Topology of DNA.

Duplex DNA is made up of two molecular strands that are twisted together in a right handed helix.

The two strands are joined together by bonds.

Each strand is made up of alternating sugars and phosphates, and each sugar has one of four bases attached to it.

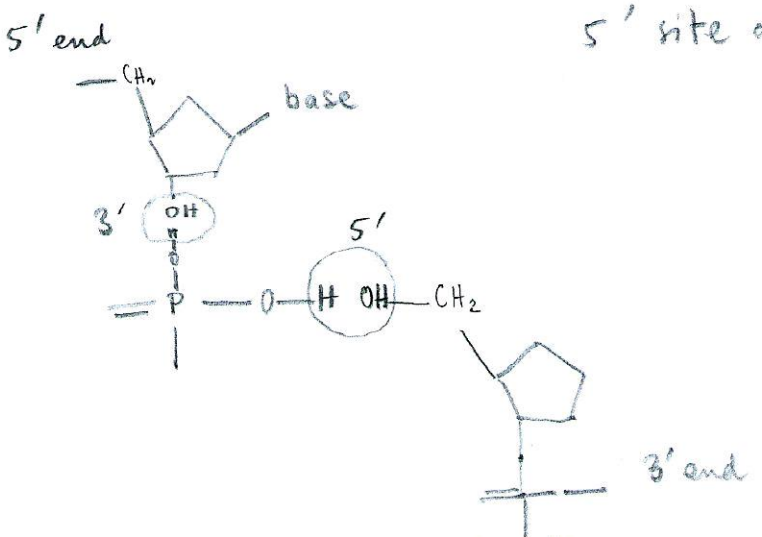
A : adenine T : thymine

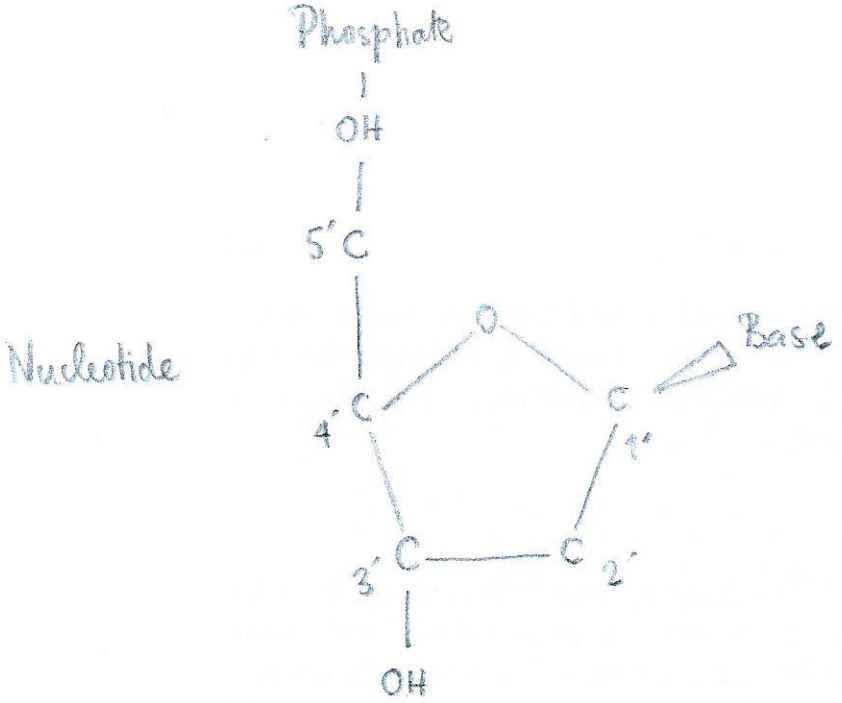
C : cytosine G : guanine

The base A on one strand forms a hydrogen bond to a base T on the other strand, and similarly a base C forms a bond to a base G.

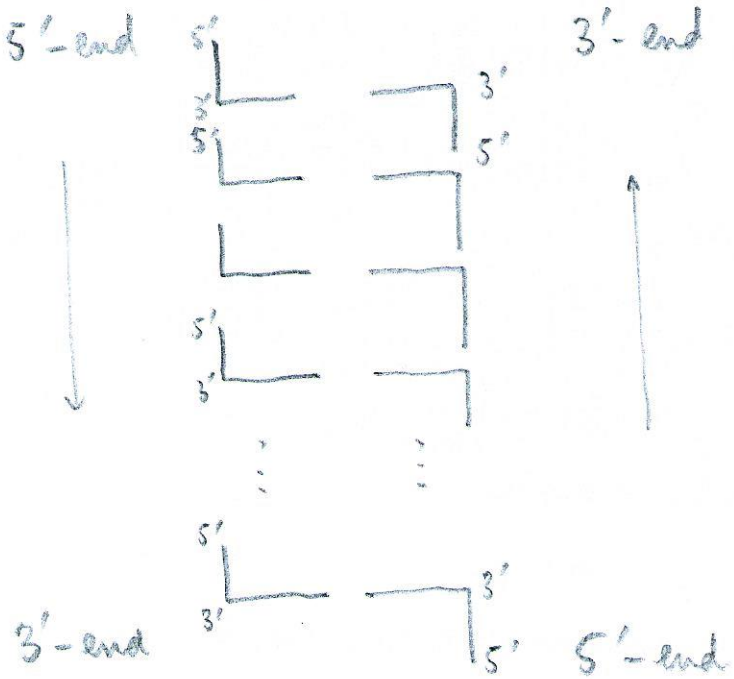
A pair of bases A — T or C — G bonded together are a <u>base pair</u> of the molecule.

The sugar molecule in the strand of DNA has a site called <u>3 prime</u>, 3', and a site called 5 prime, 5'. Each phosphate molecule is joined to the 3' site of one sugar and to the 5' site of another sugar.



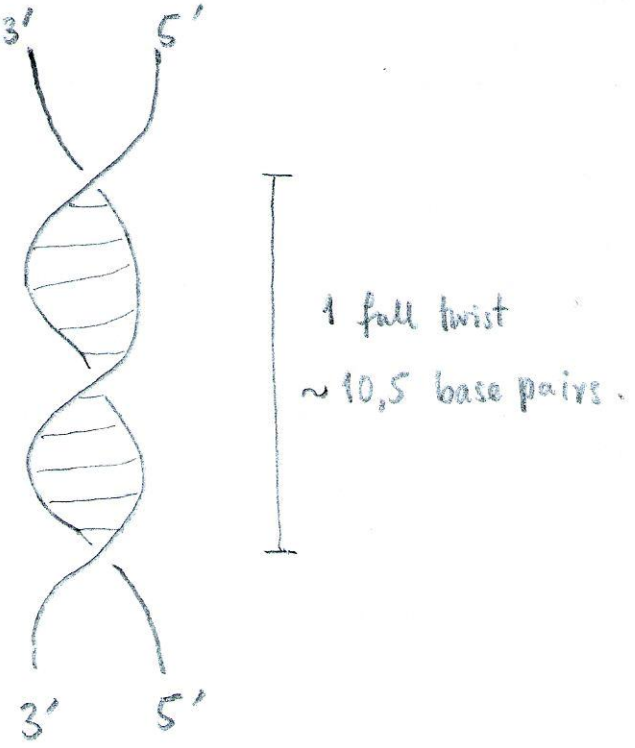


Schematically:



One end of a strand of DNA contains the 3' site of a sugar and is called the 3'-end, and the other end contains the 5' site of a different sugar, and is called the 5'-end.

The two strands of duplex DNA are oriented in opposite directions, so that one end of the DNA molecule consists of the 3'-end of one strand and the 5'-end of the other strand, while the other end consists of the 5'-end of the first strand and the 3'-end of the second strand.

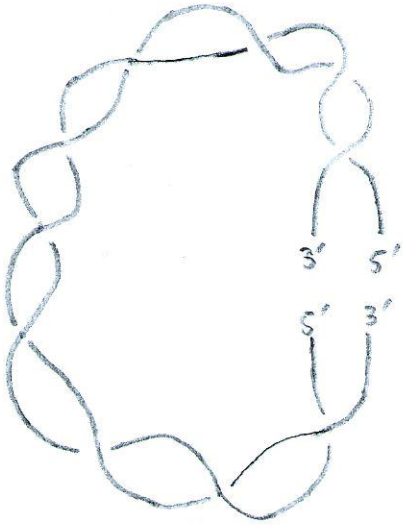


Since 3' can only join to 5', if the two ends of a duplex molecule are joined together to produce a circular DNA molecule, each strand is joined to

itself, rather than to the other strand. So a circular DNA molecule always has the form of a twisted annulus rather than the form of a twisted Möbius band.

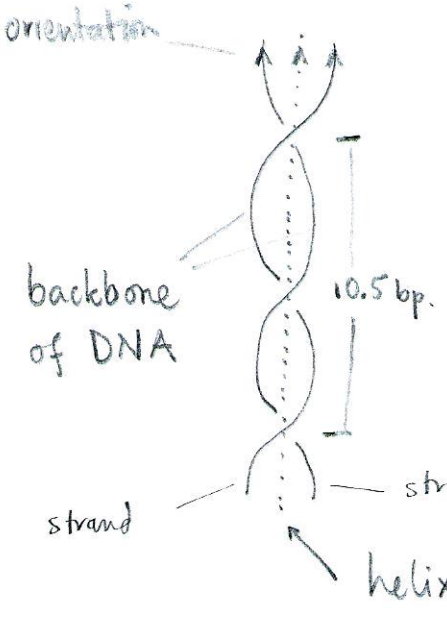
The circular DNA can be knotted, and two or more circular DNA molecules may form a link.

Apart from the linking of duplex molecules, we may also consider the linking of the two strands within a single DNA molecule.



This is a link with two components, called a torus link, and is always topologically chiral. Hence circular DNA molecules are always chiral.

Supercoiling.



A DNA molecule is in a relaxed state if the backbone twists one full turn for every 10,5 base pairs. In this state the helix axis is roughly planar.

If more or fewer twists are introduced between the strands, the molecule will become supercoiled, in order to minimize the torsional stress.

(Try it with a piece of rope!)

Supercoiling is important for many biological, chemical and physical properties of DNA. In particular it facilitates recombination.

The mathematics of supercoiling involves both the topology and the differential geometry of DNA.

Supercoiling.

Consider a circular DNA molecule. Choose the same orientation, for the helix axis and each strand. Let l be the linking number between the two strands. This is a topological invariant and will not change during deformations.

Because the strands are oriented in the same direction, for a right-handed helix all the crossings are positive, while for a left-handed helix all crossings are negative.

So the linking number of the circular molecule is

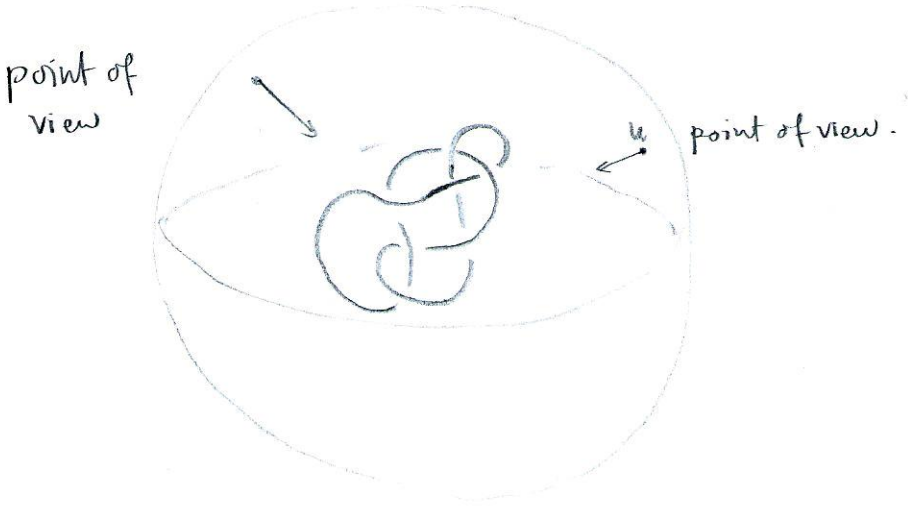
$$l = \frac{1}{2} \left(\text{number of crossings} \right).$$

If the molecule is relaxed, there are 2 crossings for every 10,5 bp, so the linking number of a relaxed circular molecule with N bp is the closest integer to $N/10,5$.

The linking number of DNA can be split into a part that comes from the wrapping of the helix axis, and another part which comes from the twisting of the backbone around the axis. This is expressed by the following two numbers:

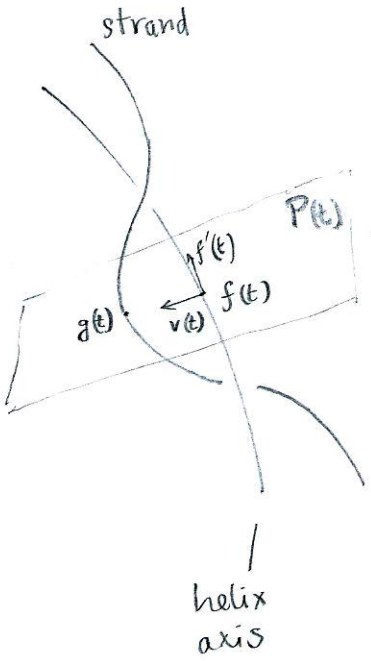
The <u>average writhe</u> of a DNA molecule is the average of all the values of the writhe of the helix axis over all possible projections.

This can be expressed as an integral over the surface of the sphere, of the writhe of the link from each different point of view:



$$W = \frac{1}{4\pi} \int_{S^2} w(u)\, dS$$

To define the twist we parametrize the helix axis and the two strands of the backbone.



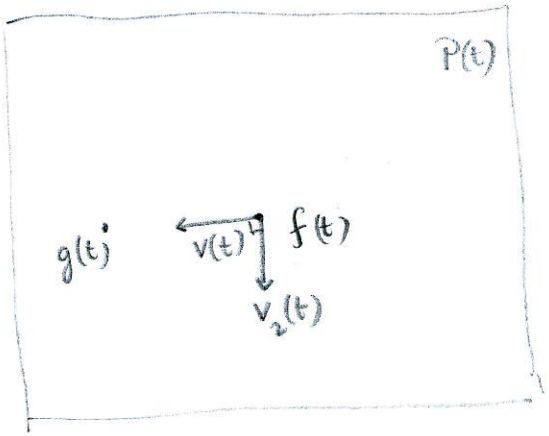
$f(t)$ parametrization of helix axis

$g(t)$ parametrization of one strand, so that $g(t)$ belongs to the plane $P(t)$ through $f(t)$, perpendicular to the helix axis.

$$v(t) = \frac{g(t) - f(t)}{|g(t) - f(t)|}$$ unit vector

Since $|v(t)|$ is constant, $v(t) \cdot v'(t) = 0$.

let $v'(t) = v_1(t) + v_2(t)$ with $v_1(t) \perp P(t)$
$v_2(t)$ in $P(t)$



$f'(t)$ up from the page.

Then $v(t) \cdot v_1(t) = 0$

and $v(t) \cdot v'(t) = 0$

hence $v(t) \cdot v_2(t) = 0$.

let $w(t) = \dfrac{f'(t) \times v(t)}{|f(t) \times v(t)|}$

be the unit vector forming

a right handed system $f'(t)$, $v(t)$, $w(t)$.

We define $\lambda(t) = v_2(t) \cdot w(t)$, and note that

$\lambda(t) = v'(t) \cdot w(t)$.

Then the twist is defined as

$$T = \frac{1}{2\pi} \int_C v'(t) \cdot w(t)\, dt .$$

The relation between the 3 quantities is given by:

Theorem (White, 1969).

$$\ell = T + W.$$

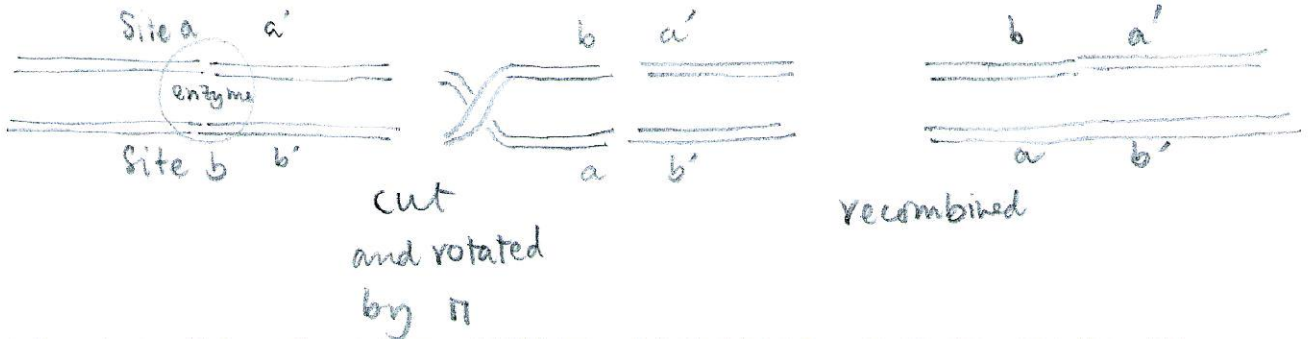
//.

Physically, the linking number is related to the compactness of a molecule, and can be measured experimentally by how fast a molecule moves through a gel (gel electrophoresis).

The action of enzymes on DNA.

One way enzymes act on DNA is through recombination: the creation of new genetic sequences out of pieces of existing genetic sequences.

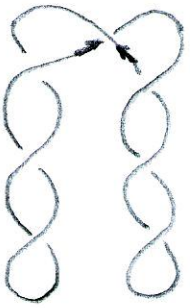
Site specific recombination: The sites are two short pieces of linear DNA, in a supercoiled DNA molecule. They are brought together by the enzyme, they are cut and recombined in a different way.



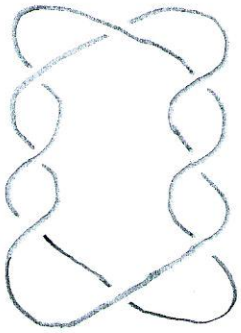
The overall effect depends on the supercoiling in the piece of DNA between the two sites.

Ex.

supercoiled unknot



link



unlink

One experiment which has been studied successfully by molecular biologists Wasserman, Duncan & Cozzarelli (1985) and by mathematicians Ernst & Sumners (1990) is the action of enzyme TN3 resolvase on unknotted, supercoiled, closed circular DNA.

In 95% of cases the enzyme causes 1 recombination

But in 5% it causes multiple recombination.

The experimental results showed that

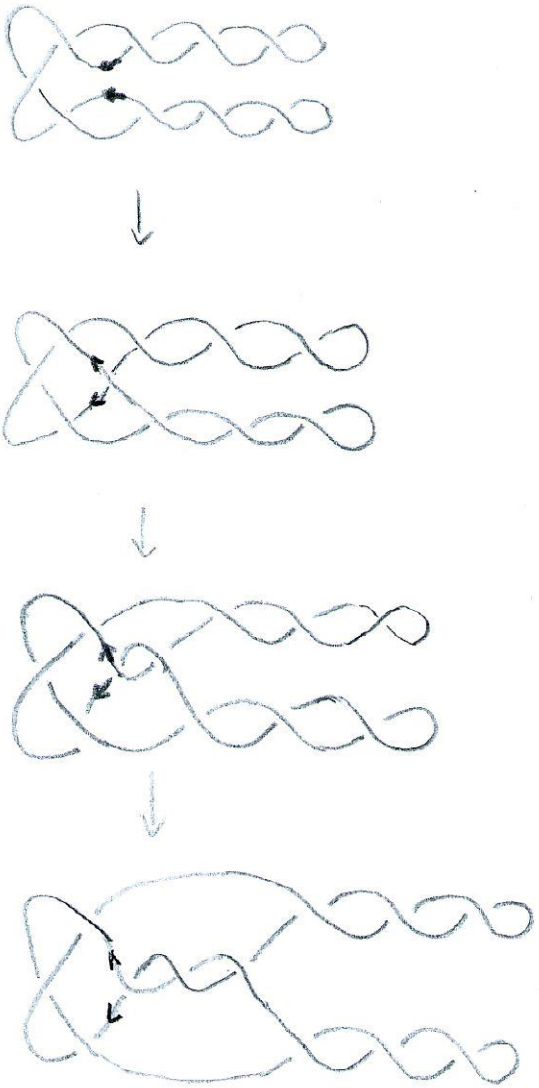
after 1 recombination — Hopf link

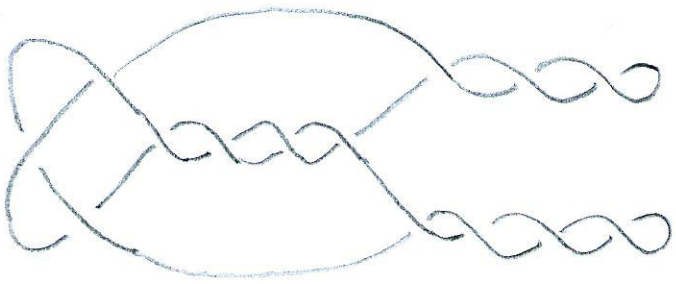
after 2 recombinations — figure 8

after 3 recombinations — Whitehead link

after 4 recombinations — 6_2 knot.

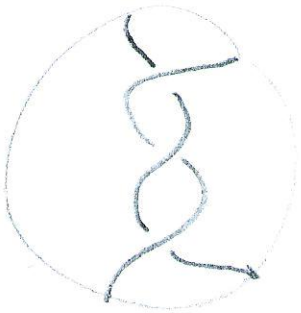
The biologists made the following / hypothetical model of how the enzyme acted on the substrate to create these products:





6_2

Ernst & Sumners modelled the recombination event using Conway's theory of <u>tangles</u> (a way to build up complicated knots and links from more simple ones) and proved that to build a sequence U, H, S, W one would have to start with a tangle



S

and successively add tangles

R

by recombination events. They also predicted that a 4th recombinatin would produce 6_2.