Lecture -1-

THE microbial genetics

Dr. Yasir Alabdali

Structure of nucleic acids

DNA

In bacteria, the genetic material is double-stranded DNA, although bacteriophages (viruses that infect bacteria) may have double-stranded or single-stranded DNA, or RNA. The components of DNA are 2´-deoxyribose (forming a backbone in which they are linked by phosphate residues), and four heterocyclic bases: two purines (adenine and guanine), and two pyrimidines (thymine and cytosine). The sugar residues are linked by phosphodiester bonds between the 5´ position of one deoxyribose and the 3´ position of the next, while one of the four bases is attached to the 1´ position of each deoxyribose. It is the sequence of these four bases that carries the genetic information.

The two strands are twisted around each other in the now familiar double helix, with the bases in the center and the sugar-phosphate backbone on the outside. The two strands are linked by hydrogen bonds between the bases. The only arrangement of these bases that is consistent with maintaining the helix in its correct conformation is when adenine is paired with thymine and guanine with cytosine

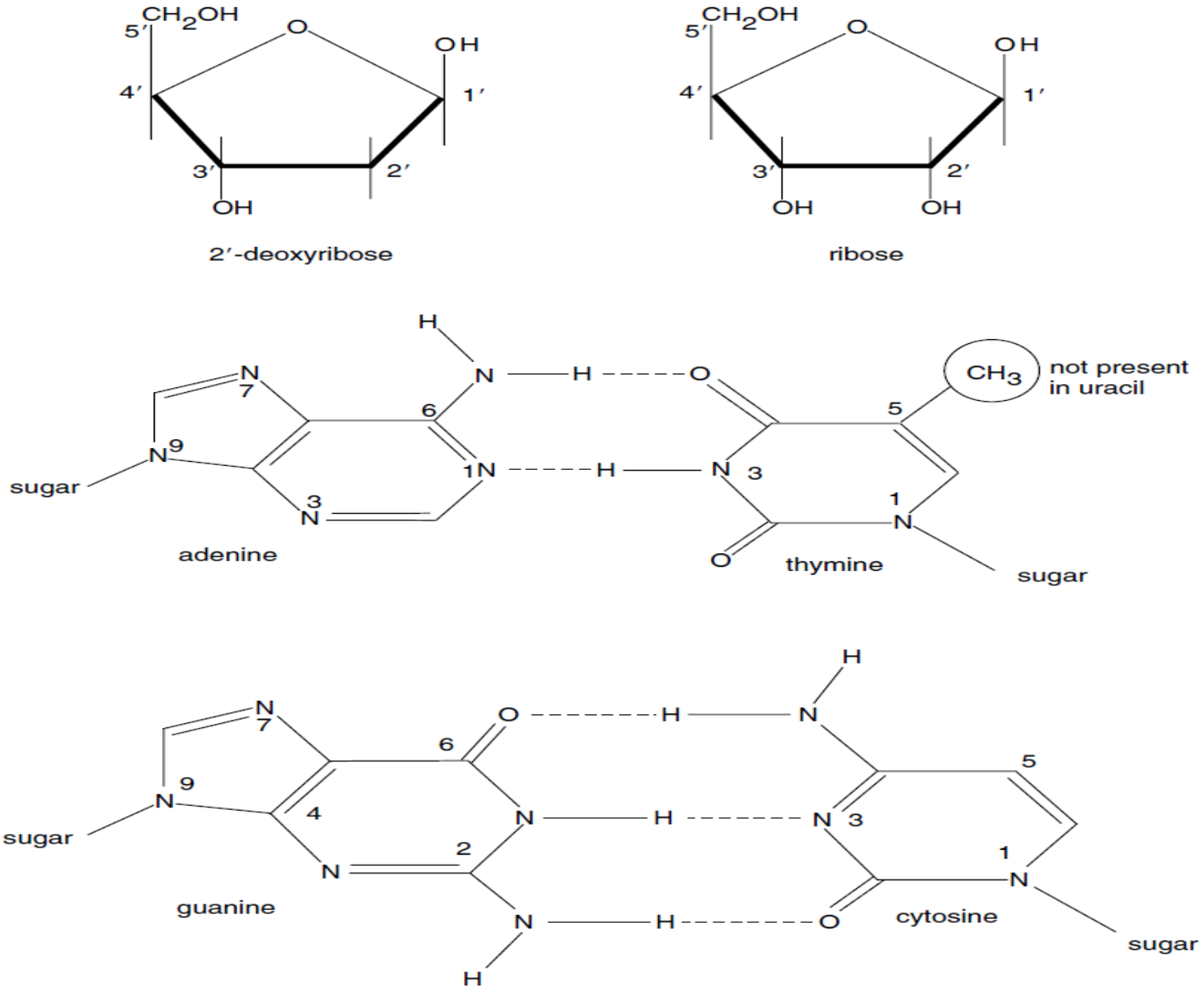
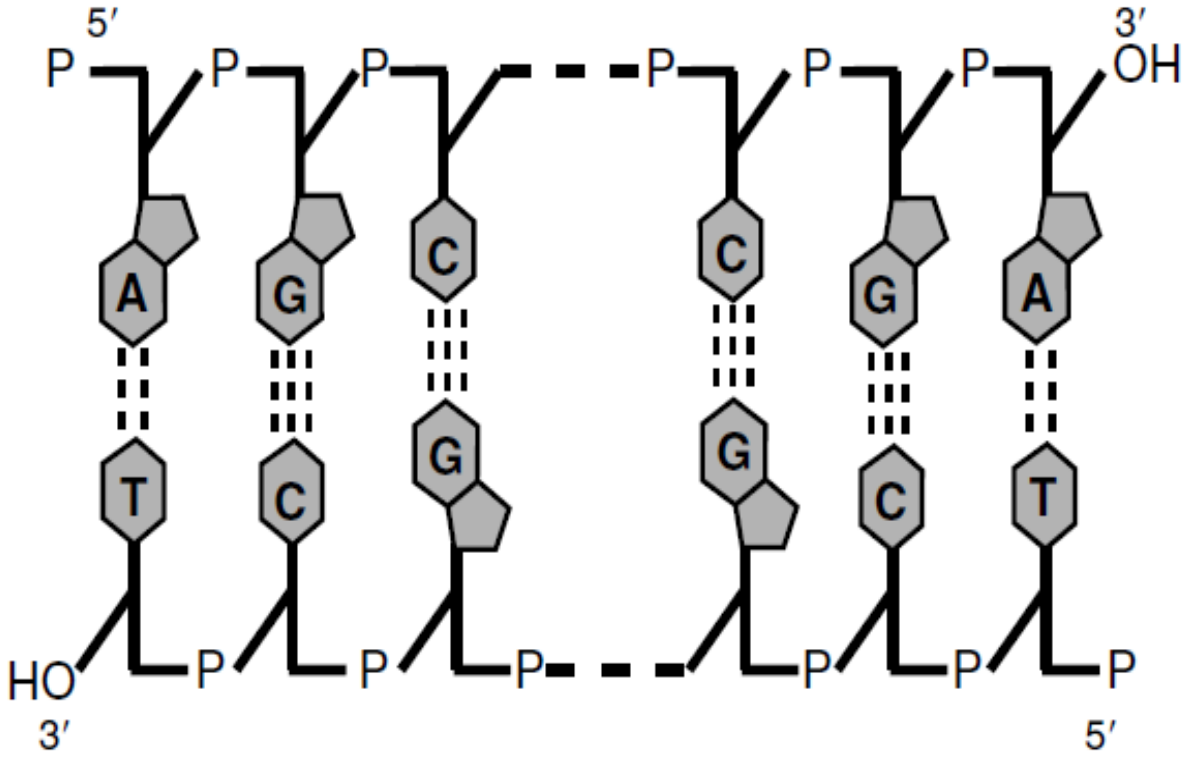
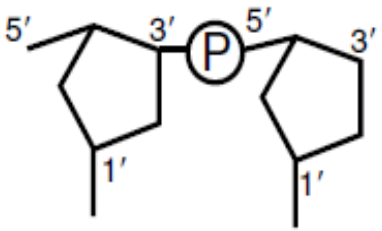


Figure 1.1 Structure of the basic elements of DNA and RNA. RNA contains ribose rather than deoxyribose, and uracil instead of thymine

One strand therefore consists of an image of the other; the two strands are said to be complementary. Note that the purines are larger than the pyrimidines, and that this arrangement involves one purine opposite a pyrimidine at each position, **so the distance separating the strands remains constant.**



The sugar-phosphate 'backbone'



is represented diagrammatically as

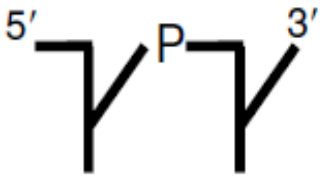


Figure 1.2 Diagrammatic structure of DNA

RNA

The structure of RNA differs from that of DNA in that it contains the sugar ribose instead of deoxyribose, and uracil instead of thymine . It is usually described as single stranded, but only because the complementary strand is not normally made. There is nothing inherent in the structure of RNA that prevents it forming a double-stranded structure: an RNA strand will pair with (hybridize to) a complementary RNA strand, or with a complementary strand of DNA. Even a single strand of RNA will fold back on itself to form double stranded regions. In particular, transfer RNA (tRNA), and ribosomal RNA (rRNA) both form complex patterns of base-paired regions.

Hydrophobic interactions

Although geneticists emphasize the importance of the hydrogen bonding between the two DNA strands, these are not the only forces influencing the structure of the DNA. **The bases themselves are hydrophobic**, **and will tend to form structures in which they are removed from the aqueous environment**. This is partially achieved by stacking the bases on top of one another. The double-stranded structure is stabilized by additional hydrophobic interactions between the bases on the two strands.

The hydrogen bonding not only holds the two strands together **but also allows the corresponding bases to approach sufficiently closely for the hydrophobic forces to operate. The hydrogen bonding of the bases is however of special importance because it gives rise to the specificity of the base pairing between the two chains.**

Although the bases are hydrophobic, and therefore very poorly soluble in water, **nucleic acids are quite soluble, due largely to the hydrophilic nature of the backbone, and especially the high concentration of negatively-charged phosphate groups**. This will also tend to favour a double helical structure, in which the hydrophobic bases are in the center, shielded from the water and the hydrophilic phosphate groups are exposed.

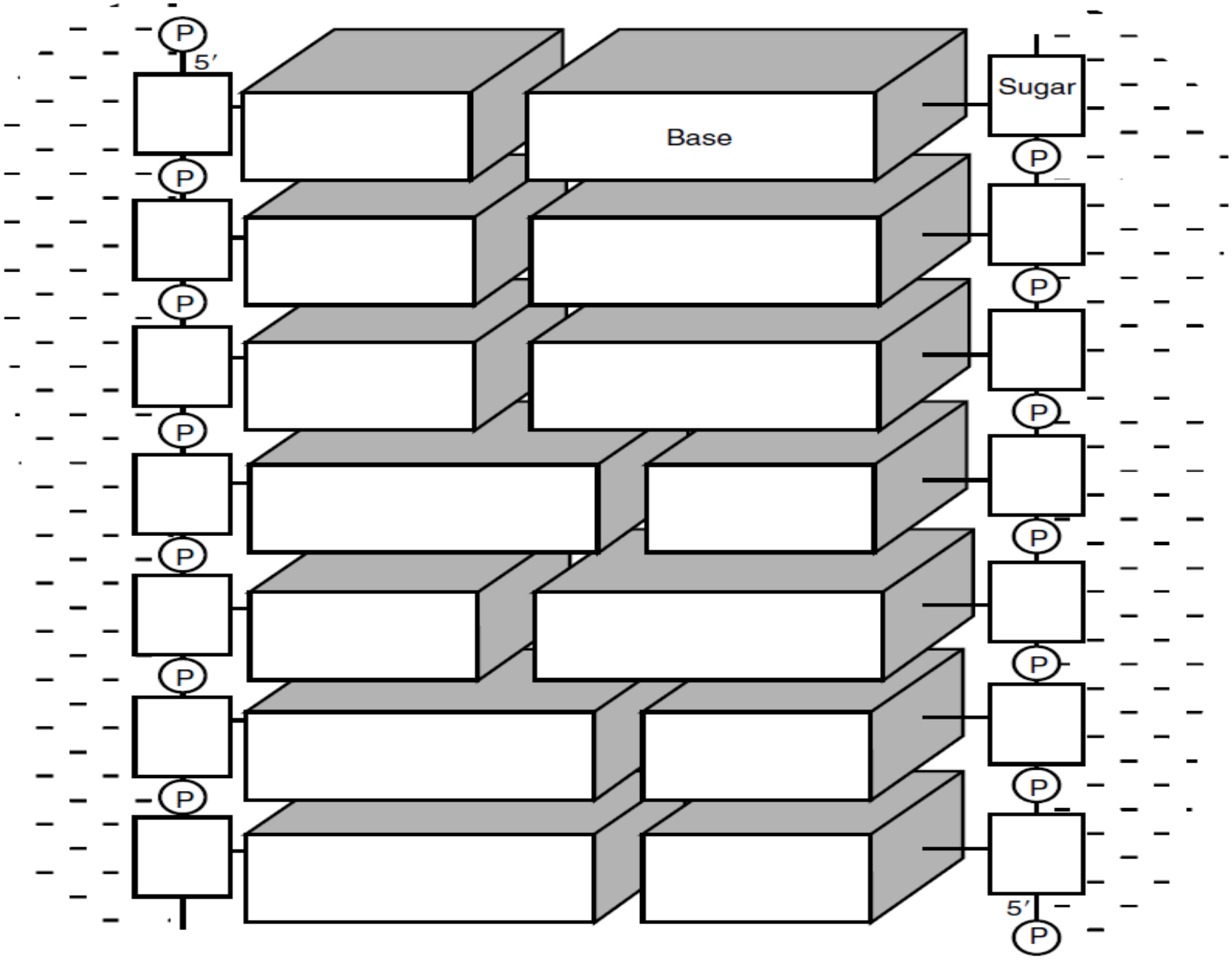


Figure 1.3 Hydrophobic interactions of bases in DNA. The hydrophobic bases stack in the centre of the helix, reducing their contact with water

Different forms of the double helix

A full consideration of DNA structure would be extremely complex, and **would have to take into account interactions with the surrounding water itself, as well as the influence of other solutes or solvents.** The structure of DNA can therefore vary to some extent according to the conditions. In vitro, two main forms are found. The Watson and Crick structure refers to the B form, which is a right-handed helix with 10 base pairs per turn. Under certain conditions, isolated DNA can adopt an alternative form known as the A form, which is also a right-handed helix, but more compact with about 11 base pairs per turn. Within the cell, DNA resembles the B form more closely, but has about 10.4 base pairs per turn

Certain DNA sequences, notably those containing alternating G and C residues, tend to form a left-handed helix, known as the Z form (since the sugar-phosphate backbone has a zig-zag structure rather than the regular curve shown in the B form). Although Z DNA was originally demonstrated using synthetic oligonucleotides, naturally occurring DNA within the cell can adopt a left-handed structure, at least over a short distance or temporarily. The switch from left- to right-handed can have important influences on the expression of genes in that region.

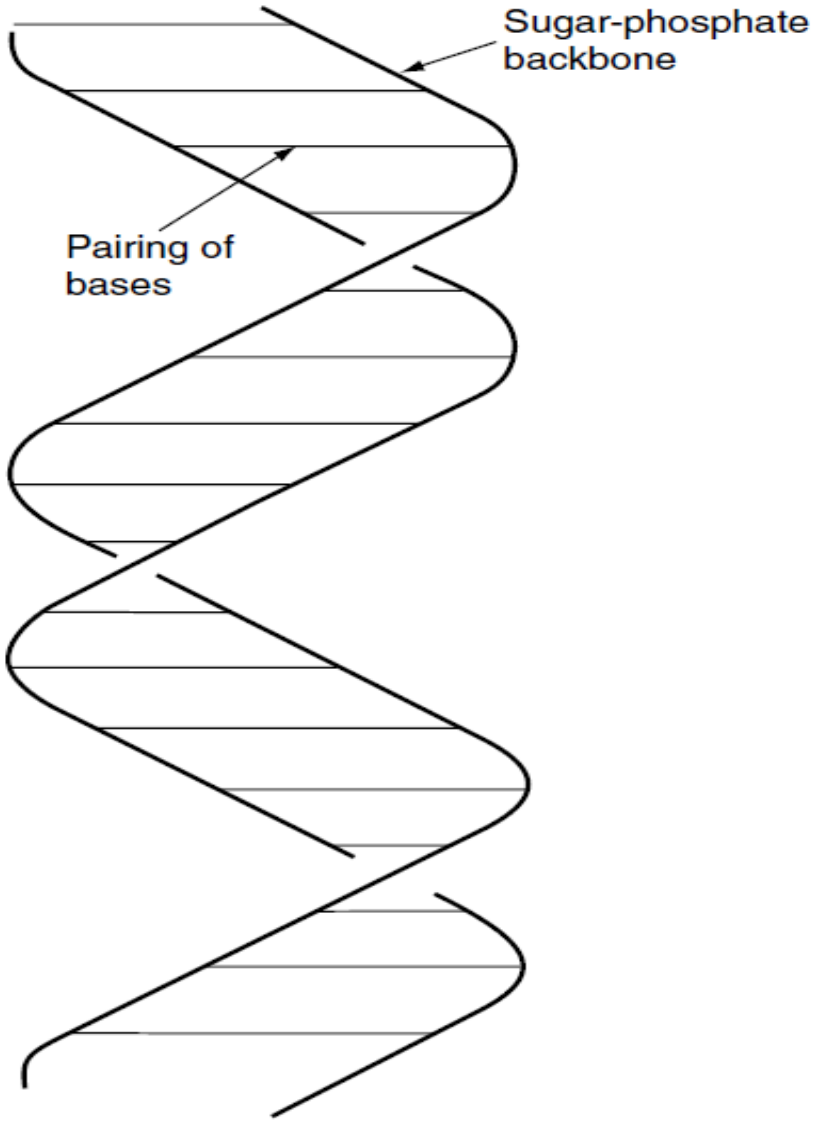
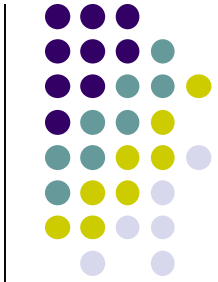


Figure 1.4 Diagrammatic structure of B-form DNA. The two anti-parallel sugar-phosphate chains form a right-handed helix with the bases in the centre, held together by hydrophobic interactions and hydrogen bonding

Supercoiling



Within the cell, the DNA helix is wound up into coils; this is known as supercoiling.

There are three parameters involved:

twist (T), linking number (L) , writhe (W).

The twist is the number of turns of the strip, while writhe (essentially a measure of the degree of supercoiling) can be considered as the number of times the strip crosses over itself in a defined direction. These two parameters vary according to the conformation:

in **(a),** there is one twist (T=1) but no supercoiling (W= 0), while in **(b)** there is no twist (T= 0) and the strip crosses itself once (W=1). The sum of these two parameters is used to define the linking number, i.e. L = T + W.

The linking number is therefore a measure of the overall twisting of the strip, or for DNA, the total number of times that the DNA strands wrap around one another.

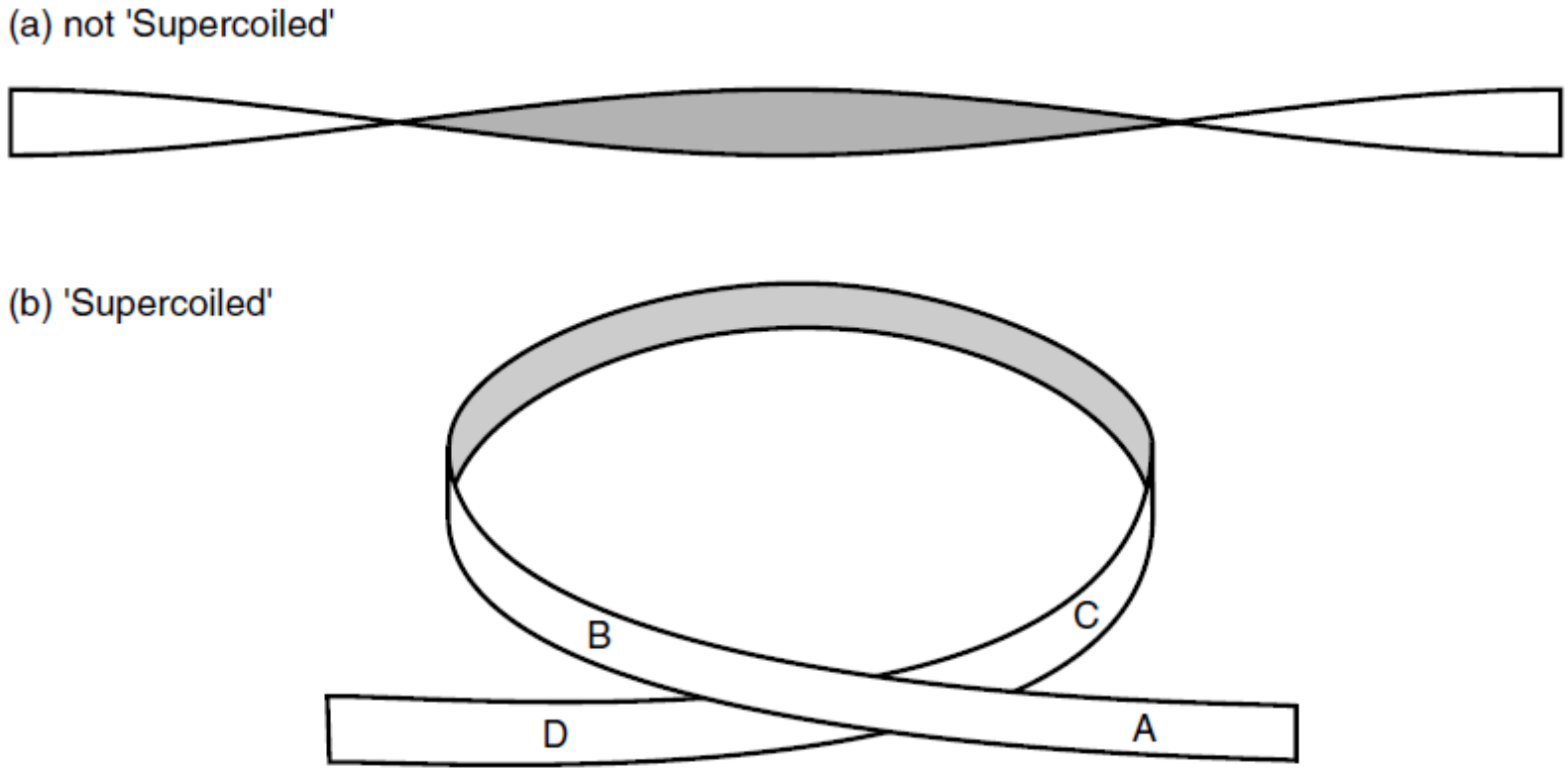


Figure 1.5 Interaction between twisting and supercoiling. (a) A ribbon with a single complete twist, without supercoiling. (b) The same ribbon, allowed to form a supercoil; the ribbon is now not twisted

If the ends of the strip are not free to rotate, then the linking number will remain **constant**. Most of the DNA molecules we will be considering are circular, and therefore do not contain ends that can rotate. Unless there is a break in the DNA, any change in the twist will be balanced by a change in supercoiling, and vice versa .

In **(a),** the strip (or DNA molecule) is not supercoiled (W = 0) but contains one complete twist (T=+1); the linking number (L) is +1.

In **(b),** the overall shape has been changed by rotating one end of the structure (i.e. introducing a degree of supercoiling). The strip crosses itself once, and by convention a crossover in this direction is assigned a negative value, so W = -1. At the same time the twist has changed; there are now two complete twists, so T = +2. Since L = T + W, we see that L remains the same (+1).

In (c), which is obtained by a further rotation of one end of the structure, there are two negative supercoils (W = - 2), and three complete twists (T =+3). Once more, L stays the same.

Changing the supercoiling alters the degree of twisting correspondingly, so that the linking number remains constant. The three structures shown in Figure 1.6 are interchangeable by rotating one end, without opening the circle. With an intact circle, the twist and the writhe can be changed jointly but not separately; any change in supercoiling will involve a compensating change in the twist (and vice versa) so that the **linking number remains constant**. It is only possible to alter the linking number in circular DNA by breaking and rejoining DNA strands, for example through the action of topoisomerases.

Bacterial DNA is normally negatively supercoiled. Another way of putting it is to say that the DNA is underwound so that if the DNA was not supercoiled, the degree of twisting of the helix would be less than that seen in relaxed linear DNA.

If the DNA is nicked (i.e. one strand is broken, leaving it free to rotate) it relaxes into an open circular, non-supercoiled form. Chromosomal DNA is usually broken into linear fragments during lysis of the cell, but bacterial plasmids are usually small enough to be isolated intact in a supercoiled form.

The compact supercoiled structure of the DNA is also significant in that the chromosome, in its expanded state, would be a thousand times longer (about 1 mm) than the bacterial cell itself. To put it another way, a bacterial operon of four genes, in its non-supercoiled B-form, would stretch from one end of the cell to the other.

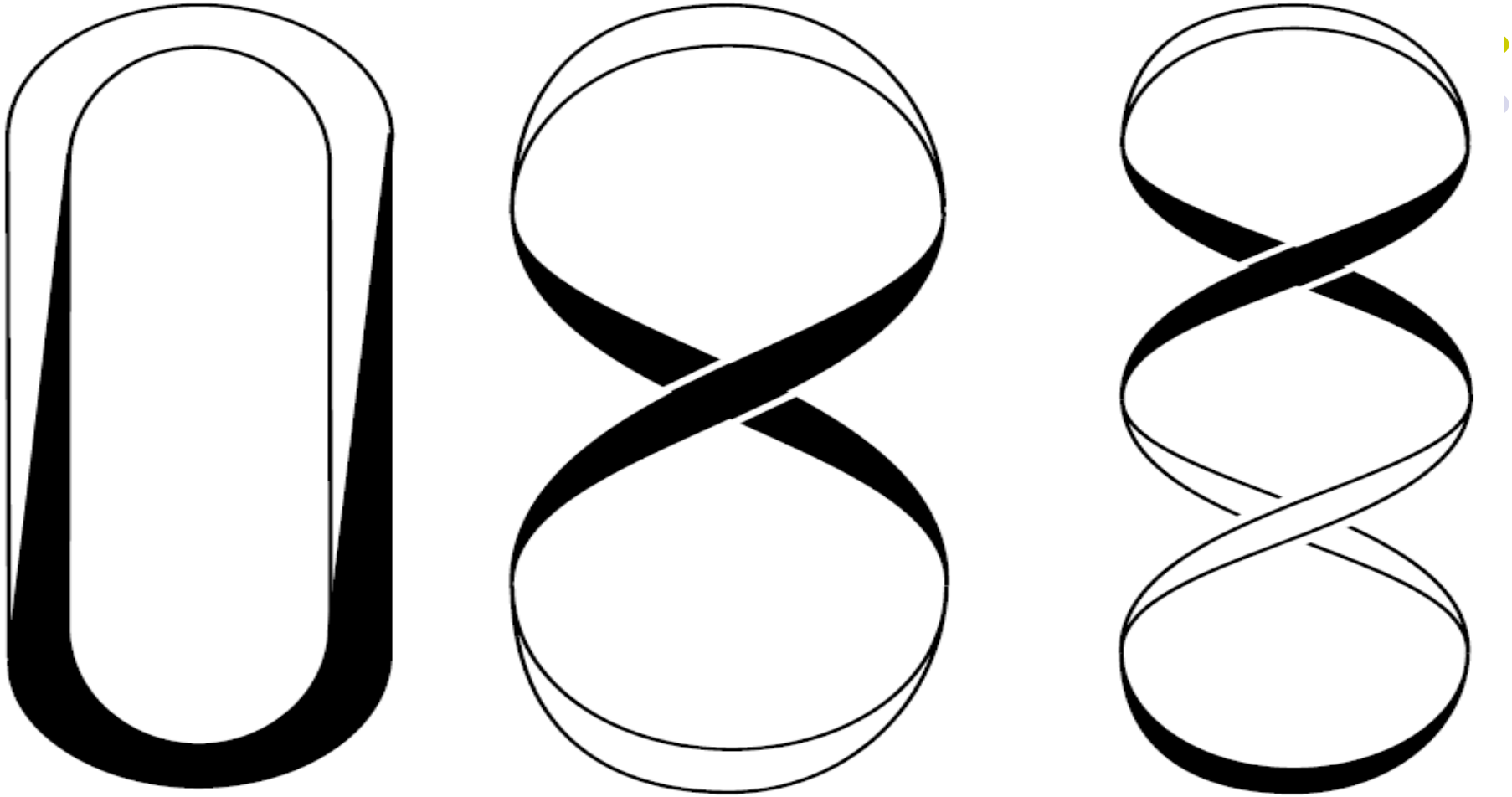


Figure 1.6 Supercoiling of a circular molecule. (a) The 'molecule' has one twist and no supercoils. Rotating one end of the molecule (b and c) introduces negative supercoils and increases the amount of twisting. Since the linking number (L) remains the same, the three forms are interchangeable without breaking the circle (see the text for further explanation)

Supercoiling is only the start of story, as the bacterial chromosome consists of a large number of supercoiled loops arranged on a core to produce a highly compact and organized structure known as the nucleoid. Supercoiling (and other structural features) of the DNA are also important in the regulation of gene expression

Action of topoisomerases

Supercoiling of bacterial DNA is not achieved by physically twisting the circular molecule. Instead the cell uses enzymes known as DNA topoisomerases to introduce (or remove) supercoils from DNA, by controlled breaking and rejoining of DNA strands.

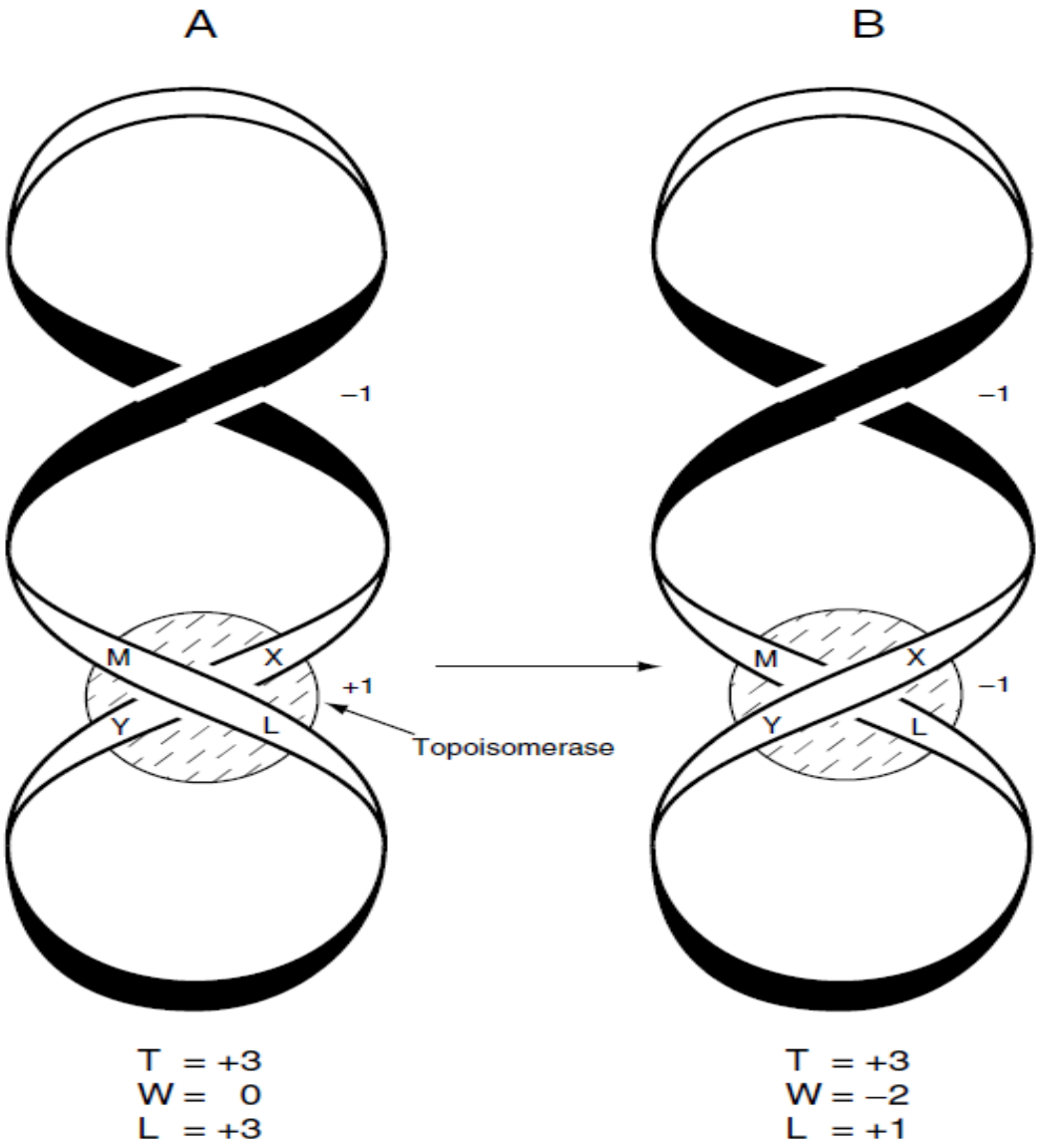


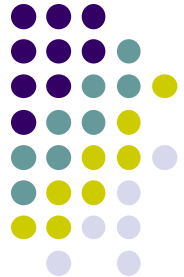
Figure 1.7 Action of Type II topoisomerase. Structure A is not supercoiled: the two crossing points are of opposite sign and cancel one another. The topoisomerase makes a double strand break between L and M, passes the X–Y region through the gap and re-seals the break between L and M. This changes the sign of W at that point, so structure B is now negatively supercoiled

DNA topoisomerases can be considered in two classes:

–Type I topoisomerases act on a segment of DNA by breaking one of the strands and passing the other strand through the gap, followed by resealing the nick. Since this increases the number of times the two strands cross one another, the linking number is increased by 1 which results in an increase in either T or W. **The E. coli topoisomerase I acts only on negatively supercoiled DNA; the increase in the value of W means that the degree of negative supercoiling is reduced (the DNA becomes relaxed).**

–**Type II topoisomerases** break both strands and pass another duplex region through the gap. In Figure 1.7, the structure A looks at first glance to be supercoiled, but closer inspection shows that the two crossover points are of opposite sign, and therefore cancel each other out. Structure A is a nonsupercoiled circle that is drawn in this way to show the action of the topoisomerase.

If both strands of the helix are broken between points L and M and the lower strands (X–Y) are moved through the gap followed by re-sealing the strands between L and M, structure B is formed. As a consequence of this reaction, the sign of W is changed at that point (changing W from +1 to -1), which has the overall effect of a change of 2 in the value of W and hence in the value of the linking number L. An important example of this type of enzyme is **DNA gyrase, which is able to introduce negative supercoils into newly replicated DNA.**



Denaturation and hybridization

Since the two strands of DNA are only linked by non-covalent forces,

they can easily be separated in the laboratory, for example **by increased temperature or high pH**. Separation of the two DNA strands, **denaturation**, **is readily reversible**

Reducing the temperature, or the pH, will allow hydrogen bonds between complementary DNA sequences to re-form; this is referred to as **re-annealing**

If DNA molecules from different sources are denatured, mixed and allowed to re-anneal, it is possible to form hydrogen bonds between similar DNA sequences **(hybridization).** This forms the basis of the use of DNA probes to detect specific DNA sequences.

The specificity of the reaction can be adjusted by altering the temperature and/or the ionic strength. Higher temperature, or lower ionic strength, gives greater stringency of hybridization. High stringency hybridization is used to detect closely-related sequences, or to distinguish between sequences with only small differences, while low stringency conditions are used to detect sequences that are only remotely related to the probe. This technique forms an important part of modern molecular biology

Temporary separation of localized regions of the two DNA strands also occurs as an essential part of the processes of replication and transcription.

Note that there are three hydrogen bonds linking guanine and cytosine while the adenine– thymine pairing has only two hydrogen bonds. The two DNA strands are therefore more strongly attached in those regions with a high G : C content. Because of this, such regions are more resistant to denaturation and conversely re-anneal more readily. The influence of base composition on the ease of separation of two nucleic acid strands may play an important role in the control of processes such as the initiation of RNA synthesis where an A-T rich region may facilitate the initial separation of the DNA strands.

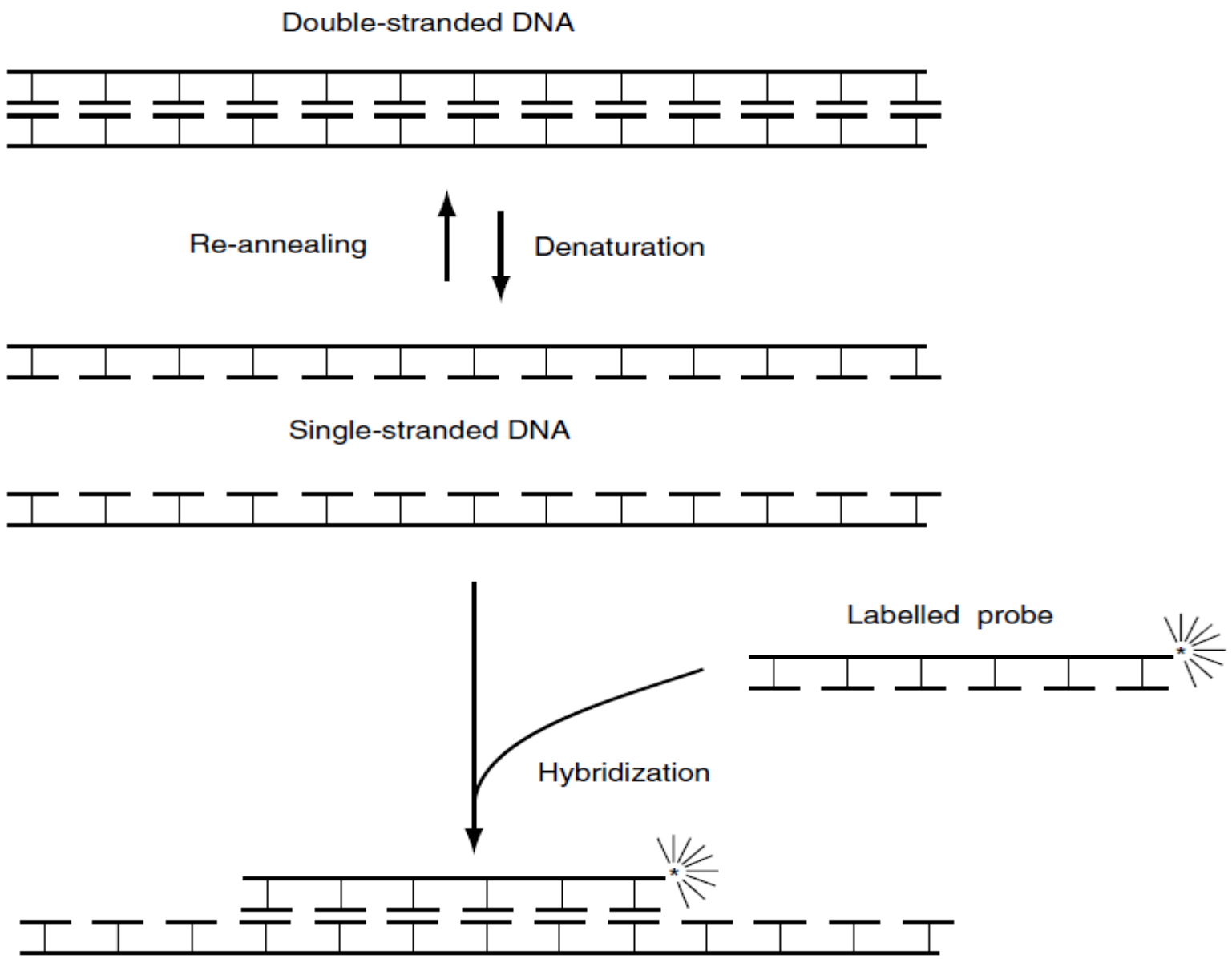


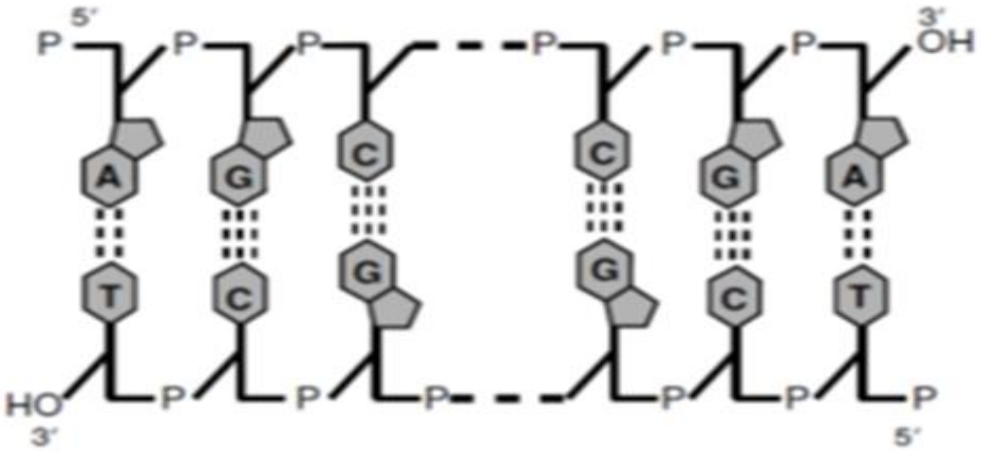
Figure 1.8 Denaturation and hybridization of DNA

Orientation of nucleic acid strands

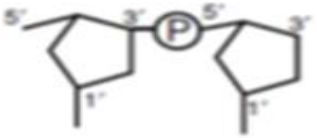
Feature of the helix is that each strand can be said to have a **direction**, **based on the orientation of the linkages in the sugar-phosphate backbone.** Each phosphate group joins the **5´ position** of one sugar residue to the **3´ position** of the next deoxyribose.

the upper strand has a free 5´ group at the left-hand end and a 3´ OH group at the right-hand end. It is therefore said to run (from left to right) in the 5´ to 3´ direction. Conversely, for the lower strand, the 5´ to 3´ direction runs from right to left.

By convention, if a single DNA (or RNA) strand is shown, it reads in the 5´ to 3´ direction from left to right (unless otherwise stated). If both strands are shown, the upper strand reads (left to right) from the 5´ to 3´end. All nucleic acids are synthesized in the 5´ to 3´ direction. That is, the new strand is elongated by the successive addition of nucleotides to the free 3´ OH group of the preceding nucleotide. The phosphate to make the link is provided by the substrate which is the nucleoside 5´-triphosphate (ATP, GTP, CTP, UTP for **RNA**; dATP, dGTP, dCTP, dTTP for **DNA**).



The sugar-phosphate 'backbone'



is represented diagrammatically as

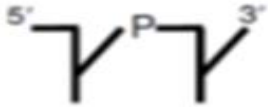


Figure 1.2 Diagrammatic structure of DNA

Genome

Genome is the complete set of organism's hereditary information. It is encoded either in DNA or in RNA for many types of viruses. The genome includes both the genes and the non-coding sequences of the DNA/RNA.

Genome size: is the total number of DNA base pairs in one copy of haploid genome. The genome size is positively correlated with the morphological complexity among prokaryotes and lower eukaryotes.

Organism	**Genome size (base pairs)**
HIV	9,749
Phage λ	48,502
Escherichia coli	4,600,000
Haemophilus influenzae	1,830,000
Saccharomyces cerevisiae	12,100,000
Homo sapiens	3,200,000,000

References

Molecular Genetics of Bacteria. By Jeremy W. Dale and Simon F. Park2004 , 4 edition, John Wiley & Sons Ltd.

p.: 1-12