha(s_1)|}, \right. \\
&\quad \left. - \frac{\langle \alpha(s_2) - \alpha(s_1), \alpha'(s_2) \rangle}{|\alpha(s_2) - \alpha(s_1)|^3} (\alpha(s_2) - \alpha(s_1)) + \frac{\alpha'(s_2)}{|\alpha(s_2) - \alpha(s_1)|}, \frac{\alpha(s_2) - \alpha(s_1)}{|\alpha(s_2) - \alpha(s_1)|} \right) = \\
&= \det \left(- \frac{\alpha'(s_1)}{|\alpha(s_2) - \alpha(s_1)|}, \frac{\alpha'(s_2)}{|\alpha(s_2) - \alpha(s_1)|}, \frac{\alpha(s_2) - \alpha(s_1)}{|\alpha(s_2) - \alpha(s_1)|} \right) = \\
&= \frac{1}{|\alpha(s_1) - \alpha(s_2)|^3} \begin{vmatrix} \alpha'_1(s_1) & \alpha'_2(s_1) & \alpha'_3(s_1) \\ \alpha'_1(s_2) & \alpha'_2(s_2) & \alpha'_3(s_2) \\ \alpha_1(s_1) - \alpha_1(s_2) & \alpha_2(s_1) - \alpha_2(s_2) & \alpha_3(s_1) - \alpha_3(s_2) \end{vmatrix} = -\mu(s_1, s_2)
\end{aligned}$$

quedando demostrado que $\lambda(s_1, s_2) = -\mu(s_1, s_2)$ y, en consecuencia, que ambas definiciones de Wr son equivalentes.

Antes de pasar a demostrar la fórmula, enunciaremos el siguiente resultado, útil en la demostración.

Teorema 2.3.1 (Stokes). *Sea M^n una variedad n -dimensional, quizás con $\partial M \neq \emptyset$. En ∂M , la variedad M induce la siguiente orientación: dado un $p \in \partial M \subset M$, y ν el conormal exterior, con $\nu \in T_p M$, entonces*

$$\begin{aligned}
\{v_1, \dots, v_{n-1}\} \text{ está positivamente orientada en } T_p \partial M &\iff \\
\{\nu, v_1, \dots, v_{n-1}\} \text{ está positivamente orientada en } T_p M &
\end{aligned}$$

En esta situación, se verifica:

$$\int_M d\omega = \int_{\partial M} \omega$$

Después de todo lo que hemos visto, estamos en condiciones de afrontar la demostración de la fórmula de White, sin embargo, antes destacaremos que esta fórmula fue probada por Călugăreanu y White, y aplicada al ADN por Fuller. Sin más preámbulos, veamos dicha demostración:

Teorema 2.3.2 (Călugăreanu-White-Fuller). $Lk(C, C_V) = Tw(C, V) + Wr(C)$

Demostración. Consideremos la región siguiente:

$$Q_\varepsilon := \{(s, t, u) \in [0, \ell] \times [0, \ell] \times I \mid |s - t|^2 + u^2 \geq \varepsilon^2\}$$

en $[0, \ell] \times [0, \ell] \times I$, donde $I = [0, 1]$, y siendo ε suficientemente pequeño. Veamos su representación gráfica:

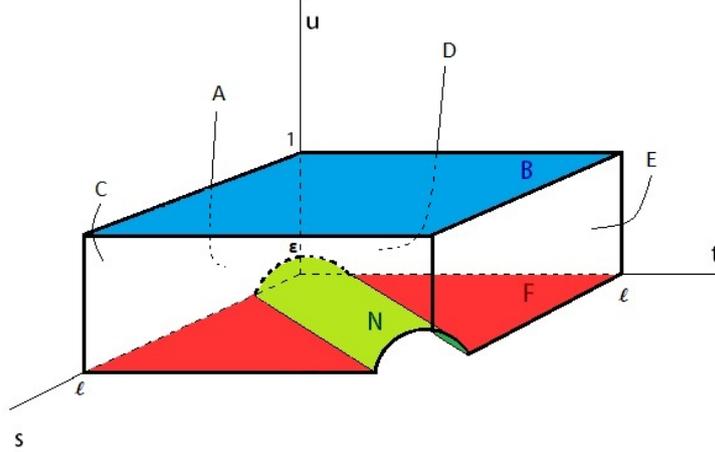


Figura 2.8: Q_ε

siendo A, B, C, D, E, F y N las caras de Q_ε , que es un paralelepípedo al que se le extrae, de la parte inferior, un *semi-cilindro* de radio ε (cuyo eje viene dado por la recta $(s, s, 0)$), denotado por N .

Como las curvas α y β son curvas cerradas, y la topología de $[0, \ell]$ es la de la identificación en la que unimos los extremos del intervalo, en realidad Q_ε es un toro sólido cuyo *espesor* viene determinado por el parámetro u . Por tanto, debemos identificar las caras: $A \equiv E$ y $C \equiv D$. De esta forma, la frontera de Q_ε queda determinada por:

$$\partial(Q_\varepsilon) = B \cup F \cup N$$

donde

$$\begin{aligned} B &:= \{(s, t, 1) \mid s \in [0, \ell], \quad t \in [0, \ell]\} \\ F &:= \{(s, t, 0) \mid s \in [0, \ell], \quad t \in [0, \ell]\} \\ N &:= \{(s, s - \varepsilon \cos(\theta), \varepsilon \sin(\theta)) \mid s \in [0, \ell], \quad \theta \in [0, \pi]\} \end{aligned}$$

Consideremos también $g : Q_\varepsilon \longrightarrow \mathbb{R}^3 \setminus \{0\}$ definida mediante

$$g(s, t, u) := \alpha(s) - \alpha(t) - uV(t)$$

De este modo, se aprecia que

$$\begin{aligned} g(s, t, 0) &= \alpha(s) - \alpha(t) = g_0(s, t) \\ g(s, t, 1) &= \alpha(s) - \alpha(t) - V(t) = g_1(s, t) \end{aligned}$$

Recordemos ahora la definición de la 2-forma $\omega \in \Lambda^2(\mathbb{R}^3 \setminus \{0\})$

$$\omega := \frac{1}{(x^2 + y^2 + z^2)^{3/2}}(xdy \wedge dz - ydx \wedge dz + zdx \wedge dy)$$

De esta forma, si consideramos $r^2 := x^2 + y^2 + z^2$, podemos redefinir ω como:

$$\omega := \frac{1}{r^3}(xdy \wedge dz - ydx \wedge dz + zdx \wedge dy)$$

y así:

$$\begin{aligned} d\omega &= \frac{\partial}{\partial x} \left(\frac{x}{r^3} \right) (dx \wedge dy \wedge dz) - \frac{\partial}{\partial y} \left(\frac{y}{r^3} \right) (-dx \wedge dy \wedge dz) + \frac{\partial}{\partial z} \left(\frac{z}{r^3} \right) ((-1)^2 dx \wedge dy \wedge dz) = \\ &= 3r^{-3} dx \wedge dy \wedge dz - 3r^{-4} \left[\left(x \frac{\partial r}{\partial x} + y \frac{\partial r}{\partial y} + z \frac{\partial r}{\partial z} \right) dx \wedge dy \wedge dz \right] = \\ &= 3r^{-3} dx \wedge dy \wedge dz - 3r^{-4} \left[\left(\frac{x^2}{r} + \frac{y^2}{r} + \frac{z^2}{r} \right) dx \wedge dy \wedge dz \right] = \\ &= 3r^{-3} dx \wedge dy \wedge dz - 3r^{-4} \left[\left(\frac{1}{r} r^2 \right) dx \wedge dy \wedge dz \right] = \\ &= 3r^{-3} dx \wedge dy \wedge dz - 3r^{-3} dx \wedge dy \wedge dz = 0 \end{aligned}$$

Por otro lado, recordando el teorema de Stokes, mencionado al principio de esta sección, elegimos una parametrización positivamente orientada para B , F y N , tomando el normal exterior como orientación, con el fin de aplicar dicho resultado. Con todo, conseguimos las siguientes igualdades:

$$\begin{aligned} \int_B g^*(\omega) - \int_F g^*(\omega) - \int_N g^*(\omega) &= \int_{\partial(Q_\varepsilon)} g^*(\omega) = \int_{\partial(Q_\varepsilon)} dg^*(\omega) = \\ &= \int_{\partial(Q_\varepsilon)} g^*(d\omega) = \int_{\partial(Q_\varepsilon)} 0 = 0 \end{aligned}$$

Luego:

$$\underbrace{\int_B g^*(\omega)}_{(1)} - \underbrace{\int_F g^*(\omega)}_{(2)} - \underbrace{\int_N g^*(\omega)}_{(3)} = 0$$

Observemos lo siguiente: Si tenemos en cuenta la definición de B , la integral de g^* sobre B es lo mismo que la integral de g_1^* en $[0, \ell] \times [0, \ell]$. Más aún, si observamos detenidamente la definición de Lk , obtenemos la última igualdad de la siguiente expresión:

$$(1) = \int_B g^*(\omega) = \int_{(s,t) \in [0,\ell] \times [0,\ell]} g_1^*(\omega) = -4\pi Lk(C, C_V)$$

Análogamente, si recordamos ahora la definición de F , la integral de g^* sobre F es lo mismo que la integral de g_0^* en $[0, \ell] \times [0, \ell]$, con la condición de que $|s - t| \geq \varepsilon$.

$$(2) = \int_F g^*(\omega) = \int_{(s,t) \in [0,\ell] \times [0,\ell]; |s-t| \geq \varepsilon} g_0^*(\omega)$$

Aún más, si tomamos límites cuando ε tiende a cero, tenemos que :

$$\lim_{\varepsilon \rightarrow 0} (2) = -4\pi W r(C)$$

por consiguiente, si probamos que

$$\lim_{\varepsilon \rightarrow 0} (3) = \lim_{\varepsilon \rightarrow 0} \int_N g * (\omega) = -4\pi T w(C, C_V)$$

acabaremos la demostración del teorema.

Bien, para ello vamos a evaluar, en primer lugar, $g^*(\omega)$. Como $g^*(\omega) \in \Lambda^2(Q_\varepsilon)$, las coordenadas de Q_ε son $\{ds, dt, du\}$, y $dx_i \wedge dx_j = 0$ si $i = j$, entonces existen funciones diferenciables $a(s, t, u), b(s, t, u), c(s, t, u)$ tales que

$$g^*(\omega) = a(s, t, u) ds \wedge dt + b(s, t, u) ds \wedge du + c(s, t, u) dt \wedge du$$

Determinemos ahora $a(s, t, u), b(s, t, u), c(s, t, u)$. En primer lugar, si consideramos $\alpha(s) = (\alpha_1(s), \alpha_2(s), \alpha_3(s))$ y $V(s) = (V_1(s), V_2(s), V_3(s))$, entonces g viene determinada por:

$$\begin{aligned} g(s, t, u) &= (\alpha_1(s) - \alpha_1(t) - uV_1(t), \alpha_2(s) - \alpha_2(t) - uV_2(t), \alpha_3(s) - \alpha_3(t) - uV_3(t)) = \\ &= (x_1(s, t, u), x_2(s, t, u), x_3(s, t, u)) \end{aligned}$$

De este modo:

$$\begin{aligned} a(s, t, u) &= g^*(\omega)(\partial s, \partial t) = \omega_{g(s,t,u)}(dg(\partial s), dg(\partial t)) = \\ &= \omega_{g(s,t,u)}(\alpha'(s), -\alpha'(t) - uV'(t)) = \\ &= \frac{1}{(x_1^2(s, t, u), x_2^2(s, t, u), x_3^2(s, t, u))^{3/2}} \det \left(g(s, t, u), \alpha'(s), (-\alpha'(t) - uV'(t)) \right) = \\ &= \frac{-1}{|g|^3} \begin{vmatrix} x_1(s, t, u) & x_2(s, t, u) & x_3(s, t, u) \\ \alpha'_1(s) & \alpha'_2(s) & \alpha'_3(s) \\ -\alpha'_1(t) - uV'_1(t) & -\alpha'_2(t) - uV'_2(t) & -\alpha'_3(t) - uV'_3(t) \end{vmatrix} = \\ &= \frac{-1}{|g|^3} \det \left(g(s, t, u), \alpha'(s), \alpha'(t) + uV'(t) \right) \end{aligned}$$

Análogamente:

$$b(s, t, u) = \frac{-1}{|g|^3} \det \left(g(s, t, u), \alpha'(s), V(t) \right)$$

$$c(s, t, u) = \frac{-1}{|g|^3} \det \left(g(s, t, u), \alpha'(t) + uV'(t), V(t) \right)$$

Luego:

$$\begin{aligned} -g^*(\omega) &= \frac{1}{|g|^3} \det \left(g(s, t, u), \alpha'(s), \alpha'(t) + uV'(t) \right) (ds \wedge dt) + \\ &+ \frac{1}{|g|^3} \det \left(g(s, t, u), \alpha'(s), V(t) \right) (ds \wedge du) + \\ &+ \frac{1}{|g|^3} \det \left(g(s, t, u), \alpha'(t) + uV'(t), V(t) \right) (dt \wedge du) \end{aligned}$$

Buscamos ahora expresar $g(s, t, u)$ en función de s y θ . En primer lugar, a partir de la definición de N , tenemos:

$$t = s - \varepsilon \cos(\theta) \qquad u = \varepsilon \cos(\theta)$$

Usando los desarrollos de Taylor en torno al punto s , nos queda:

$$\alpha(t) = \sum_{i=0}^{\infty} \frac{(-\varepsilon \cos(\theta))^i}{i!} \alpha^{(i)}(s)$$

$$V(t) = \sum_{i=0}^{\infty} \frac{(-\varepsilon \cos(\theta))^i}{i!} V^{(i)}(s)$$

$$\alpha'(t) = \sum_{i=0}^{\infty} \frac{(-\varepsilon \cos(\theta))^i}{i!} \alpha^{(i+1)}(s)$$

$$V'(t) = \sum_{i=0}^{\infty} \frac{(-\varepsilon \cos(\theta))^i}{i!} V^{(i+1)}(s)$$

Por otro lado, como $dt = a(s, \theta)ds + b(s, \theta)d\theta$, entonces:

$$a(s, \theta) = dt(\partial s) = 1 \quad b(s, \theta) = dt(\partial \theta) = \varepsilon \operatorname{sen}(\theta)$$

tal que $dt = ds + \varepsilon \operatorname{sen}(\theta)d\theta$. Análogamente, $du = \varepsilon \operatorname{cos}(\theta)d\theta$.

Con esto

$$\begin{aligned} g(s) &= \alpha(s) - \left(\alpha(s) + (-\varepsilon \operatorname{cos}(\theta))\alpha'(s) + \frac{(-\varepsilon \operatorname{cos}(\theta))^2}{2}\alpha''(s) + \dots + \frac{(-\varepsilon \operatorname{cos}(\theta))^i}{i!}\alpha^{(i)}(s) + \dots \right) + \\ &\quad - \varepsilon \operatorname{sen}(\theta) \left(V(s) + (-\varepsilon \operatorname{cos}(\theta))V'(s) + \frac{(-\varepsilon \operatorname{cos}(\theta))^2}{2}V''(s) + \dots + \frac{(-\varepsilon \operatorname{cos}(\theta))^i}{i!}V^{(i)}(s) + \dots \right) = \\ &= \varepsilon \left(\operatorname{cos}(\theta)\alpha'(s) + \frac{\varepsilon \operatorname{cos}^2(\theta)}{2}\alpha''(s) + \dots + \frac{\varepsilon^{i-1}\operatorname{cos}^i(\theta)}{i!}\alpha^{(i)}(s) + \dots \right) + \\ &\quad - \operatorname{sen}(\theta)V(s) \left(\varepsilon \operatorname{cos}(\theta)\operatorname{sen}(\theta)V'(s) + \frac{(\varepsilon \operatorname{cos}(\theta)\operatorname{sen}(\theta))^2}{2}V''(s) + \dots + \right. \\ &\quad \left. + \dots + \frac{(\varepsilon \operatorname{cos}(\theta)\operatorname{sen}(\theta))^i}{i!}V^{(i)}(s) + \dots \right) \end{aligned}$$

Si separamos $\alpha(s)$ y $V(s)$ en sus componentes, nos queda, para $j = 1, 2, 3$

$$\begin{aligned} x_j^2(s) &= \operatorname{cos}^2(\theta)\alpha_j(s) + \frac{\varepsilon^2 \operatorname{cos}^4(\theta)}{4}(\alpha_j''(s))^2 + \dots + \frac{\varepsilon^{2i-2}\operatorname{cos}^{2i}(\theta)}{(i!)^2}(\alpha^{(i)}(s))^2 + \dots + \\ &\quad + \dots + \operatorname{sen}^2(\theta)V_j^2(s) + \varepsilon^2 \operatorname{cos}^2(\theta)\operatorname{sen}^2(\theta)(V_j'(s))^2 + \dots + \\ &\quad + \frac{\varepsilon^{2i}\operatorname{cos}^{2i}(\theta)\operatorname{sen}^2(\theta)}{(i!)^2}(V_j^{(i)}(s))^2 + \text{dobles productos} + \dots \end{aligned}$$

de manera que, apartando los términos $o(\varepsilon)$,

$$\begin{aligned} x_1^2(s) + x_2^2(s) + x_3^2(s) &= \operatorname{cos}^2(\theta)|\alpha'(s)|^2 + \operatorname{sen}^2(\theta)|V(s)|^2 + o(\varepsilon) \\ &= \operatorname{cos}^2(\theta) + \operatorname{sen}^2(\theta) + o(\varepsilon) = 1 + o(\varepsilon) \end{aligned}$$

tal que $|g|^3 = \varepsilon^3$.

Por otro lado, apartando de nuevo los términos $o(\varepsilon)$:

$$\begin{aligned}
& \frac{1}{|g|^3} \det\left(g(s, t, u), \alpha'(s), \alpha'(t) + uV'(t)\right)(ds \wedge dt) = \\
& = \frac{1}{\varepsilon^3} \det\left(\alpha(s) - \alpha(t) - uV(t), \alpha'(s), \alpha'(t) + uV'(t)\right)(ds \wedge dt) = \\
& = \frac{1}{\varepsilon^3} \det\left(\alpha(s) - \alpha(t), \alpha'(s), \alpha'(t)\right)(\varepsilon \operatorname{sen}(\theta) ds \wedge d\theta) + \\
& \quad \frac{1}{\varepsilon^3} \det\left(-uV(t), \alpha'(s), \alpha'(t)\right)(\varepsilon \operatorname{sen}(\theta) ds \wedge d\theta) + \\
& \quad \frac{1}{\varepsilon^3} \det\left(\alpha(s) - \alpha(t), \alpha'(s), uV'(t)\right)(\varepsilon \operatorname{sen}(\theta) ds \wedge d\theta) + \\
& \quad \frac{1}{\varepsilon^3} \det\left(-uV(t), \alpha'(s), uV'(t)\right)(\varepsilon \operatorname{sen}(\theta) ds \wedge d\theta) = \\
& = \frac{-\varepsilon \operatorname{sen}(\theta)}{\varepsilon^3} \det\left(V(t), \alpha'(s), \alpha'(t)\right)(\varepsilon \operatorname{sen}(\theta) ds \wedge d\theta) + \\
& \quad \frac{-\varepsilon^2 \operatorname{sen}^2(\theta)}{\varepsilon^3} \det\left(V(t), \alpha'(s), V'(t)\right)(\varepsilon \operatorname{sen}(\theta) ds \wedge d\theta) = \\
& = \frac{1}{\varepsilon^3} \left(\varepsilon^3 \operatorname{sen}^2(\theta) \cos(\theta) \det\left(V(t), \alpha'(s), \alpha''(s)\right) - \varepsilon^3 \operatorname{sen}^3(\theta) \det\left(V(t), \alpha'(s), V'(t)\right) \right) (ds \wedge d\theta) = \\
& = \left(\operatorname{sen}^2(\theta) \cos(\theta) \det\left(V(t), \alpha'(s), \alpha''(s)\right) - \operatorname{sen}^3(\theta) \det\left(V(t), \alpha'(s), V'(t)\right) \right) (ds \wedge d\theta)
\end{aligned}$$

y, de manera análoga:

$$\begin{aligned}
& \frac{1}{|g|^3} \det\left(g(s, t, u), \alpha'(s), V(t)\right)(ds \wedge du) = \\
& = \left(-\operatorname{sen}(\theta) \cos^2(\theta) \det\left(V(t), \alpha'(s), V''(s)\right) + \frac{1}{2} \cos^3(\theta) \det\left(\alpha''(s), \alpha'(s), V(t)\right) - \right. \\
& \quad \left. - \operatorname{sen}(\theta) \cos^2(\theta) \det\left(V'(t), \alpha'(s), V(s)\right) \right) (ds \wedge d\theta)
\end{aligned}$$

y

$$\begin{aligned} \frac{1}{|g|^3} \det\left(g(s, t, u), V(t), \alpha'(s)\right)(dt \wedge du) = \\ = \left(\cos^3(\theta) \det\left(\alpha'(s), V(s), \alpha''(s)\right) - \operatorname{sen}(\theta) \cos^2(\theta) \det\left(\alpha'(s), V(t), V'(s)\right) - \right. \\ \left. - \operatorname{sen}(\theta) \cos^2(\theta) \det\left(V(t), V'(s), \alpha'(s)\right) + \frac{1}{2} \cos^3(\theta) \det\left(\alpha''(s), V(t), \alpha'(s)\right) - \right. \\ \left. - \operatorname{sen}(\theta) \cos^2(\theta) \det\left(V'(t), V(s), \alpha'(s)\right) \right) (ds \wedge d\theta) \end{aligned}$$

Así, con todo, tenemos que:

$$\int_N g^*(\omega) = -2 \int_{[0, \ell]} \det\left(V'(s), \alpha'(s), V(s)\right) ds + 0(\varepsilon)$$

y, por tanto,

$$\lim_{\varepsilon \rightarrow 0} \int_N g^*(\omega) = -4\pi Tw(C, C_V)$$

como se quería demostrar. \square

2.4. Interpretación de la fórmula de White

Consideremos una molécula de ADN circular sin superenrollamiento, es decir, una molécula de ADN relajada. En estas condiciones, al número de enlace de dicha molécula lo denotaremos⁵ por Lk^0 . El Lk^0 para esta molécula es la cantidad de pares de bases dividida por 10.5 (que es el número de pares de bases que tiene el ADN por vuelta de la hélice). Así, por ejemplo, para una molécula de ADN circular de 10500 pares de bases, $Lk^0 = +1000$ (el signo es positivo porque el giro del ADN es dextrógiro⁶). En este contexto, se tiene que $Wr = 0$, y que $Lk = Tw$, tal que, si tampoco hay tensiones, $Lk = Lk^0 = Tw$.

Supongamos ahora que se corta el ADN y uno de los extremos se gira una vuelta a derechas antes de volver a sellar los extremos, entonces se obtiene una molécula de ADN con una vuelta menos de la que le correspondería por el número de pares de bases, es decir, $Lk < Lk^0$. En esta situación, la molécula ya no es estable porque el número de vueltas total es menor y hay más pares de bases por vuelta de los que correspondería. Por tanto, hay que aumentar Tw en una unidad para volver al número ideal de pares de bases

⁵Abusando de la notación, omitiremos las curvas de las que depende cada parámetro, y sólo emplearemos las letras mayúsculas que representan al número de enlace, al número de enrollamiento, y al número de retorcimiento.

⁶Se dice que la doble hélice es una espiral dextrógira, es decir, cada una de las cadenas de nucleótidos gira a derechas.

por vuelta o, lo que es lo mismo, acercar Tw a Lk^0 para liberar la tensión producida por el desenrollamiento en una vuelta, y esto se consigue haciendo que el ADN se retuerza la vuelta que falta.

Debido a la relación establecida por la fórmula de White, como lo que se busca es subir el valor de Tw , lo que se debe hacer es disminuir el valor de Wr . La extensión del superenrollamiento se mide por la diferencia entre Lk y Lk^0 , $\Delta Lk = Lk - Lk^0$, que se conoce como *diferencia de enlace*. Si la diferencia de enlace es significativamente diferente de cero, entonces el ADN estará superenrollado, distinguiendo dos casos:

- Si $\Delta Lk = Lk - Lk^0 < 0$. Supongamos, por ejemplo, que $Lk = Lk^0 - 1$. Entonces $Tw < Lk^0$ (pues en la molécula estable $Lk = Tw$). Como Lk es invariable a menos que se rompan las hebras, para acercarse al número ideal de vueltas hay que hacer $T + 1$, tal que Wr tendrá que bajar a $Wr - 1$ para que Lk siga constante. Es decir, Wr contrarresta Tw para que Lk permanezca invariable. Así, puesto que Tw aumenta y Wr disminuye, se dice que se produce un superenrollamiento negativo.
- Si $\Delta Lk = Lk - Lk^0 > 0$. Con un razonamiento análogo al anterior, vemos que Wr debe aumentar para que la molécula sea estable. Por tanto, en este caso, el superenrollamiento es positivo.

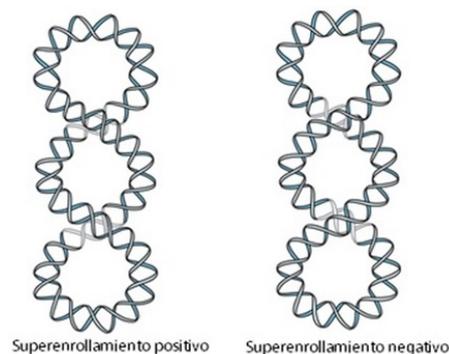


Figura 2.9: Superenrollamiento

Con esto, los superenrollamientos a derechas (i.e. izquierdas) suponen decrecimientos (i.e. crecimientos) de Wr , y por eso reciben el nombre de superenrollamientos negativos (i.e. positivos), haciendo que Tw aumente (i.e. disminuya) en tantas unidades como superenrollamientos.

Gracias a la interpretación de la fórmula de White y a la observación de los cambios en los distintos parámetros (usando las técnicas de electroforesis en gel), también se puede demostrar que realmente la molécula de ADN tiene una periodicidad helicoidal de alrededor de 10.5 pares de bases por vuelta [25]

Capítulo 3

Cambios en la geometría y topología por la acción de enzimas, y diseño de fármacos.

El ADN de todas las células se encuentra bajo tensión helicoidal o superenrollamiento. Este fenómeno de superenrollamiento ya ha sido descrito en este documento, y hemos destacado que su principal función es la de *empaquetar* la molécula de ADN para que ocupe menos espacio del que realmente debería ocupar.

En apartados anteriores, para visualizar este proceso, hemos puesto de ejemplo el cable de un teléfono. Bien, pues otro modelo, que se asemeja más al caso que queremos estudiar, es el de una cuerda enrollada sobre sí misma. Si los extremos de la cuerda permanecen fijos, por ejemplo cogidos por nuestras manos, y giramos la cuerda sobre sí misma, generamos una tensión que, después de unas cuantas vueltas, se libera mediante la formación de superenrollamientos.

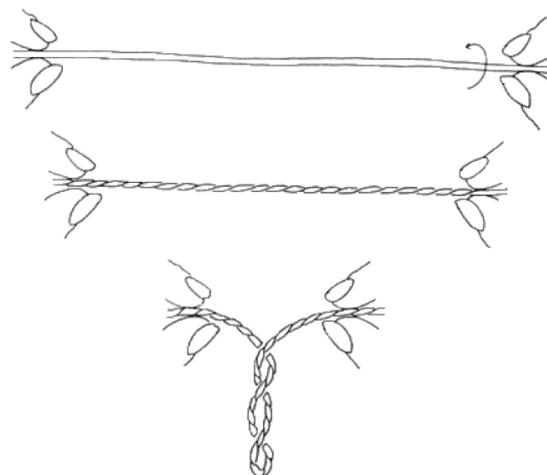


Figura 3.1: Superenrollamiento

El superenrollamiento del que estamos hablando es conocido como **superenrollamiento plectonémico**, y es generado, principalmente, por la acción de unas enzimas llamadas topoisomerasas. Estas enzimas se encargan de introducir tensiones estructurales en la molécula de ADN, modificando de este modo la topología de la molécula.

El superenrollamiento también permite que se acerquen regiones de la molécula alejadas en su estructura primaria (estructura lineal del polímero), algo clave en los procesos de replicación y transcripción. Como nota final, las moléculas pueden sufrir superenrollamiento tanto positivo como negativo, dependiendo del sentido de la torsión.

3.1. Topoisomerasas

Ya ha aparecido el nombre de topoisomerasas en el texto. Es el momento ahora de describir con detalle estas enzimas y su función. Exceptuando los procesos de replicación o transcripción, el resto del tiempo el ADN debe mantenerse en una configuración lo más compacta posible, sin embargo cuando ocurren dichos procesos, el ADN debe ser separado en hebras simples, de forma que los estados de enrollamiento dificultan seriamente este proceso.

Cuando la naturaleza de doble hélice del ADN fue resuelta por Watson y Crick, los autores notaron que debía haber algún mecanismo que pudiera resolver o solventar los enredos que surgirían de esta característica estructural. Así, la necesidad de la enzima en cuestión fue reconocida incluso antes de ser descubierta.

Las **topoisomerasas** son enzimas que regulan la topología del ADN, ya sea enredándolo, para permitir que se almacene de manera más compacta; o desenredándolo para que controle la síntesis de proteínas y para facilitar la replicación del mismo. Son esenciales, por tanto, para el metabolismo del ADN.

Las topoisomerasas, en base al número de cadenas de ADN que cortan por ronda de acción, se clasifican en dos grandes grupos:

En primer lugar, consideraremos la **topoisomerasa I**. Este tipo de topoisomerasa se une a la molécula de ADN, corta una de las cadenas del ADN, permitiendo de esta forma liberar las tensiones internas debidas a un excesivo enrollamiento o a un enrollamiento deficiente; y luego pasa esta cadena a través de la otra antes de volver a sellar los extremos de la cadena de ADN cortada. La energía liberada por la rotura se usa para sellar el DNA. Este proceso provoca una disminución de una unidad en el número de enrollamiento Tw , y por tanto, debido a la relación establecida por la fórmula de White, un cambio en el número de superenrollamientos Wr .

La topoisomerasa I actúa sobre ADN con superenrollamiento negativo (no positivo). Por tanto, podemos decir que es una enzima *relajante*.

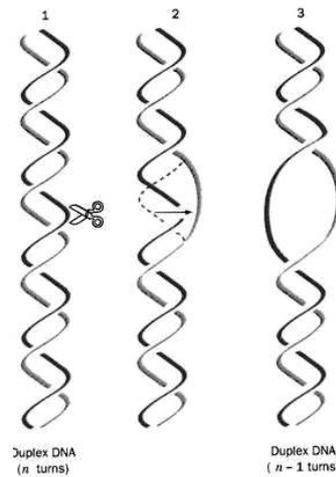


Figura 3.2: Acción de la topoisomerasa I

Veamos ahora las **topoisomerasas II**. Estas topoisomerasas se distinguen de las anteriores por ser unas enzimas que cortan ambas cadenas de la doble hélice de ADN simultáneamente. Cuando se produce la rotura de la doble hélice del ADN, uno de los extremos se adhiere a la enzima, y el otro extremo pasa por encima de la cadena que permanece intacta. Al igual que la topoisomerasa de tipo I, las topoisomerasas son enzimas *relajantes*.

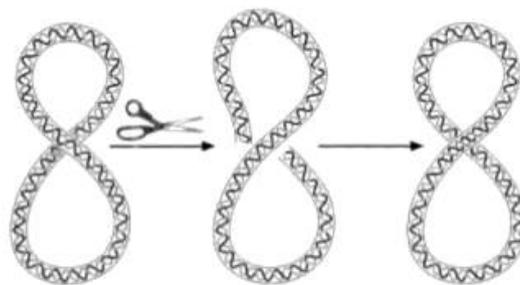


Figura 3.3: Acción de la topoisomerasa II

3.2. Recombinación. Recombinación específica de sitio

La recombinación del ADN se refiere al fenómeno por el cual dos filamentos parentales de ADN se juntan dando como resultado un intercambio de las porciones de sus respectivos filamentos. Este proceso conduce a nuevas moléculas de ADN que contienen una mezcla de la información genética de cada filamento parental.

La **recombinación específica de sitio**, o site-specific recombination (en inglés), es una de las maneras que tiene la naturaleza de alterar el código genético de un organismo, bien moviendo un fragmento de ADN a otra posición dentro de la misma molécula, o bien mediante la inserción de un fragmento de ADN ajeno en un genoma huésped.

La recombinación específica de sitio tiene un papel vital en el ciclo de vida de ciertos virus, que utilizan este proceso para insertar el ADN viral en el ADN de un organismo huésped y así infectarlo. Estas reacciones de recombinación tienen lugar en prácticamente todas las células, pero las funciones están especializadas y varían en gran medida según la especie.

Algunas estrategias de ingeniería genética común requieren una modificación permanente del genoma, por lo que la introducción de genes debe ser algo sofisticado y específico. Si la integración del gen se hace de forma aleatoria, en general, la expresión de los genes es impredecible y pueden aparecer mutaciones asociadas indeseadas. Por ejemplo, los requisitos moleculares en el campo de células madre son muy estrictos. Por tanto, la recombinación específica de sitio es un proceso que aporta toda la especificidad que necesita el experimento.

Todos los sistemas de recombinación específica de sitio requieren una enzima denominada **recombinasa** y una secuencia única de ADN corta (entre 20 y 200 pares de bases) sobre la que actúa la recombinasa (el sitio de la recombinación). La enzima se encarga de reorganizar la secuencia de bases del segmento de ADN expuesto a su acción.

La recombinación específica de sitio ocurre cuando un par de *zonas* de la misma molécula de ADN, o incluso de distintas moléculas de ADN, se unen en presencia de la recombinasa. Esta etapa de la reacción se llama **sinapsis**. La parte de la molécula de ADN que queda expuesta a la recombinasa recibe el nombre de **sustrato**. Así, el complejo formado por este sustrato y la enzima recibe el nombre de **sinaptosoma** o **complejo sináptico**.

La siguiente imagen muestra una micrografía de electrones de un complejo sináptico formado por la enzima de recombinación Tn3 resolvasa cuando se hace reaccionar con una molécula de ADN circular.

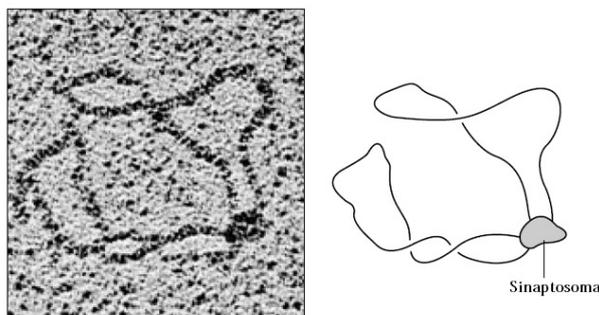


Figura 3.4: Recombinación específica en sitio (vista al microscopio, y una representación)

Después de formar el sinaptosoma, se produce el proceso de recombinación: la enzima rompe la doble cadena en ciertos puntos o sitios específicos y combina los extremos resultantes del corte también de una forma específica. Después de esto, el complejo sináptico se disocia poco a poco y la enzima recombinasa libera la molécula de ADN. El resultado es un cambio visible de la topología del ADN. A esta nueva topología la llamaremos **producto** de la reacción.

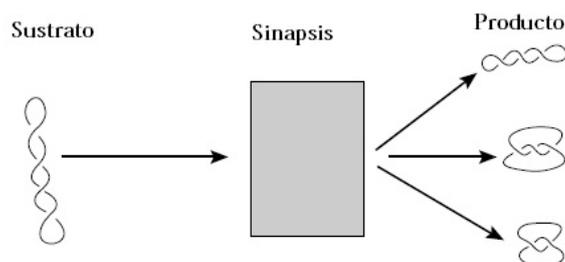


Figura 3.5: Sustrato-Sinapsis-Producto

Este proceso de unión de la enzima con el sustrato, la sinapsis, y posteriormente la obtención del producto, puede producirse una sola vez, lo que se conoce como **recombinación procesiva**; o puede ocurrir que cuando se disocia el sinaptosoma, la enzima libera esta *parte* de la molécula de ADN y se une a otra zona de la misma molécula. A este proceso le llamaremos **recombinación distributiva**. La mayoría de las recombinasas sólo pueden hacer un tipo de recombinación, pero también hay algunas enzimas capaces de realizar ambas recombinaciones.

Como ya se ha dicho antes, la recombinación específica de sitio implica cambios topológicos en el sustrato (molécula de ADN que entra en contacto con la enzima recombinasa). Para identificar estos cambios topológicos, se llevan a cabo experimentos sobre sustrato de ADN circular, obtenidas mediante el uso de técnicas de clonación. Se realiza este experimento en un gran número de moléculas circulares con el fin de obtener una cantidad observable de productos.

El experimento *in vitro* sigue el siguiente procedimiento: se prepara un sustrato circular, controlando la cantidad de superenrollamiento de las moléculas de sustrato. Estas moléculas de sustrato se hacen reaccionar con una alta concentración de enzima purificada, y los productos de reacción se fraccionan por electroforesis en gel. Dado que todas las moléculas tienen el mismo peso molecular, la electroforesis discrimina entre unas moléculas y otras basándose en diferencias sutiles de la geometría (superenrollamiento) y topología (nudo y el tipo de enlace) de las moléculas de ADN.

Es importante destacar que en el lugar en que ocurre la recombinación específica de sitio, la secuencia de pares de bases no es palindrómica en general, es decir, la lectura de izquierda a derecha de la secuencia de bases es distinta de la lectura de derecha a izquierda. Esto induce una orientación local en el sustrato, que favorece la posterior interpretación que se observa en el producto.

La estrategia experimental que se sigue para estudiar estas reacciones es el enfoque topológico para la enzimología. El enfoque topológico para la enzimología usa los poderes descriptivos y analíticos de la topología y la geometría en un esfuerzo indirecto por determinar el mecanismo de acción de la enzima sobre la molécula de ADN.

El enfoque topológico para la enzimología supone un reto interesante para las matemáticas: a partir de los cambios observados en la geometría y la topología del ADN, ¿cómo podemos deducir el mecanismo enzimático que se ha llevado a cabo en la recombinación específica de sitio?. Esta pregunta requiere la construcción de modelos matemáticos que simulen la acción de la enzima, y el uso de éstos para analizar los resultados de los experimentos realizados. La cantidad de nudos, o la cantidad de enlaces, en la molécula de ADN producto nos aporta información sobre la actuación de la enzima. Además, el uso de modelos matemáticos ayuda a determinar las conclusiones biológicas finales extraídas de los resultados experimentales. Así, la respuesta a la pregunta anterior se encuentra en la *Teoría de nudos*, que se estudiará en el apéndice final.

3.3. Recombinasas

Las recombinasas son las enzimas que aparecen en la recombinación específica de sitio, como ya hemos visto en la sección anterior. Estas enzimas se dividen principalmente, atendiendo al aminoácido particular del polímero de la proteína que cataliza la reacción de escisión, en dos familias: las serinas, o más conocidas como resolvasas; y las tirosinas, también conocidas como integrasas.

Las recombinasas serina y tirosina también difieren en su mecanismo de corte y sellado del ADN en los sitios específicos de esta recombinación.

3.4. Enfermedades y diseño de fármacos

El descubrimiento del ADN, su análisis y la comprensión de su estructura y niveles de complejidad, han permitido a muchos científicos llevar a cabo todo tipo de investigaciones y avanzar en muchos campos de la ciencia cuyos fines son mejorar las condiciones de vida de los seres vivos. Estos avances han sido posibles gracias al estudio *in vitro* de las reacciones de ciertas enzimas sobre la molécula de ADN, acciones que han podido ser deducidas con el apoyo de la teoría de nudos y su modelización matemática.

Entre estos avances debemos mencionar los logros en genética, y en la resolución de ciertas afecciones, pues conociendo la composición de un individuo se pueden averiguar sus deficiencias y buscar entonces alternativas que permitan curar enfermedades.

Por un lado, las topoisomerasas han sido el núcleo para el tratamiento de ciertas enfermedades. La girasa bacteriana (un tipo de enzima dentro del grupo de las topoisomerasas II) y la topoisomerasa IV son la base de dos tipos de medicamentos antibióticos: quinolonas y cumarinas. Estos antibióticos se usan para la cura de varias enfermedades, como la neumonía, la tuberculosis y la malaria, mediante la inhibición de la replicación del ADN de la bacteria responsable de dicha enfermedad.

Las quinolonas son un grupo de agentes quimioterapéuticos sintéticos, es decir, no proceden de organismos vivos. Se usan para tratar infecciones urinarias, gastroenteritis (diarrea del viajero), infecciones por *Salmonella*, infecciones respiratorias como neumonías, teniendo un papel relevante en el tratamiento de la tuberculosis cuando no pueden administrarse los fármacos primarios [28]. Su acción se basa en la inhibición de la acción de las girasas.

Por su parte, las cumarinas provocan evidentes mejoras en tratamientos antitumorales, antiarrítmicos, antiinflamatorios, antisépticos, analgésicos, y contra la hipertensión, la osteoporosis y el VIH.

Por otro lado, las topoisomerasas eucariotas I y II (las topoisomerasas de las células animales y, por tanto, de las células humanas) son la base de una cantidad, a día de hoy creciente, de fármacos contra el cáncer que actúan para inhibir estas enzimas mediante el bloqueo de la reacción que sella las roturas en el ADN, destruyendo así las células cancerosas.

En el tratamiento del cáncer, tanto la topoisomerasa I como la topoisomerasa II pueden ser dianas terapéuticas¹. La camptotecina y sus derivados² modifican la función de la topoisomerasa I. Se ha observado una muy buena correlación entre la actividad antitumoral de varios derivados de la camptotecina en el tratamiento de la leucemia de ratón y la interferencia de la actividad de la topoisomerasa. Sin embargo, las camptotecinas pueden causar lesiones letales en las células. De hecho, se sigue profundizando en este tipo de tratamientos, pues los inhibidores de las topoisomerasas I y II pueden afectar indistintamente al ADN huésped como al ADN objetivo.

Otro caso a determinar es el tratamiento del VIH. Hasta ahora, la infección con el virus de inmunodeficiencia humana (VIH) era considerada irreversible. Este agente es un retrovirus, que inserta su ADN en el genoma de las células humanas a las que infecta. Según [27], Hauber (científico experto del Instituto Heinrich Pette de Virología e Inmunología Experimental de la ciudad alemana de Hamburgo) existe una pequeña esperanza de desarrollar, en unos 10 años, una terapia para humanos. El próximo paso sería, según el experto, realizar experimentos con ratones durante tres años, y seguidamente se iniciarán amplios estudios con pacientes en Hamburgo.

El genoma del virus del sida tiene en ambos extremos dos secuencias idénticas, perfectamente conocidas, que casi no sufren modificaciones aún en el caso de mutaciones del VIH. Puesto que la recombinasa Cre reconoce una secuencia de ADN similar a la que

¹Una diana terapéutica o blanco molecular puede definirse como el lugar del organismo donde un fármaco ejerce su acción.

²La camptotecina es un fármaco empleado para tratar varios tipos de tumores malignos. Se obtiene de la corteza y las hojas del árbol de origen chino, *Camptotheca acuminata*, y su acción se debe a la inhibición de la enzima topoisomerasa I. En los estudios realizados se ha comprobado su eficacia en el tratamiento de los tumores, pero su utilidad práctica queda limitada por sus efectos adversos, por lo que se han desarrollado diferentes derivados (los más usados, topotecán y el irinotecán) que mejoran sus propiedades.

aparece en el genoma del VIH, biólogos moleculares “cultivaron”, a partir de la recombinasa Cre, la recombinasa Tre, que ataca exactamente las secuencias de los extremos del virus del sida.

Ya sabemos que las recombinasas reconocen una secuencia de bases de ADN específica y actúa justamente en ese lugar. Como si de una tijera se tratase, estas proteínas cortan el ADN en determinados sitios y vuelven a unir. *El fragmento recortado es degradado por la propia célula*, explicó Hauber. La célula queda libre del virus.

Este tratamiento, en caso de que el método pueda ser desarrollado en una terapia, sería una buena vía para tratar el SIDA, aunque sería un tratamiento complicado. Según explicó Hauber, se deberían obtener células madre de la sangre del paciente y limpiarlas de los virus en el laboratorio. Las células tratadas se reintroducirían en el paciente para regenerar su sistema inmune. De este modo, aunque no se pueda eliminar el VIH totalmente del cuerpo, se espera lograr un retroceso de la carga viral tal que la infección quede bajo control.

Sin embargo, el principal problema de este tratamiento es que se produzcan mutaciones en las secuencias de los extremos del virus del sida, pues las recombinasas podrían ser adaptadas rápidamente. En estos casos, se dispondría de una serie de enzimas diferentes, que podrían ser combinadas y desarrolladas rápidamente. Es por ello por lo que este método de cura está aún por estudiar y desarrollar.

Apéndice A

Teoría de nudos

A.0.1. Introducción

Las palabras nudos y enlaces¹ designan objetos cotidianos, cuya utilidad práctica no necesita ser detallada, que han sido usados por el hombre desde los tiempos más antiguos, pero el significado y estudio de los nudos y enlaces como elementos u objetos matemáticos es relativamente reciente.

El primer científico que usó un concepto estrechamente relacionado con lo que se conoce, en la actualidad, como un enlace matemático, fue Gauss, para trabajar en el campo magnético inducido. Pero además de Gauss, aparecen otros nombres como los precursores de la Teoría de nudos, como Tait y Kelvin.

En 1867 Lord Kelvin propuso su teoría del “átomo vorticial”, inspirado por un artículo sobre vórtices de Helmholtz y por el trabajo de Riemann sobre funciones abelianas. Kelvin albergaba la esperanza de justificar las propiedades químicas de los átomos a partir de los anudamientos producidos entre los vórtices de éter.

Diez años después, el también británico Peter G. Tait publicó su primer artículo sobre la enumeración de los tipos posibles de nudos. Tait comenzó a colaborar con el matemático C. N. Little, quienes consiguieron, por el año 1900, clasificar los nudos con diez entrecruzamientos. Tait llegó a distinguir nudos diferentes pero no pudo demostrar matemáticamente esa diferencia.

A comienzos del siglo XX, los matemáticos J. W. Alexander II y Kurt Reidemeister fueron los principales impulsores del desarrollo de la Teoría de nudos, destacando, en la segunda mitad del siglo, los matemáticos John H. Conway y L. H. Kauffman.

La teoría de nudos es la rama de la topología que se encarga de estudiar el objeto matemático que abstrae la noción cotidiana de nudo, y cuyo principal objetivo es saber distinguir un nudo de otro.

La motivación original de los *fundadores* de la Teoría de nudos era crear un cuadro de nudos y enlaces. A día de hoy, más de seis mil millones de nudos y enlaces han sido tabulados desde los inicios de esta teoría.

¹Los enlaces están formados por nudos enlazados entre sí.

Un nudo matemático no es más que la idea abstracta de un nudo *normal* en el que pegamos los extremos de la cuerda que estamos empleando. Así, bajo esta idea, se puede desarrollar toda una teoría de nudos que abarca desde un estudio topológico hasta el estudio de sus aplicaciones en campos como la biología, la química o la física.

A.0.2. Nudos y primeras definiciones

El concepto matemático de nudo es una abstracción, como ya sabemos, de la siguiente imagen:

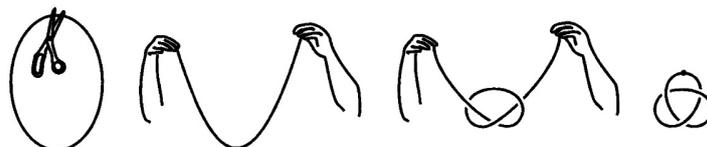


Figura A.1: Idea de un nudo con una cuerda.

Matemáticamente, no tiene importancia que la cuerda que usemos sea más o menos gruesa, ni más o menos larga. Lo esencial es que sea una línea recta, pues, topológicamente hablando, lo que tendremos será una circunferencia en el espacio.

Definición A.0.1. Se define un **nudo** como una aplicación continua e inyectiva de la forma $f : \mathbb{S}^1 \rightarrow \mathbb{R}^3$, y que sea un homeomorfismo en la imagen, es decir, un embebimiento topológico.

Denotaremos un nudo como $N = f(\mathbb{S}^1)$.

Una definición más intuitiva de nudo es la de encajar una circunferencia en el espacio.

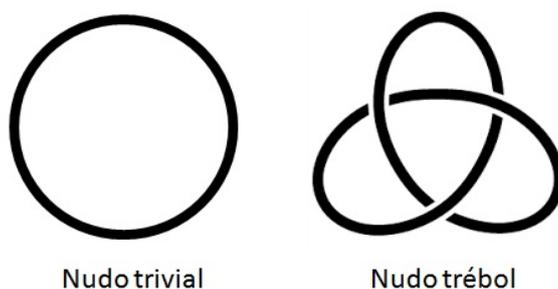


Figura A.2: Nudo trivial y nudo trébol.

Dado un nudo hecho con una cuerda, cualquier modificación que se haga de la cuerda, sin cortarla ni volverla a pegar, origina un nudo equivalente al de partida. Para definir la equivalencia de nudos de forma matemática, introducimos los siguientes conceptos:

Definición A.0.2. Sean X e Y dos espacios topológicos, y $f, g : X \rightarrow Y$ dos funciones continuas. Consideremos el intervalo $I = [0, 1]$ como un subespacio de \mathbb{R} , con la topología heredada; y el $X \times I$, con la topología producto. Se define una **homotopía** como una función continua dada por $F : X \times I \rightarrow Y$, tal que

$$\begin{cases} F(x, 0) := f(x) \\ F(x, 1) := g(x) \end{cases}$$

En este contexto, diremos que f y g son **homotópicas**.

Definición A.0.3. Una homotopía $F : X \times I \rightarrow Y$ se dice **isotopía** si $F|_{X \times \{t\}}$ es un homeomorfismo para todo $t \in I$.

Definición A.0.4. Si $f, g : Y \rightarrow X$ son dos embebimientos de Y en X , decimos que f y g son un **ambiente isotópico** si existe una isotopía $F : X \times I \rightarrow X$ tal que

$$\begin{cases} F(x, 0) := x & \forall x \in X \\ F(f(y), 1) := g(y) & \forall y \in Y \end{cases}$$

En este contexto, diremos que X es el **espacio ambiente**, y que F es la **isotopía ambiente**.

De este modo, podemos definir matemáticamente la equivalencia entre nudos:

Definición A.0.5. Dos **nudos** $f, g : \mathbb{S}^1 \rightarrow \mathbb{R}^3$ se dicen **equivalentes** si son un ambiente isotópico. Dicho de otra forma, si existe un homeomorfismo de \mathbb{R}^3 a \mathbb{R}^3 que transforma un nudo en otro.

La colección de nudos equivalentes se llama **tipo de nudo**.

Dejando caer una cuerda con un nudo sobre una superficie plana, obtenemos una representación bidimensional del nudo. Por tanto, la mejor forma de representar un nudo es mediante su proyección en un plano, dibujando los cruces de forma que sea visible qué trozo del nudo pasa por debajo o por encima; y distinguiendo el tramo superior en un cruce, marcándolo con un trazo continuo, y el tramo inferior, dibujándolo discontinuo o cortado.

Estas representaciones se conocen por el nombre de **diagramas planos del nudo**. Atendiendo al diagrama del nudo, podemos distinguir dos tipos de nudos. Si el nudo puede representarse mediante un diagrama con un número finito de cruces, se dice que es un **nudo dócil**, y su **proyección** se llama **regular**. De lo contrario, si la única representación que admite el nudo tiene infinitos cruces, diremos que es un **nudo salvaje**. Centraremos toda la teoría que desarrollaremos en los nudos dóciles.

La Teoría de nudos también estudia los enlaces, que responde a la idea de varias cuerdas cerradas, que pueden estar anudadas y/o enlazadas entre sí. Se define entonces, matemáticamente, un enlace como:

Definición A.0.6. Un **enlace** E de n componentes es un encaje de n circunferencias disjuntas en el espacio. Es decir, es la imagen de la siguiente aplicación $l : \cup_{i=1}^n \mathbb{S}^1 \rightarrow \mathbb{R}^3$.

Así, un nudo es un enlace con una sola componente.

A continuación, y para terminar, definiremos una operación de nudos, llamada suma conexas.

Definición A.0.7. Dados dos nudos, orientados, N_1 y N_2 , se define la **suma conexas** de los dos **nudos**, denotada por $N_1 \# N_2$, como el nudo resultante de suprimir un segmento de cada nudo y pegar los extremos resultantes de ambos nudos, a lo largo de este segmento, de forma que las orientaciones se conserven.

De este modo, diremos que un **nudo** es **primo** si no puede expresarse como la suma conexas de dos nudos triviales.

La suma conexas tiene las siguientes propiedades:

1. La suma conexas está bien definida, salvo equivalencia. Es decir, si $N_1 \cong M_1$, y $N_2 \cong M_2$, entonces $N_1 \# N_2 \cong M_1 \# M_2$
2. Es una propiedad asociativa, es decir: $(N_1 \# N_2) \# N_3 \cong N_1 \# (N_2 \# N_3)$
3. Es conmutativa, es decir: $N_1 \# N_2 \cong N_2 \# N_1$

donde \cong denota la equivalencia de nudos.

A.0.3. Equivalencia de nudos

Equivalencia de nudos. Movimientos de Reidemeister

Ya hemos visto que el diagrama de un nudo es la proyección sobre un plano de un nudo tridimensional. Ahora bien, bajo ciertas condiciones, puede probarse que si dos nudos tienen el mismo diagrama, entonces ambos nudos son equivalentes. Pero, ¿cómo podemos confirmar la igualdad de los diagramas de dos nudos?.

Para comprobar la equivalencia de los diagramas de dos nudos, a finales de los años veinte, Kurt Reidemeister publicó unas transformaciones, llamadas con su propio nombre, “desplazamientos de Reidemeister”.

Los movimientos de Reidemeister son tres y cada uno de ellos opera en una pequeña región del diagrama del nudo, de la siguiente forma:

- I. Girar o desenroscar en cualquier dirección.
- II. Agregar o remover dos cruces por arriba, o por debajo, consecutivos.
- III. Movimiento triangular.

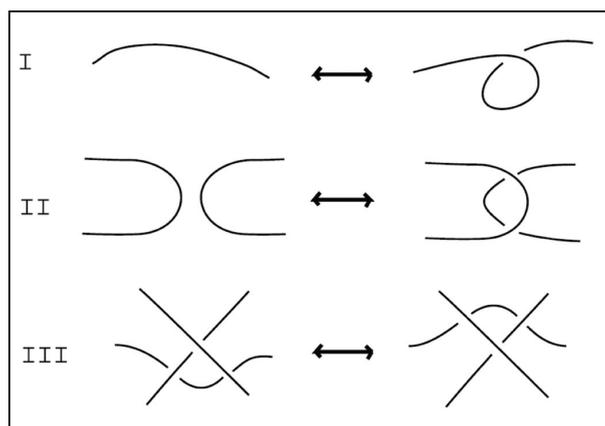


Figura A.3: Movimientos de Reidemeister.

Así, podemos escribir formalmente el resultado anteriormente comentado (cuya demostración puede ver en [7]):

Teorema A.0.8. Teorema de Reidemeister. *Dos nudos son equivalentes si y sólo si existe una secuencia finita de movimientos de Reidemeister que permiten pasar del diagrama de un nudo al diagrama del otro.*

Equivalencia de nudos. Invariantes de los nudos

Una forma de determinar si dos nudos son equivalentes o no es encontrando alguna propiedad de los nudos que no cambie cuando estos se deforman, y que sirva para distinguir nudos no equivalentes. A dichas propiedades se les llama **invariantes de nudos**. En general, estos invariantes son unidireccionales, es decir:

Si dos nudos son equivalentes \Rightarrow Sus invariantes son iguales.

Gracias a la ayuda de los movimientos de Reidemeister se han podido definir varios invariantes de nudos. Algunos de ellos se describen a continuación.

El número de enlace. Ya conocemos la definición de número de enlace. Recordemos entonces que, dado un enlace E de dos componentes, el número de enlace de E viene dado por:

$$Lk(E) = \frac{1}{2} \sum_{\text{cruces}} \pm 1$$

Con esto, se establece el siguiente resultado, cuya demostración se puede ver en [6], [21].

Teorema A.0.9. *Si dos enlaces, con una orientación fija, son equivalentes, entonces todos sus diagramas tienen el mismo número de enlace.*

Número mínimo de cruces. Consideremos un diagrama regular D de un nudo (o enlace) N , que tiene un número finito de cruces. Denotemos a esta última cantidad por $c(D)$. Se tiene entonces que $c(D)$ no es un invariante. Sin embargo, si consideramos el conjunto de todas las representaciones o diagramas regulares del nudo N , denotado por \mathcal{D} ; y denotamos por $c(N)$ el número mínimo de cruces de todos los diagramas regulares, es decir:

$$c(N) = \min_{\mathcal{D}} \{c(D)\}$$

entonces $c(N)$ sí es un invariante de nudos (ver [6], [21]).

Además, se sabe lo siguiente:

1. El nudo trivial es el único nudo que posee diagramas regulares tales que $c(D) = 0, 1, 2$.
2. Entre todos los nudos y enlaces, el nudo trébol es el único con $c(N) = 3$.
3. Conjetura: Sean N_1 y N_2 dos nudos arbitrarios. Entonces, $c(N_1 \# N_2) = c(N_1) + c(N_2)$

Número de puentes. Consideramos, de nuevo, un diagrama regular D de un nudo (o enlace) N , de manera que en cada punto de intersección de D cortamos un pequeño segmento AB del cruce. De este modo, el resultado de llevar a cabo este proceso repetidas veces es un conjunto de curvas disconexas. A estos segmentos AB que eliminamos los llamaremos **puentes**. Así, la colección de puentes de D se denota por $br(D)$.



Figura A.4: Puentes en una representación D de un nudo.

Tal y como sucede con el número de cruces, el número de puentes de un diagrama regular D , no es un invariante de nudos, pero sí lo es, como pasaba en el caso anterior, el mínimo número de puentes de todos los diagramas regulares del nudo N dado, denotado por:

$$br(N) = \min_{\mathcal{D}}\{br(D)\}$$

siendo \mathcal{D} el conjunto de todas las representaciones o diagramas regulares del nudo N (ver [6]).

Sobre esta *característica* se conoce lo siguiente:

1. Un nudo N es trivial $\iff br(N) = 1$.
2. Todos los nudos con número de puente 2 forman una familia especial de nudos que ha sido totalmente clasificada.
3. Si N_1 y N_2 son dos nudos (o enlaces) arbitrarios, entonces,

$$br(N_1 \# N_2) = br(N_1) + br(N_2) - 1$$

4. Conjetura: Si N es un nudo (o enlace) arbitrario, entonces, $c(N) \geq 3(br(N) - 1)$

Equivalencia de nudos. Polinomios

Otro tipo de invariantes de nudos son los polinomios que podemos asociar a los mismos. Los polinomios más destacados son el polinomio de Alexander, el polinomio de Conway, el polinomio de Jones, y el polinomio de Kauffman.

En particular, el polinomio de Alexander, asocia a cada nudo N un polinomio $\Delta_N(t)$ en una variable t . Además, Alexander probó que si dos nudos (o enlaces) N_1 y N_2 orientados eran equivalentes, entonces $\Delta_{N_1}(t) \approx \Delta_{N_2}(t)$, donde \approx significa igualdad salvo un múltiplo de $\pm t^n$, para algún entero n .

No desarrollaremos la teoría de estos polinomios, pues excedería la extensión de este documento, pero se puede ampliar sobre ellos en [6], [8], [7], [19].

A.0.4. Nudos en biología

Volviendo a la recombinación específica de sitio, ya introducida en este texto, supongamos que conocemos el sustrato y la topología del producto obtenido en dicha recombinación. Nuestro objetivo ahora es utilizar argumentos topológicos para precisar el proceso o mecanismo de recombinación que se ha producido. Para ello nos vamos a ayudar de la teoría de nudos, y del modelo de ovillos, el cual desarrollaremos en la próxima sección.

Cabe destacar el hecho de que las *soluciones* matemáticas dadas para explicar la acción de las enzimas en la recombinación específica de sitio pueden ser descartadas por consideraciones biológicas.

Nudos en biología. Modelo de ovillos

Comencemos dando algunas definiciones extraídas de [1], [2], [4], [6], [16], [17]:

Definición A.0.10. Un n -ovillo es una pareja (\mathbb{S}^2, T) , donde \mathbb{S}^2 es la esfera unidad en \mathbb{R}^3 , y T es un conjunto de n hebras propiamente encajadas en \mathbb{S}^2 .

En lo que sigue, consideraremos sólo 2-ovillos, y nos referiremos a ellos simplemente por ovillos. Además, en lugar de trabajar con los ovillos en la esfera, consideraremos su **diagrama de ovillo** asociado a T , que no es otra que la proyección de sus hebras en el plano YZ .

Las hebras del ovillo, en la esfera, tienen *puntos de origen*, denotados por $\{NO, NE, SO, SE\}$. La siguiente imagen ilustra las definiciones dadas hasta ahora:

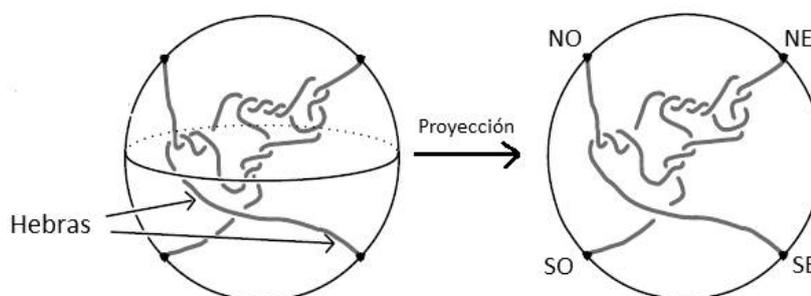


Figura A.5: 2-ovillo y su diagrama.

Definición A.0.11. Diremos que dos **ovillos**, O_1 y O_2 , son **equivalentes** si existe una isotopía entre O_1 y O_2 que deja fija la frontera de la esfera unidad.

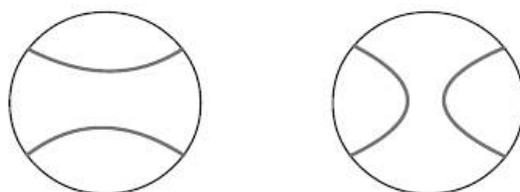
En otras palabras, O_1 y O_2 son equivalentes si podemos deformar las hebras de O_1 y sobreponerlas en las de O_2 , sin romper dichas hebras.

Con esta definición, y recordando los movimientos de Reidemeister, tenemos:

Teorema A.0.12. Dos ovillos, O_1 y O_2 , son equivalentes \iff podemos pasar del diagrama de O_1 al diagrama de O_2 usando, para ello, una sucesión finita de movimientos de Reidemeister.

Por otro lado, los ovillos pueden clasificarse en tres tipos:

1. Los ovillos anudados localmente, que son ovillos que poseen una hebra anudada.
2. Los ovillos racionales. Un ovillo es racional si, mediante una isotopía que no deja fija la frontera de \mathbb{S}^2 , puede ser transformado en uno de los ovillos siguientes:

Figura A.6: El ovrillo 0, y el ovrillo ∞

La familia de los ovillos racionales ha sido extensamente estudiada al existir un correspondencia biyectiva entre los números racionales extendidos, $\mathbb{Q} \cup \{\infty\}$, y los ovillos racionales. Es decir, dos ovillos se les asocia el mismo número racional \iff son ovillos equivalentes.

Con esto, viendo más sobre ello en [1], [4], [5], [6], [32], dado un número racional, el ovrillo asociado a él puede ser construido a partir de la expansión de dicho número como una fracción continua. En este sentido, las fracciones continuas determinan los diagramas de ovillos.

3. Los ovillos primos, que son aquellos ovillos que no son ni racionales ni localmente anudados.

Veamos ahora las operaciones que podemos realizar con los ovillos. Son las siguientes:

- (S). **Suma de ovillos.** Dados dos ovillos O_1 y O_2 , la suma de estos ovillos consiste en identificar los puntos NE_{O_1} con NO_{O_2} , y SE_{O_1} con SO_{O_2} . De este modo, el elemento neutro de esta operación es (0).

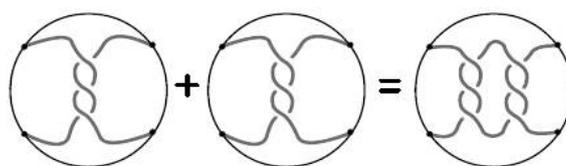


Figura A.7: Suma de dos ovillos

- (N). **Numerador de un ovrillo.** Esta operación consiste en unir con una hebra, sin nudos, los puntos del norte y los del sur, esto es, unir NO con NE , y SO con SE .
- (D). **Denominador de un ovrillo.** Esta operación consiste en unir con una hebra, sin nudos, los puntos occidentales y los orientales, esto es, unir NO con SO , y NE con SE .

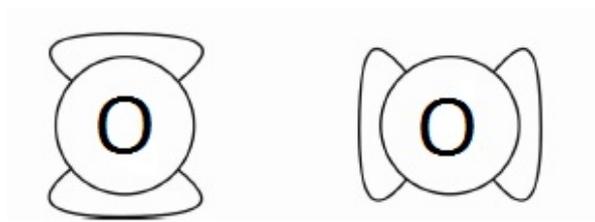


Figura A.8: El numerador y denominador de un ovillo.

Por último, veamos un ejemplo de un ovillo racional particular.

Definición A.0.13. Un **4-plat** es un nudo o enlace que admite un diagrama basado en el trenzado de cuatro hebras cerradas.

La clasificación de los 4-plats muestra que cada 4-plat K se caracteriza por un vector $\langle c_1, c_2, \dots, c_{2n+1} \rangle$ (el vector de Conway) que tiene un número impar de entradas de números enteros, de manera que, c_1 y c_{2n+1} son distintos de cero, siendo n el número de cruces del nudo, y cada c_i determina el tipo de nudo o enlace.

A partir del vector de Conway, podemos obtener un número racional que clasifique a un 4-plat K , de la siguiente forma:

$$\frac{\beta}{\alpha} = \frac{1}{c_1 + \frac{1}{c_2 + \dots + \frac{1}{c_{2n+1}}}}$$

donde $0 < \beta < \alpha$. Luego, los 4-plats son ovillos racionales.

Nudos en biología. Usando el modelo de ovillos

Si recordamos la recombinación específica de sitio, dentro de la región controlada por la enzima, ésta rompe el ADN en cada sitio y recombina los extremos mediante el intercambio de los mismos. Consideremos entonces esta enzima como la esfera unidad. Por otro lado, el sinaptosoma, que consiste en la unión de la enzima con el ADN, forma un ovillo de dos hebras, es decir, un 2-ovillo. Por tanto, podemos modelizar el proceso de la recombinación específica de sitio con el modelo de ovillos. Aunque, para ello debemos asumir una serie de suposiciones, biológicas y matemáticas.

S.1. *El mecanismo de la enzima en un solo evento de recombinación específica de sitio, se supondrá constante, e independiente de la geometría (superenrollamiento) y la topología del sustrato.*

Por otra parte, la recombinación tiene lugar en su totalidad dentro del sinaptosoma, y la configuración del sustrato que permanece fuera de este sinaptosoma permanece invariante, mientras que las hebras que están en el interior de éste se rompen y recombinan en su interior y en su límite.

S.2. *El sinaptosoma es un 2-ovillo que puede ser descompuesto matemáticamente en la suma de dos ovillos $O_B + O_P$, siendo el ovillo O_P el ovillo padre que sufre la acción de la enzima en la recombinación específica de sitio; y O_B , la parte del ovillo en la que, aun estando en el interior del sinaptosoma, no actúa la enzima.*

Como O_P es la parte que sufre el cambio, éste será reemplazado por otro ovillo denotado por O_R , obtenido después de que la enzima libere el ADN y concluya así su acción sobre él. Con esto:

$$\begin{aligned} \text{pre-recombinación} &= O_B + O_P. \\ \text{post-recombinación} &= O_B + O_R. \end{aligned}$$

S.3. Denotando por O_F la parte del sustrato que permanece fuera del sinaptosoma, y por tanto invariante, podemos suponer que *todo el complejo sináptico se obtiene del numerador del ovillo resultado de la suma de los ovillos $O_F + \text{sinaptosoma}$.*

De este modo, un único evento de recombinación específica de sitio en una molécula de sustrato circular produce dos ecuaciones de recombinación:

$$\begin{aligned} \text{Ecuación del sustrato: } &N(O_F + O_B + O_P) = \text{sustrato.} \\ \text{Ecuación del producto: } &N(O_F + O_B + O_R) = \text{producto.} \end{aligned}$$

Con todo lo que hemos visto, podría surgir una pregunta, *¿es posible, conociendo la topología tanto del sustrato como del producto de varios eventos de recombinación específica de sitio, deducir el mecanismo de la acción de la enzima?* La respuesta es que sí, y las enzimas estudiadas han sido, por ejemplo Tn3 resolvasa (Ver [2], [4], [5], [6]), tirosina recombinasa (Ver en [1]), serina recombinasa (Ver en [1]), y topoisomerasas (Ver en [17]), phaage lambda (Ver en [4]).

A.0.5. Software TangleSolve

Atendiendo a [33] [34], *TangleSolve* es un aplicación, programada en lenguaje Java, que implementa el modelo de ovillos, y ofrece una interfaz gráfica fácil de usar para el análisis y la visualización de los mecanismos de recombinación.

La interfaz de usuario consta de tres paneles: entrada (donde se introducen los datos experimentales), selección (donde aparecen las posibles soluciones que el programa aporta) y pantalla (donde, una vez seleccionada la mejor solución, se visualiza la misma).

El algoritmo de solución se basa en unos resultados, recientemente desarrollados, de topología tridimensional. Además, *TangleSolve* asume que los sustratos y productos de la recombinación específica de sitio son 4-plats, que los sitios de recombinación están encerrados dentro del ovillo padre, y que tras dicha recombinación el ovillo padre se convierte en un ovillo recombinante.

De este modo, *TangleSolve* permite que el modelo de ovillos sea una herramienta eficiente utilizada en el diseño de experimentos.

Bibliografía

- [1] DOROTHY BUCK. *DNA Topology*. Proceedings of Symposia in Applied Mathematics, 2009, VOL. 66.
- [2] J.C.GÓMEZ LARRAÑAGA, H. CABRERA IBARRA. *Nudos en biología*. Miscelánea matemática, 2007, 44, PÁGS. 53-66.
- [3] A. FERRÁNDEZ, M.A.HERNÁNDEZ CIFRE, J.A.PASTOR. *Algunos aspectos matemáticos de la doble estructura helicoidal del ADN*. La gaceta de la RSME, 2003, VOL. 6.3, PÁGS. 1-14.
- [4] C. ERNST, D.W.SUMNERS. *A calculus for rational tangles: applications to DNA recombination*. Mathematical Proceedings of the Cambridge Philosophical Society, 1990, 108, PÁGS. 489-515.
- [5] WITT SUMNERS. *Lifting the Curtain: Using Topology to probe the hidden action of enzymes*. Notices of the AMS, 1995, VOL.42, NUM. 5.
- [6] K. MURASUGI. *Knot theory and its applications*. Birkhäuser, Boston, 1996, ISBN: 0-8176-3817-2, 3-7643-3817-2.
- [7] COLIN ADAMS, ROBERT FRANZOSA. *Introduction to Topology. Pure and Applied*. Pearson, 2009, ISBN: 978-81-317-2692-1.
- [8] COLIN C. ADAMS. *The knot book. An elementary introduction to the mathematical theory of knots*. W. H. Freeman and Company, 2004, ISBN: 978-0821836781.
- [9] RICHARD H. CROWELL, RALPH H. FOX. *Introduction to knot theory*. Springer-Verlag, 1963, ISBN: 0-387-90272-4, 3-510-90272-4.
- [10] THOMAS M. DELVIN. *Bioquímica. Libro de texto con aplicaciones clínicas*. Reverté, 2004, ISBN: 84-291-7208-4.
- [11] M.H.EDGAR *On White's formula*. Journal of knot theory and its ramifications, 2000, VOL. 9, NUM.5, PÁGS. 611-615.
- [12] WITT SUMNERS. *Untangling DNA*. The mathematical intelligencer, 1990, VOL. 12, NUM.3.
- [13] CONTENIDO WEB. Ver: [en. wikipedia. org/wiki/DNA](http://en.wikipedia.org/wiki/DNA)

- [14] MIGUEL VICENTE. *La dama ausente: Rosalind Franklin y la doble hélice*.
- [15] RAFAEL ANDRÉS ALEMÁN BERENGUER, ESTRELLA JORNET GIL. *La fascinante matemática de los nudos*. Revista de Didáctica de las Matemáticas, 2011, VOL. 76, PÁGS. 47-54, ISBN: 1887-1984.
- [16] ROBERT PLANQUÉ. *The role of knot theory in DNA research*. 2000.
- [17] ISABEL K. DAREY. *Modeling protein-DNA complexes with tangles*. Mathematics department, University of Iowa.
- [18] WENJING ZHENG, CHRISTINE GALLOY, BERNARD HALLET, MARIEL VÁZQUEZ. *The Tangle Model for Site-specific Recombination: A Computer Interface and the TnpI-IRS Recombination System*. Mathematical Subject Classification 2000.
- [19] MARÍA TERESA LOZANO IMÍZCOZ *La teoría de nudos en el siglo XX*. Universidad de Zaragoza.
- [20] LUPING WANG XIAO *Introducción a la teoría de nudos*. Universidad de Cantabria.
- [21] JOSÉ LUIS CISNEROS MOLINA *Introducción a la teoría de nudos*. V jornadas de física y matemáticas, Universidad autonómica de Ciudad Juárez, 2011.
- [22] RENZO L. RICCA, BERNARDO NIPOTI *Gauss's linking number revisited*. Journal of knot theory and its ramifications, 2011, VOL. 20, NUM. 10, PÁGS. 1325-1343.
- [23] CONTENIDO WEB. *Quimioterapias para la Replicación del ADN*. <http://themedicalbiochemistrypage.org/es/dna-sp.php>
- [24] MANFREDO P. DO CARMO. *Differential forms and applications*. Springer-Verlag, 1994, ISBN: 0-387-57618-5, 3-540-57618-5.
- [25] JAMES D. WATSON, TANIA A. BAKER, STEPHEN P. BELL, ALEXANDER GANN, MICHAEL LEVINE, RICHARD LOSICK. *Biología molecular del gen*. Medica Panamericana, 2008, ISBN: 978-84-7903-505-1.
- [26] HELENA CURTIS, N. SUE BARNES, ADRIANA SCHNEK, GRACIELA FLORES. *Invitación a la biología*. Medica Panamericana, 2006, ISBN: 978-9500604475.
- [27] CONTENIDO WEB. *Científicos alemanes eliminan el VIH de células infectadas*. <http://www.jornada.unam.mx/2007/06/29/index.php?section=ciencias&article=a03n1cie>
- [28] JOHNSON J, HADAD D, BOOM W Y COLABORADORES. *Uso de Fluoroquinolonas en el Tratamiento de la Tuberculosis*. <http://www.bago.com/BagoArg/Biblio/infectoweb451.htm>
- [29] CONTENIDO WEB. *El enlace puente de hidrógeno*. https://en.wikipedia.org/wiki/Hydrogen_bond

-
- [30] Lehninger DAVID L. NELSON, MICHAEL M. COX. *Lehninger: Principio de bioquímica*. Ediciones Omega, 2009, EDIC. 5, ISBN: 978-84-282-1486-5
- [31] ERIC S. LANDER, MICHAEL S. WATERMAN. *Calculating the secrets of life. Applications of the Mathematical Sciences in Molecular Biology*. National academy press. Washington, DC, 1995.
- [32] CONTENIDO WEB. *Mathematical Background*. <http://ewok.sfsu.edu/TangleSolve/MathematicalBackground.html>
- [33] YUKI SAKA, MARIWL VÁZQUEZ. *TangleSolve: topological analysis of site-specific recombination*. Mathematics Department, U.C, USA, 2002.
- [34] CONTENIDO WEB. *TangleSolve*. <http://bio.math.berkeley.edu/TangleSolve/>