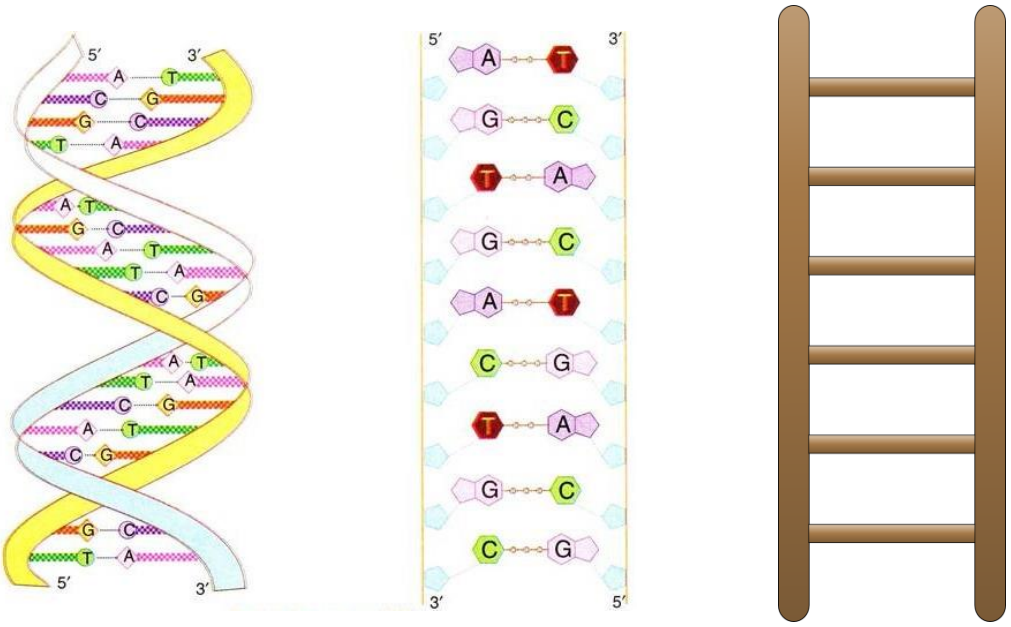
Genetics & molecular biology

◉ Sheet ○ Slide

Number:		2
Done by:	-	- Aya Dabbah - Fatima alaa - Raneen Alshareef
Corrected by:	-	- ذكرى البزور -Nora Haj-ali
Doctor:	-	Suhir Ereqat Rula abd alghani Emad matouq

The backbone of DNA is a deoxyribose sugar plus a phosphate group.

The rungs of the ladder are the nitrogenous bases.

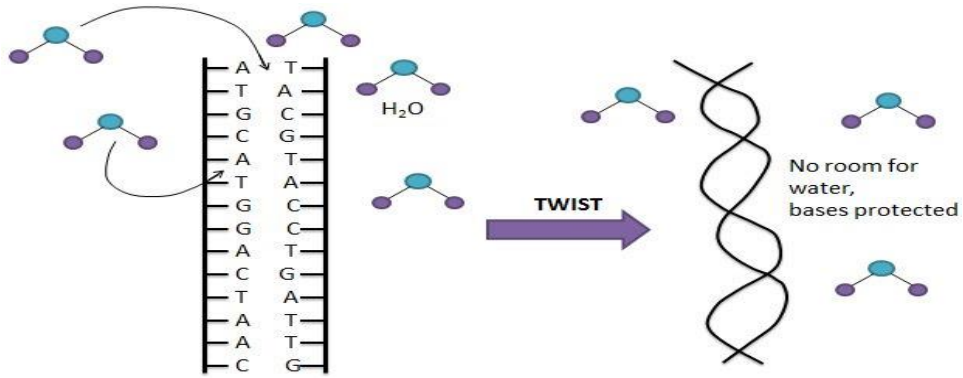


Stability of the DNA double helix structure (it's very stable) arises mainly from:

1**) Hydrogen bonds**: between A and T (two hydrogen bonds), and C and G (three hydrogen bonds). A hydrogen bond on its own is weak, but collectively within a DNA strand that contains billions of hydrogen bonds it becomes very strong and firm.

2) **Hydrophilic-hydrophobic interactions**: (hydro: water, philic: loving, phobic: scared)

In the structure of a nucleotide, one part loves water, and the other runs away from it. Nitrogen bases (A, T, C, G) are hydrophobic, while the sugar (ribose) and the phosphate group are hydrophilic. Therefore, it was found that in the model of DNA, ribose (which contains a hydroxyl group) and phosphate face outwards (since the DNA lies inside an aqueous medium) while nitrogen bases face inwards away from water. This interaction contributes largely to the stability of DNA (if a nitrogen base is broken it would immediately turn away from water and easily bind to another base inside) the double helix of the DNA turns with an angle of 36°, which is important for compaction.

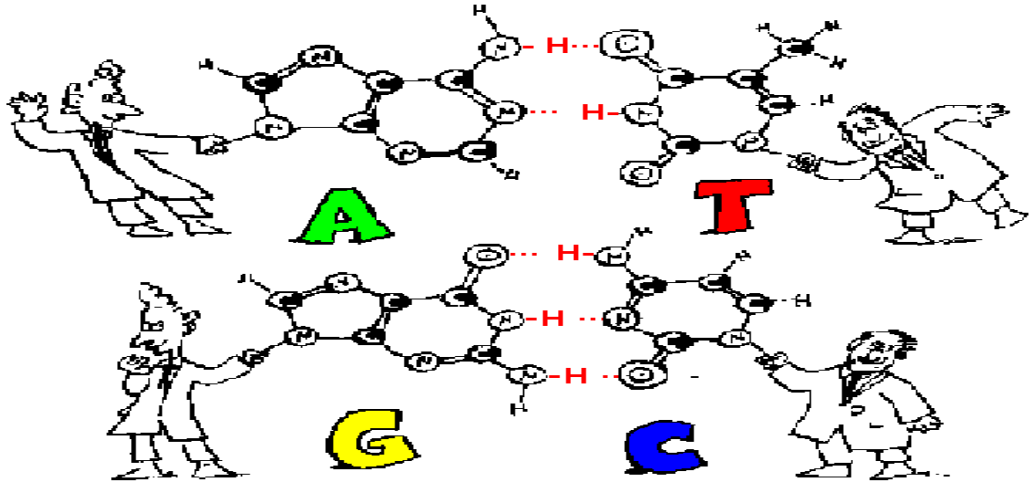


=> The distance between one pair of nitrogenous bases and another is 3.4 nanometer (3.4 Angstrom). (visualize the pairs as steps of a ladder, only a very distorted ladder)

A complete turn inside a double helix is 34 Angstrom long, and since the distance between one pair and another equals 3.4 Angstrom, the number of pairs inside a complete turn is 34/3.4 = 10 nitrogenous base pairs.

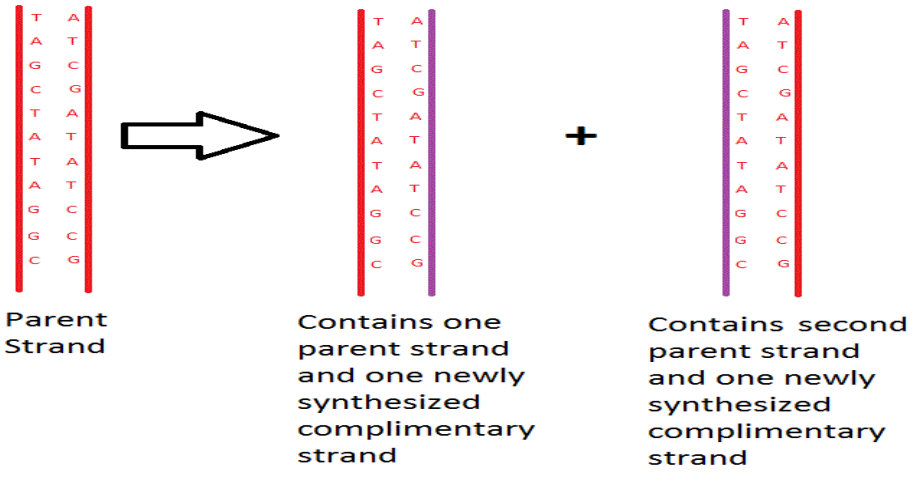
The most important principle in Watson and Crick's model of DNA; is complimentary base pairing, which means that in each strand of DNA, Adenine pairs with Thymine, while Guanine pairs with Cytosine.

Watson and Crick came to this conclusion by finding out that the number of A's equals the number of the T's, and the number of the C's equals the number of the G's, which could only mean that the A's pair with the T's , and the G's with the C's



DNA (replicates into) another DNA molecule (which transcribes into) RNA (part of which translates into) Protein, in this process:

1) Replication follows complimentary base pairing, since A (for example) doesn't produce another copy of itself, but rather replicates into T, and C replicates into G and vice versa. So the principle of complimentary base pairing helped uncover the mechanism of replication.



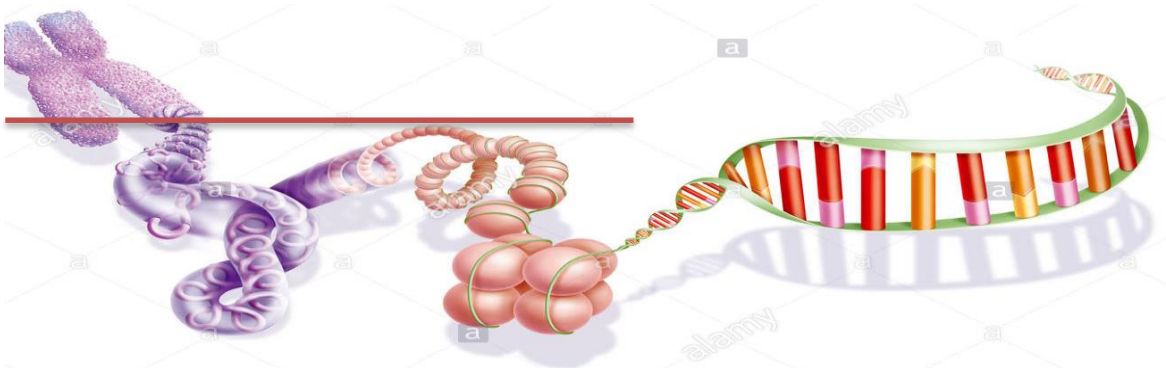
2) DNA polymerase, in transcription of DNA to RNA, also follows this principle, since it transcribes the nitrogenous bases complimentarily (C into G and T into A)

3) Also in translation, tRNA puts the complimenting amino acids into the nucleotides on the RNA strand.

In conclusion, **complimentary base pairing model is basically the fundament of Molecular biology**

Compaction process is not at all random and is extremely regulated (any problem with compaction in any part could lead to disaster). Many genetic problems are not a result of mutation in the DNA, but rather, are a result of packaging problems, even when the DNA is perfect >>

*Let's say a segment of DNA contains a number of genes, if this DNA was 100% extremely compact, RNA polymerase or DNA polymerase (or any kind of machinery) wouldn't be able to reach the genes, rendering them inactive (off). Another DNA segment is more loose than the previous one, and as a result, the genes are on, because the machinery could easily reach the genes.



Two specialized characteristics of DNA:

- The largest amount of DNA is in the nucleus.

- The arrangement of DNA is accessible.

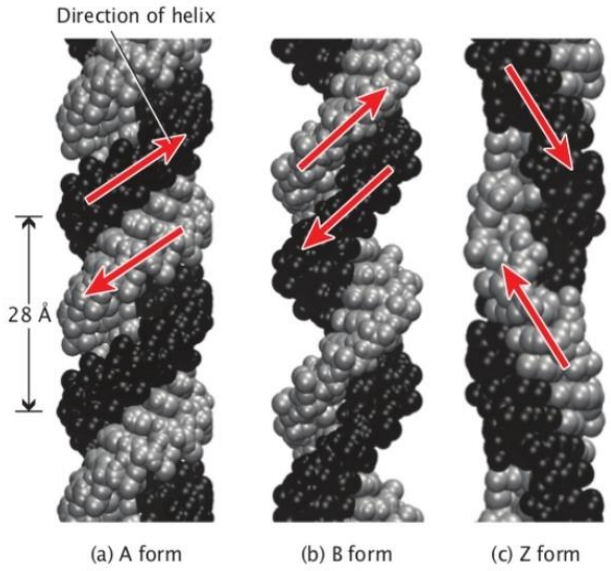
The twisting of DNA leads to major and minor grooves, we have three forms of DNA: A, B, and Z.

The most common one is B-form.

Our DNA is B-form, right handed.

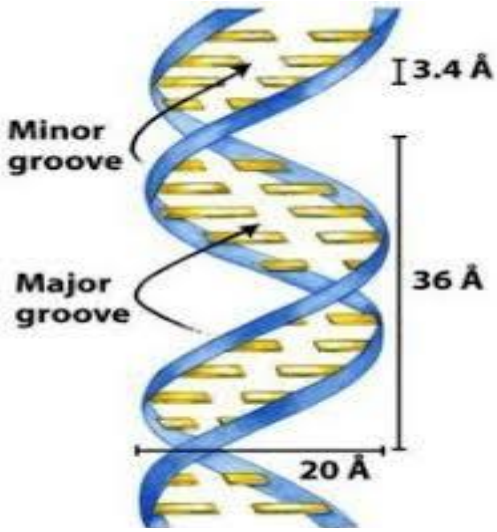
Z-form is found more in solutions.

The difference between B and Z forms is that Z is more stretched however A is more compressed than B form.



- What's the importance of major grooves: they're proteins that are important for gene expression (translation and transcription) which create bonds or connect into these groove areas. Proteins connect more into the major grooves than the minor ones. Minor grooves of DNA play an important role in the process of DNA turning, since it helps with adhesion, especially in sites rich with A-T bonds.

Grooves are not the only factor in binding proteins, but are a very important part of the process, and in case something wrong happens with them, proteins bind inefficiently, which makes living impossible, because no physiological process would be possible



=> There are 3 are parameters or characteristics of any DNA molecule double helix, which differ between the forms of DNA (A, B, Z).

Turn: is a result of twisting, it's the distance between two major grooves.

Each turn has around 10.4-10.5 base pairs (bp)

Rise: is the distance between two base pairs.

Tilt: is the angle created by the twisting of DNA.

DNA is a helix structure, there are 3 of DNA depend on 2 factors: helix sense and base pairs/turn.

Helix sense:

اصبع الابهام كمحور وهمي عند لف الاصابع الاخرى اذا كانت مع عقارب الساعة – left handed، عكس عقارب الساعة – right handed

Base pairs :كم عدد النيوكلوتيد باللفة الواحدة

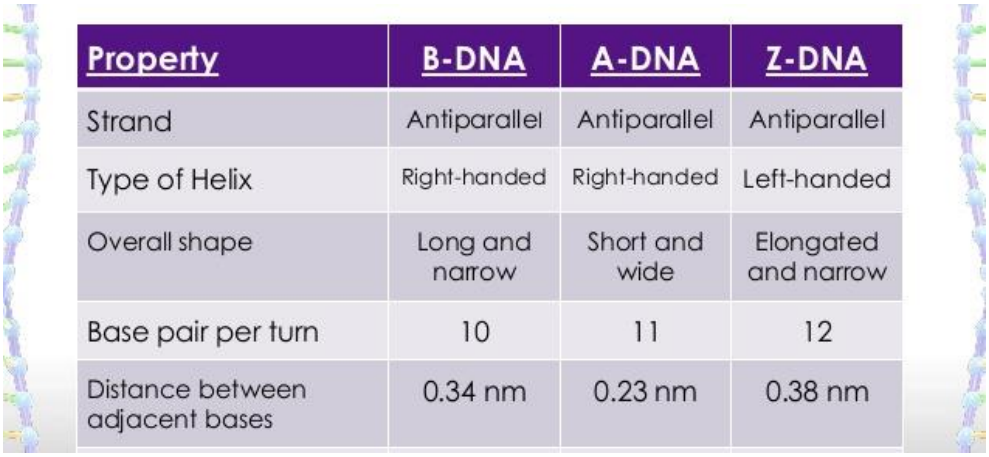
يوجد 3 وضعيات حسب العالمين

A-DNA: (right handed + 11 bp/turn)

B-DNA: (right handed + 10 bp/turn)

Z-DNA: (Left handed + 12bp/turn)

A>B>Z (بالنسبة للضغط)



Property	B-DNA	A-DNA	Z-DNA
Strand	Antiparallel	Antiparallel	Antiparallel
Type of Helix	Right-handed	Right-handed	Left-handed
Overall shape	Long and narrow	Short and wide	Elongated and narrow
Base pair per turn	10	11	12
Distance between adjacent bases	0.34 nm	0.23 nm	0.38 nm

Separation of the DNA is when the two strands of the double helix separate from each other which is necessary for the process of replication as well as transcription. This is one of the most important implications of Watson and Crick's model. Separation of the double helix in nature requires a number of enzymes, the most important of which is **Helicase.**

In a laboratory, separation of the DNA is possible, and the process is called **Denaturation** or **Melting**. Denaturation of the DNA is completely different from the denaturation of proteins, because the denaturation of DNA is reversible unlike proteins where it's irreversible (like when you boil an egg, proteins coagulate and you can't bring them back to how they were before). Therefore, if you provided the separated DNA strands the right conditions, they could easily reform into a double helix.

Melting temperature is the temperature required for 50% separation of the DNA (not all of the DNA). It depends mainly on the (C-G) content: if a DNA strand is rich in C-G pairs, we need a higher temperature, on the other hand, if the A-T content is higher, less temperature is required. Therefore, there is not a fixed melting temperature for all DNA, but it rather depends on composition.

The TM relates to two things:

1. G-C content has more hydrogen bonds meaning more heat is needed.

2. pH.

**Comparison should be only done if the fragments of DNA sequence have the same length.

This process could be done in the lab by simple heating (breaking hydrogen bonds only) (breaking 100% of the hydrogen bonds of course requires a higher temperature than melting temperature), breaking the phosphodiester bond can't be done physically and rather requires enzymes, DMSO (dimethyl sulfoxide) and formamide and other substances could be used to weaken hydrogen bonds and assist in denaturation. High pH. and low salt concentration also help with denaturation.

Renaturation: bringing the two separate strands of DNA back into a double helix, one application of denaturation and renaturation is classification of living organisms, if we took one separated strand of DNA from a creature, and another one from another creature, brought them together and found that they wrapped around each other, it means these two creatures have very similar DNA and therefore should be classified together.

Renaturation rate increases with:

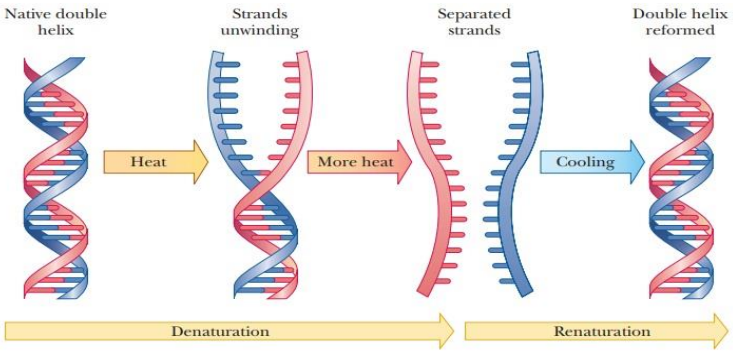
1) Removal of heat (25 degrees beneath melting temperature).

2) High concentration of DNA, because in that case the separate strands would be closer together and much more inclined to wrap back around each other faster.

3) Time, more time means more renaturation.

4) High salt (ionic) content, since low salt content triggers denaturation.

A strand can wrap around itself by forming hydrogen bonds between the nitrogen bases within it, which obstructs renaturation. It can be prevented if the temperature was higher.



■ FIGURE 9.19 **Helix unwinding in DNA denaturation.** The double helix unwinds when DNA is denatured, with eventual separation of the strands. The double helix is re-formed on renaturation with slow cooling and annealing.

One application of denaturation and renaturation is polymerase chain reaction (PCR), which is the greatest technique of the 20th century, it basically means that you can take any DNA segment of any living creature (human, fungi, bacteria…) and in a test tube replicate it into hundreds of millions of copies in less than an hour and a half.

Nucleic acid hybridization is another application (to hybridize a nucleic acid) which basically means you can check for the presence of a certain RNA inside the cell, by designing a complimentary RNA with a couple of signals, so that if it was injected inside the cell and wrapped perfectly around another RNA it sends out a signal, meaning that this certain RNA that we're looking for is in fact active inside the cell.

Spectrophotometers are used to measure denaturation of DNA by transmitting ultraviolet rays (UV), with closely varying wave lengths, such as 260 nanometers, at which both DNA and RNA are able to absorb UV, while proteins are able to absorb UV at a wavelength of 280 nm (nucleic acids and proteins absorb UV differently).

Absorption of UV is more efficient in a single strand of DNA than in a double helix, since nitrogenous bases are what absorb UV, and while they're exposed in the single strand, they are hidden away from the UV in the double helix. Which is how we check for the extent of denaturation (more absorption of UV means there are more denatured single strands of DNA).

An injected gene doesn't have to bind to cellular DNA, and sometimes if it did it could disrupt the work of another gene, so that is also taken into consideration.

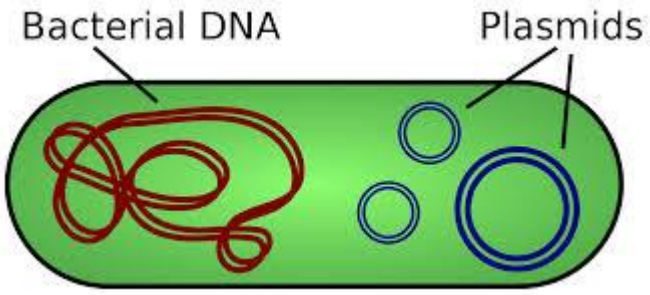
The human genome is linear and the bacterial genome is circular, but we have circular mitochondrial DNA.

We have about 20,000-25,000 genes in our body.

We have 46 chromosomes and if we stretch them it's around 2 meters of DNA.

How many genes are in a single chromosome? E coli = 4.6 million bp, encodes 4300 gene. Human = 3.2 billion bp, encodes (20000-25000) gene. We have 3.4 base pairs.

A bacteria cell contains more DNA than a virus, and moreover contains two kinds of it, chromosomal and plasmid. A plasmid has extrachromosomal DNA (isolated from bacterial DNA and replicates on its own), a plasmid is dangerous because it contains genetic material that gives bacteria the ability to resist antibiotics, and which it can give to another bacteria making it resistant too (if a human caught such bacteria from a hospital, antibiotics wouldn't work).



A human cell is more extreme, since it contains 3 billion nucleotides per cell (2 meters inside 6 micrometers). A human body (10^14 cells) contains enough DNA to travel to the sun, back and forth hundreds of times.

DNA is negatively charged because of the phosphate group which will electrostatically attract to the positively charged amino acids resulting in compaction of DNA.

How is the nucleus (which has a diameter of 0.006) capable of containing our DNA?

This means we have packing and folding processes of DNA.

Gene definition: It is a portion of chromosome or DNA that codes for a visible property (eye color, ..., etc.)

Then, scientists found that a visible property may include more than one gene so they defined it as a part of chromosome or DNA that codes for a protein.

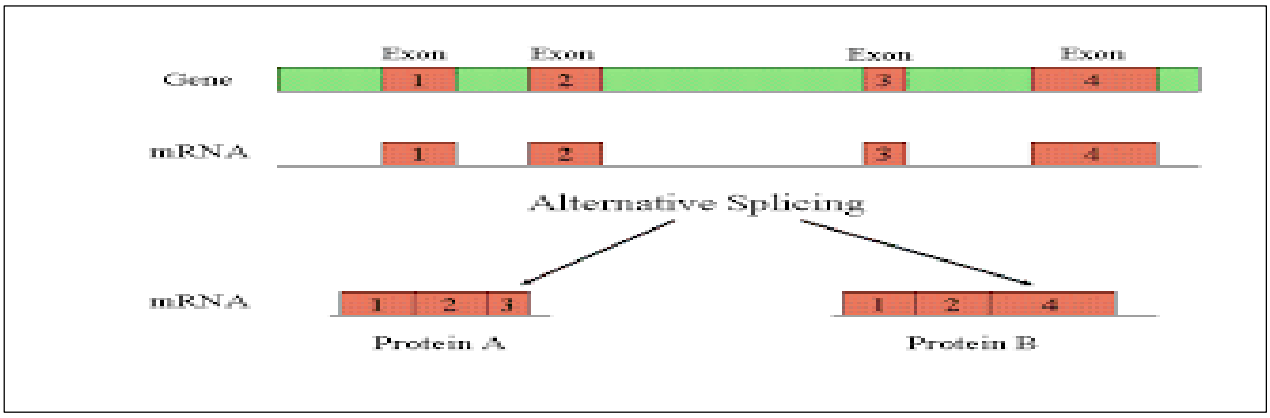
Later on, scientists discovered that a protein or enzyme may include more than one gene, so they defined a gene as a part of DNA that codes for polypeptide or RNA.

The modern biochemical definition of a gene: it's a portion of chromosome or DNA that encodes for a protein product.

Gene definition nowadays: segments of DNA that encodes into polypeptide or an RNA not a protein.

Isoforms of the same gene: A gene with several functions. It means that a gene in your liver cell can produce protein with particular property which differs from the protein produced in your skeletal muscle translated from a similar gene.

Alternative splicing: In eukaryotes, before the mRNA can be translated into proteins, non-coding portions of the sequence, called introns, must be removed and protein-coding parts, called exons, joined by RNA splicing to produce a mature mRNA. (the same DNA, the same premature mRNA and different mature mRNA which will be translated).



Telomere: sequence at the end of chromosomes that helps stabilize the chromosome. It protects the chromosome from breaking: capping, found at the end-area.

Telomerase: it's an enzyme that elongates Telomere (found in cancer cells).

Modification: is a post transcription process; splicing cutting off the introns so it becomes mature m-RNA.

30% of our genome is introns and exons (1.5% exons and 28.5% introns).

Introns: non coding, untranslated regions of DNA.

Exons: coding regions or translated regions of DNA.

The number and length of an intron or an exon doesn't reflect the importance of a gene. (in other words, we can't say that gene A is better than gene B because the number of exons in gene A is bigger).

What's the importance of introns?

1) They facilitate alternative splicing.

2) Protein diversity.

Codon gives us the amino acids and noncoding segments don't give the amino acids any translation.

The centromere area isn't translated into a polypeptide it only has the job of holding the spindle fibers, a sequence function during cell division as an attachment point during mitosis,

centromeres are rich in A=T (which makes it easier to separate the sister chromatids from each other during anaphase of mitosis; it's easier to break 2 hydrogen bonds than breaking 3).

Satellite DNA (SD): is a highly repetitive DNA sequence that can be found in centromeres. It can be less or more than 10 base pairs. 25% of the genetic material is miscellaneous (5% satellite DNA, 3% SSR (simple sequence repeat < 10 bp) 17%)

Promotor: is a segment found before the translated area of DNA, enzymes connect with it, and it is not a translated region.

Transposons: are junk DNA that has no benefit for the host cell it's also called selfish DNA. The transposon is a certain DNA fragment that will create another copy of itself (copy paste or cut paste) and transport itself to another gene which could lead to a mutation.

45% of human genetic material is transposons.

Other names of transposons:

Transposable elements.

Parasitic DNA.

Mobile genes.

Junk DNA.

Selfish DNA.

Molecular parasites.

Mutating genes.

We have two types of transposons: (they differ by length)

1) 21% LINEs: long interspersed elements 6000-8000 bp long.

2) 13% SINEs: short interspersed elements 100-300 bp long.

Supercoiling: Coiling of the coil itself which is made of two strands.

Our DNA is wrapped, which means that the double helix is also folded. We can fold them in many ways.

Super coiling is the term that explains the arrangement of the largest quantity of the DNA in the nucleus, and the super coiling and the tension increase in the translation (transcription is done so that the DNA can be accessible (حتى تتم العملية المطلوبة)

The degree of super coiling increases without any external factors.

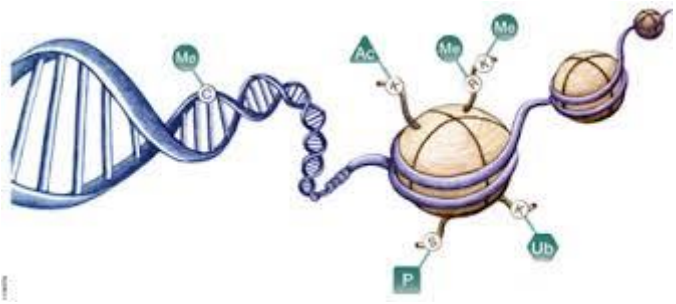
Super coiling of DNA is an intrinsic property.

Ex: the plasmid DNA in the E.coli (bacteria), (plasmid = antibiotic resistance)

استخلصوا البلازميد منها وقاموا بمراقبته بمايكروسكوب ولاحظوا كل ما تزيد الفترة البلازميد يلف على نفسه بشكل اكبر

Another example the viral sv40 DNA is being super helical and if one strand is nicked the stress is relieved.

The genome of bacteria is smaller comparatively so we can fold them in flower shaped circles and wrapping so it's more compressed, but ours is longer so it has to be in more spiral shaped and we use the histones (solenoidal) for the folding of **DNA.**



We have two types of coiling: Solenoidal and plectonemic (more in bacterial genome) and its less compact than solenoidal.

When we want to release the two strands of DNA it creates more stress resulting in supercoil, so we need an enzyme to relaxes the coiling making it more loose which is named Topoisomerase.

We have two types of topoisomers and their function is the removal of coiling. We need topoisomers for gene expression.

Topoisomerase: An enzyme that decreases the stress and the tension of the supercoiled strands by unwinding the strands to facilitate the process of translation or transcription or replication. It cuts off the backbone of the DNA, we break the bond to release the coil and then we connect them afterwards again after the removal.

Some notes about topoisomerase:

1) It breaks the phosphodiester bonds to facilitate the unwilling of the strands.

2) It converts the strands from tense to relaxed and again from relaxed to tense – play as nuclease and ligase.

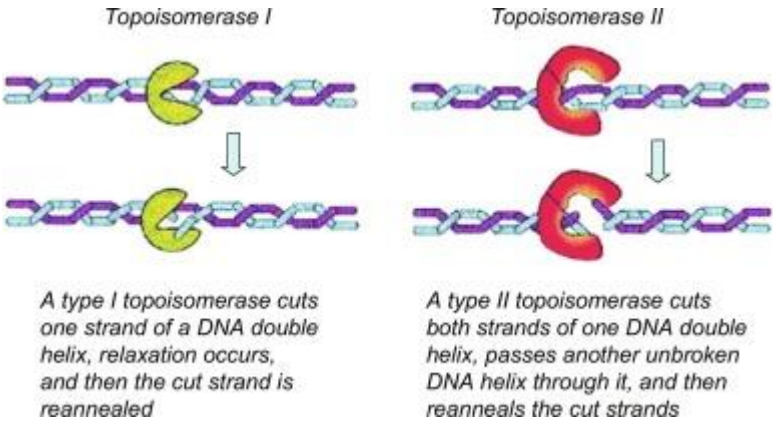
Type 1 topoisomerases: unwind one strand.

Type 2 topoisomerases: unwind two strands

Topoisomerase: Is a target for antibiotics and other drugs. Some famous topoisomerase is E.coli (DNA gyrase), giving the antibiotic like –novobiocin- to bacteria , which inhibits the DNA gyrase= topoisomerase so there is an increase in super coiling which makes it difficult to split and so the cell dies.

Eukaryotic topoisomer inhibitors: such as doxorubicin and etoposide, are used as chemotherapeutic genes in cancer therapy.

توضيح : بالخلايا السرطانية وهي من صفاتها الانقسام بشكل كبير، يعمل الدواء الكيمائي المذكور سابقا على تثبيط عمل ال topoisomers فايقاف انقسام الخلايا السرطانية وهذا يعني شيء مفيد، ولكن هذا الدواء ايضا يؤثر على الخلايا الطبيعية التي تنقسم بكثرة مثل الشعر ولذلك نلاحظ ان الشعر يتساقط عند العلاج الكيماوي ويؤثر ايضا على القناة الهضمية ويسبب اعياء وغثيان. واذا استمر العلاج الكيماوي لفترة طويلة يؤثر على الخصوبة (fertility) وللرجال يؤثر على الحيوان المنوي، وبالمرأة يؤثر على البويضة ولكن يؤثر على المرأة بشكل اكبر ومن هنا ظهرت فكرة لتجميد البويضات.



Linking number: is how many times the two strands of DNA cross each other

Writhe: is how many double strands cross each other

A nucleosome is a histone with DNA wrapped around it twice without the Clamp (H1). A nucleosome, which is a structure where DNA wraps around the histones (like a thread on a bead), contains 8 histones (two molecules of each type except for H1), forming an **octamer** which the DNA wraps around, therefore the octamer forms a base for the DNA to wrap around.

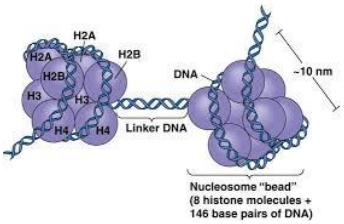
Histone is an octamer l meaning 8 subunit proteins. Histones look like beads and we wrap the DNA around it twice.

Clamp is what keeps the DNA wrapped around the histone, it is H1 histone. Histone (H1) links two nucleosomes together, therefore it's called a linker histone. The DNA segment between the two nucleosomes which doesn't wrap around the histones and connects only to H1 is called a Linker DNA.

Types of histones: H2A, H2B, H4, H3, H1 (clamp)

It's impossible to extract a DNA without extracting certain proteins with it. There are two kinds of proteins: histone proteins and non-histone proteins, also an injected gene doesn't have to bind to cellular DNA, and sometimes if it does it can disrupt the work of another gene, so that is also taken into consideration.

Histones within nucleosomes are rich in basic amino acids, making them positively charged, while the DNA is negatively charged, which results in electrostatic interaction between the histones and DNA, making the DNA wrap perfectly and firmly around the histones without budging. The basic amino acids that histones are rich in are arginine and lysine



DNA wrapping around histones is the level number one of compaction (7 folds compaction) which is not enough for all the DNA to fit inside, 7 folds compaction means that the amount of DNA inside the nucleus is seven times the amount it's capable of holding. After this level, there are levels 2, 3, 4, and 5 of compaction in order to make the DNA fit inside. In level 5 (where the chromosomes form) each chromosome performs 10,000 folds of compaction of the DNA, and since there are 46 chromosomes, a nucleus has inside it 460,000 folds (the amount of DNA it's capable of holding).

Each nucleosome has 146 base pairs of DNA wrapped around it, and DNA turns it around, 1.65 turns (the length of linker DNA is 65 base pairs).

Loose DNA is called euchromatin (active portion), compact DNA is called heterochromatin (inactive portion). Euchromatin should be able to turn into heterochromatin, and heterochromatin to euchromatin when needed, it's a very dynamic process, and histones play a very important part in this process, which means histones are not only packaging tools but also regulatory.

Chromatosome: The folding of DNA around the histone along with the clamp. Chromatosome is the nucleosome but with the clamp (H1).

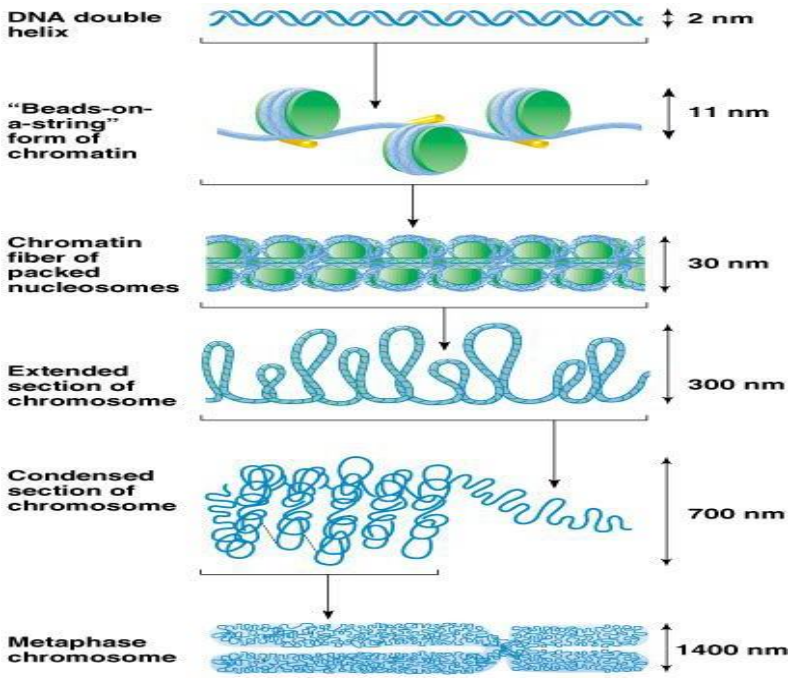
Linker DNA is the unwrapped areas of histones.

Chromatin is the repeating sequence of nucleosome.

Chromatin remodeling: tightly or loose

We have two types of chromatin: Euchromatin (loose chromatin) and heterochromatin (compressed chromatin)

Chromatosome is the folding of chromatins



We need something that opens up the coiling of DNA: histone acetylation and de-acetylation.

Adding acetyl group to the histone makes it more loose and then it's recognized by the RNA polymerase for gene expression by an enzyme, histoneactyltranfrase (HAT).

Transfers the acetyl group from acetyl-coA to a histone

Histone Deacetylate Enzyme (HDAC) removes the acetyl group from the histone.