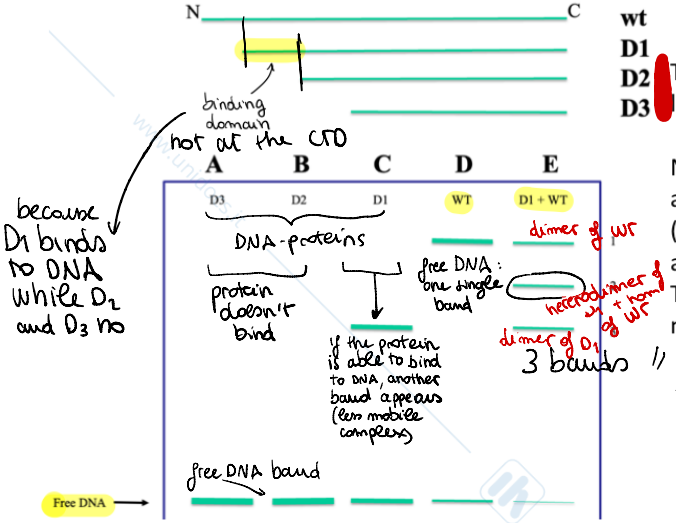
Band shifting assay problem: EMSA: electrophoretic mobility shift assay

We have to understand how the proteins bind to the DNA, what is the mechanism of their activity. Among them the transcription factors are particularly important because regulate the synthesis of the transcript. To understand which is the mechanism and how they bind we start with this problem:

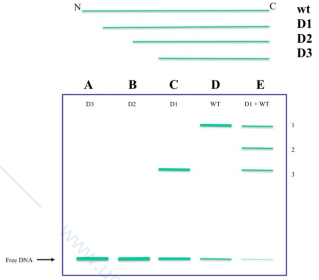


This is a band shifting gel, in the last line wee see the free DNA.

N-C is the protein and below there are three truncated version of it. (we start from the N terminal part and go on the length of the protein. There is a reduction in the molecular weight)

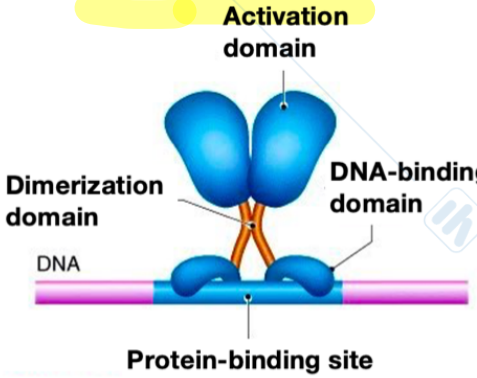
DNA bound to a protein shows less electrophoretic mobility than free DNA'

5. Transcription activation mechanisms



A band shifting assay was carried out to study the binding domain of a protein. The wt lane contains the original protein, D1, D2 and D3 are lanes containing the same protein but with different deletion sizes (D1 smallest, D3 biggest). Binding occurs only with the wild type protein and the one containing the D1 deletion. This means that in lane D2 and D3, the binding domain has been cut out, so it is positioned on the final stretch of D1: the binding domain is, therefore, close to the N-terminal domain of the protein.

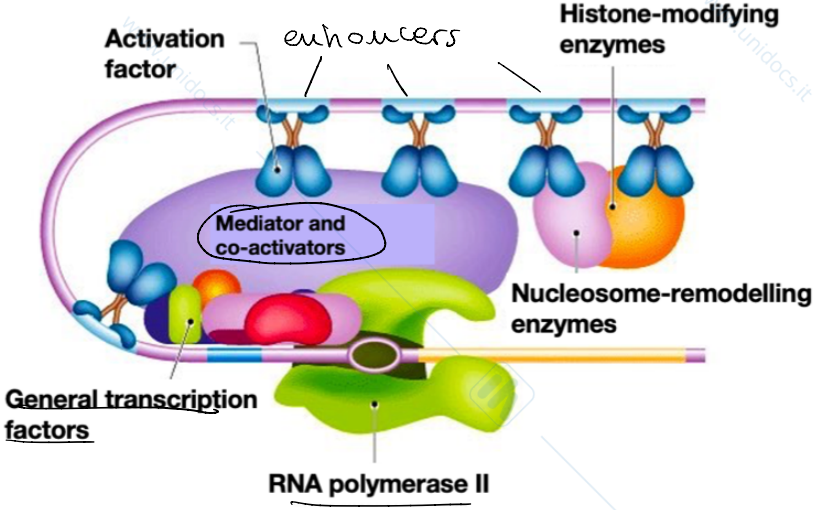
Lane E contains both the wild type protein and the one with the deletion D1. It has three bands, which is unusual, but explained by the fact that the single polypeptides work as dimers: the first band is a homodimer of the wild type proteins, the second is a heterodimer formed by a monomer of the wild type and a monomer of D1 and the final and is a homodimer of D1.



To answer to the second question, usually transcription factors active form is a **dimer**, and usually transcription factor as a monomer does not work so it cannot bind to DNA.

The protein is made of distinct **domains**, distinct portion of the polypeptide with each a specific job: one portion has to bind to DNA the other one has the job of contacting the transcription apparatus.

But there could be several transcription factor that work at the time, creating a **network** of interactions.



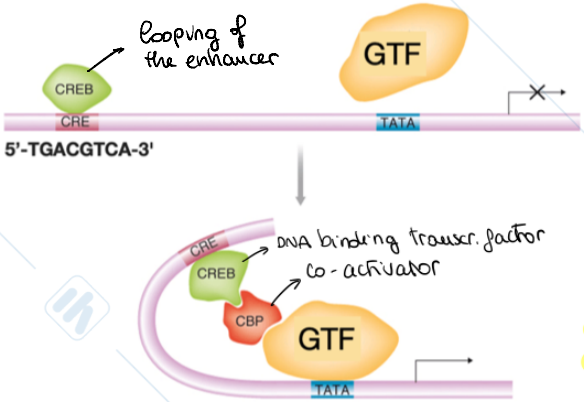
Transcription factors can bind to **enhancers** (short sequences), the enhancers contact the transcription apparatus. Transcription is then activated due to the recruitment of pol II at the transcription start site, the recruitment is increased by the **frequency**.

(frequency)

As long the interactions are increased through the protein this would favor the transcription. The interactions are all **weak.**

Some transcription factors can even affect chromatin.

Transcriptional co-activator: CBP



CREB = cAMP responsive element binding protein
CBP = CREB binding protein

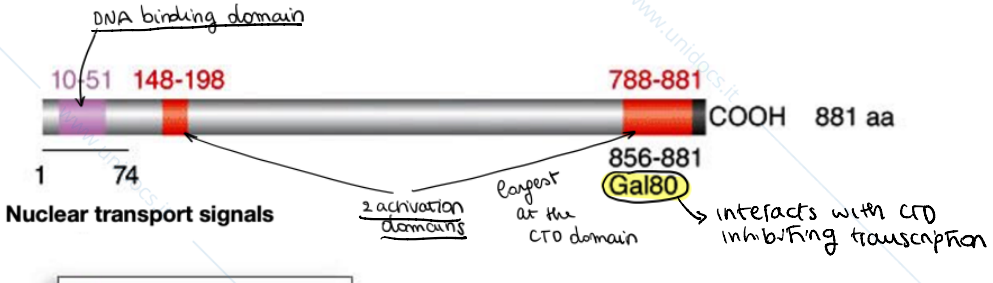
The looping for the enhancers can also occur when there are the so called **CREB motives**.

CBP is a **co-activator**, it is upstream transcription factor which do not binds to DNA. Some transcription factors miss the DNA binding site and these are called **co-activators**, they make a bridge between a DNA binding transcription factor and the DNA.

gal 4 PRESENCE activates transcription of GAL genes
gal 80 ABSENCE //

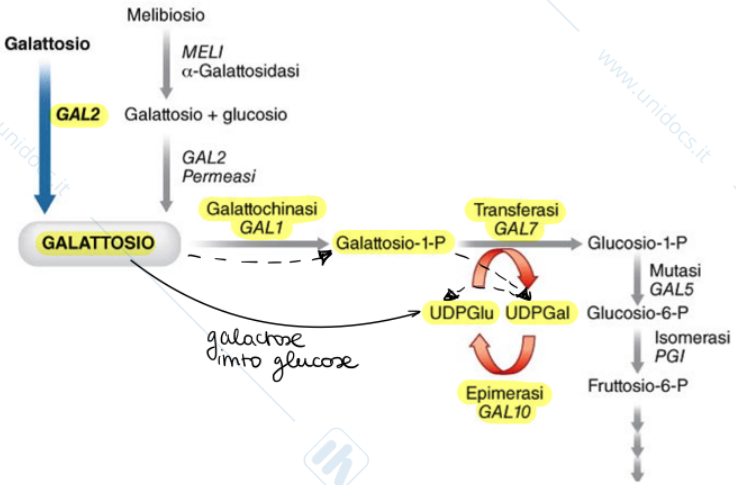
Metabolic pathway of galactose in S. cerevisiae:

The inducibility of transcription factors is an important way to link the transcription program to the need of the cell, this is the case of **gal4**, an inducible transcription factor long around 900 amino acids, it is constituted by several domains.

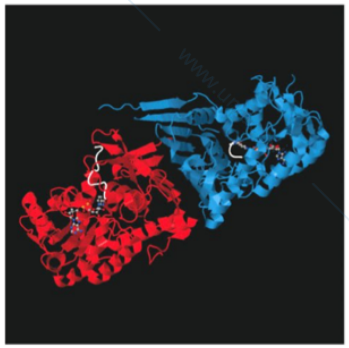


From 10 to 51 nucleotides there is the DNA binding domain then there are two activation domains, the largest is at the carboxyl terminal domain; **Gal80** can interact with this domain and so inhibit transcription.

Gal4 is induced by **galactose**, because it can sense the presence of this monosaccharide in the medium; the cerevisiae prefers eat glucose, when glucose is absent then cerevisiae will eat galactose but in order to metabolize another sugar, the cells need to modify their transcriptional program. Gal4 is able to do this, by increasing the expression of the genes in this metabolic pathway



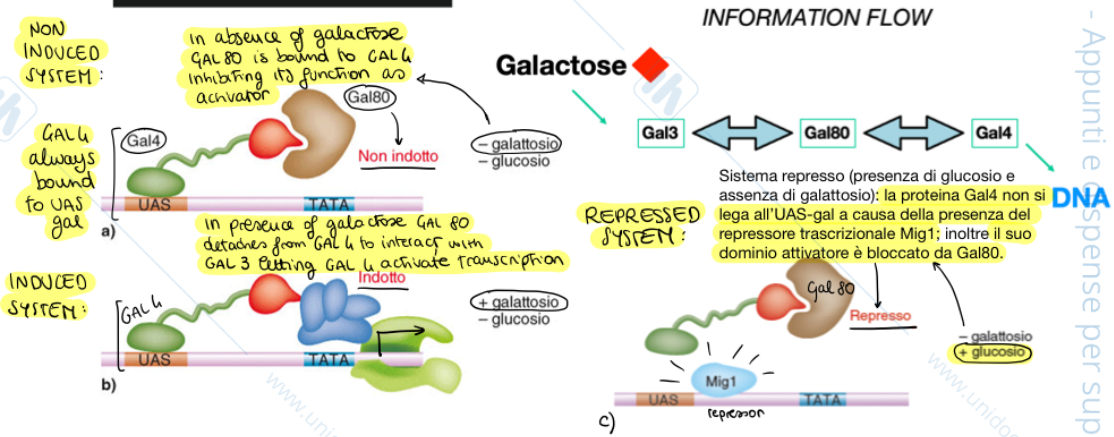
They can transform galactose into glucose.
These genes are usually shut off (they are not transcribed), by sensing the lack of glucose and the presence of galactose, gal4 we start the transcription of these genes.



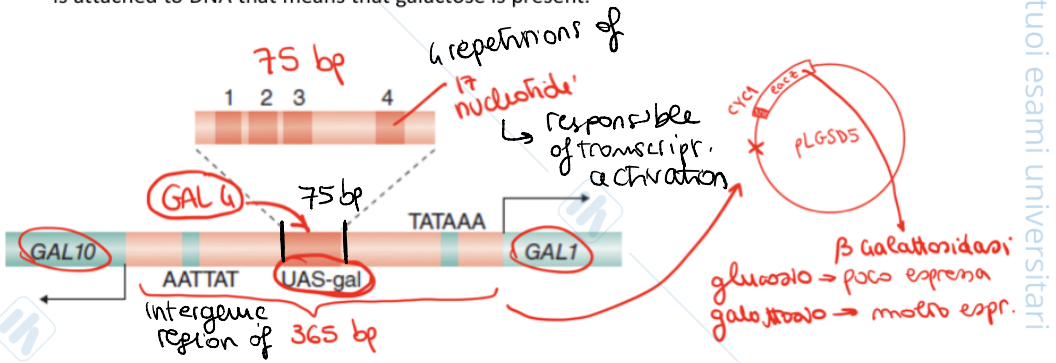
However, gal4 cannot sense the presence of galactose alone, it must bind to gal80.

Gal4 is bind directly to gal80, then the galactose removes gal80 from the domain of gal4 and so it can binds to DNA. because

How galactose is able to remove gal80? because it is attached to another protein that is **gal3** that is the true sensor of galactose.



When gal3 is alone and gal80 is attached to gal4 the information we can obtain is that there is no galactose in the environment. On the contrary, when gal3 is attached to gal80 and gal4 is attached to DNA that means that galactose is present.



I geni GAL1, GAL7 e GAL10 sono presenti sullo stesso cromosoma (cromosoma 2), sono vicini tra loro e costituiscono un cluster genico denominato GAL-cluster.
Ogni gene ha il suo promotore provvisto di una sequenza TATA box, ma i diversi geni sono co-regolati e questo dipende da una sequenza posta a monte del promotore, chiamata UAS-gal.
UAS-gal: i geni GAL1 e GAL10 sono adiacenti sul cromosoma 2 e trascritti da due promotori divergenti separati da una regione intergenica di 365 bp. la "vera" UAS-gal è una regione di 75 bp del frammento di 365 bp precedentemente isolato, nella quale si trovano 4 ripetizioni di 17 nucleotidi che sono i reali siti responsabili dell'attivazione trascrizionale.

DNA-biding structures of PROTEINS:

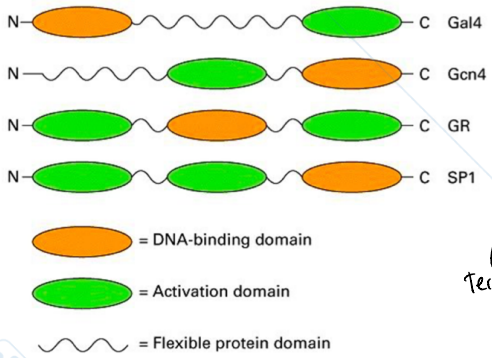
The study of protein made possible to develop new technologies.

The definition of a protein's domain: **A protein that can fold and function in an independent manner from the rest of the protein**.

What really characterize the functions of a protein are its domains, by changing them you change the fold and the activity of the protein itself.

Recombination is a driving force in evolution because it is possible to produce different protein by just exchange module of the domains.

One of the most important domains of a protein is the **DNA binding domain**. It can be in different places of a protein.



Exist three different main type of DNA binding domain in eukaryotes:

- **Zinc finger**
- **Homeodomain**
- **bZip e bHLH** (basic zip and basic helix-loop-helix)

Nel dito di zinco (zinc finger) una protuberanza (il dito) che interagisce in modo specifico con il DNA a livello di un solco maggiore della doppia elica. L'atomo di zinco è coordinato alla base del dito da quattro ligandi, che possono essere, in combinazione variabile, cisteine o istidine. Il classico zinc finger ha un atomo di zinco coordinato da due cisteine e due istidine (tipo Cys2His2) e il dito, lungo una ventina di residui amminoacidici, si struttura a formare un'α-elica da un lato e un foglietto β dall'altro. Un atomo di zinco (in rosso) è coordinato da residui di cisteina e istidina a formare una struttura in grado di interagire, di norma, con il solco maggiore del DNA.

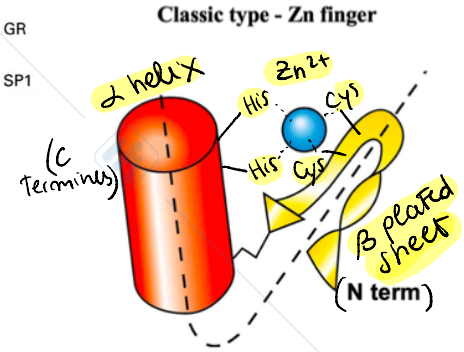
Il motivo elica-giro-elica (helix-turn-helix) è costituito da due α-eliche connesse da un breve tratto esteso, sufficientemente flessibile da permettere il ripiegamento di un'elica sull'altra. le due α-eliche formano tra loro un angolo fisso, determinato da interazioni deboli. L'α-elica C-terminale è chiamata elica di riconoscimento riconoscimento, poiché costituisce la porzione del motivo strutturale che interagisce fisicamente con il DNA, a livello di un solco maggiore. La sua sequenza amminoacidica differisce da proteina a proteina e determina la specificità del riconoscimento. L'α-elica N-terminale ha invece un ruolo strutturale, permette cioè di posizionare l'α-elica di riconoscimento in modo corretto per l'interazione con il DNA.

Zinc finger:

Is the one which have more success in evolution. It is called finger for the bidimensional structure that resemble the one of a finger. The zinc is important for the tridimensional structure which has different variations.

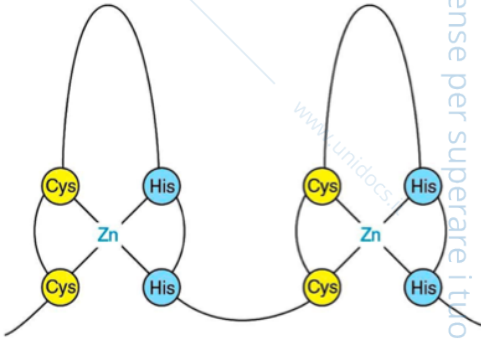
There are three types of zinc finger:

- **The classic** – it has **1 zinc ion** which is coordinated to **2 cysteine** and **2 histidine**
- **Nuclear receptor** – **1 zinc ion** and **4 cysteine**
- **Zinc cluster** – **2 zinc ions** coordinated by **6 cysteine**

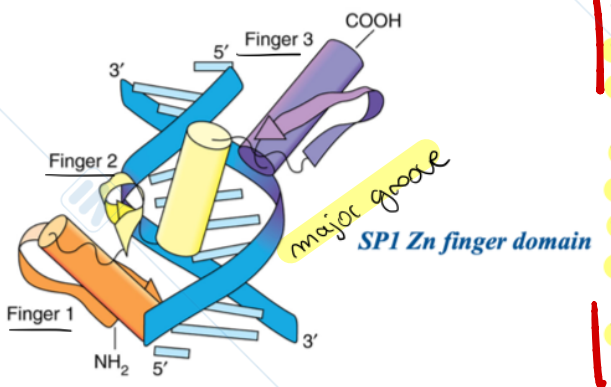


The N terminal is on the right side, there are two adjacent secondary structure which are 1 β-strand followed by an α-helix. These secondary structures will bend the zinc finger which is located at the bottom. The zinc ion will bind with the initial part of the secondary structures they bind thanks to a **metallic coordination** made possible by the amino acids.

The finger idea has been suggested because usually proteins have several of this kind of domain in tandem. Usually, we can find up to three, five, seven times the finger that are adjacent one another.



How this domain interacts with DNA:



The key point is the interaction with the major groove, the tip of the finger enters it.

The three modules of the finger will interact with the major groove, they will slightly bend to follow the pathway of it.

The interaction is **sequence specific**.

La cerniera di leucine (leucine zipper) è un motivo strutturale che deriva dalla dimerizzazione in parallelo di lunghe α-eliche avvolte una sull'altra. Le due α- eliche sono anfipatiche, cioè hanno una faccia idrofobica e l'altra polare. le α- eliche contengono una leucina ogni 7 residui. ogni due giri c'è una leucina, posizionata sempre sulla stessa faccia dell'elica. Il ruolo delle leucine è quello di permettere l'interazione tra le due α-eliche e il loro superavvolgimento attraverso la formazione di interazioni idrofobiche. Due α-eliche di due monomeri distinti formano un'elica superavvolta stabilizzata dalla presenza di residui di leucina ripetuti ogni due giri di α-elica. L'interazione mediata dalle leucine riguarda una parte delle eliche, quella definita dominio di dimerizzazione. Distalmente a esso, su ognuna delle due eliche vi è il dominio di legame al DNA, con proprietà basiche, costituito dalla porzione rimanente di α-elica che va a posizionarsi all'interno del solco maggiore del DNA.

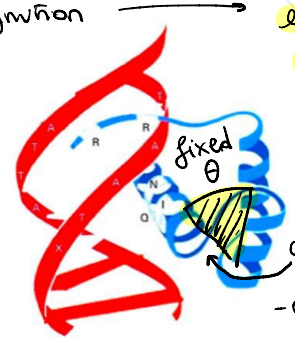
L'omeodominio è costituito da tre eliche invece che due, con due eliche strutturali e una di riconoscimento.

3α helixes { 1 structural / 2 for recognition

Homeodomain:

This domain is a **helix-turn-helix** (very common in bacteria), is called **homeodomain** because it is a domain that is part of transcription factor which determines the identity of the cell development.

It is constituted by two α-helixes, one of them will interact with the DNA entering the major groove.



2α helices connected by a flexible trait

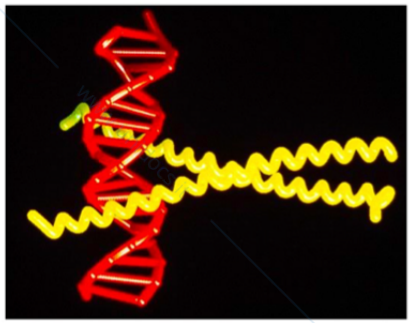
- one (N-terminal) with structural role
- one (C-terminal), for recognition, interacts with the major groove of DNA

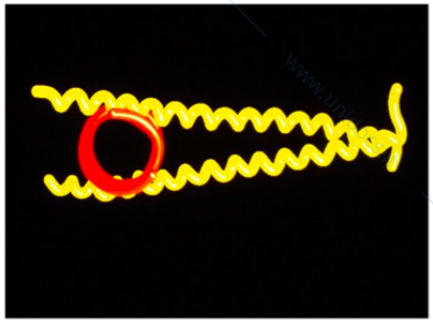
bZIP and bHLH:

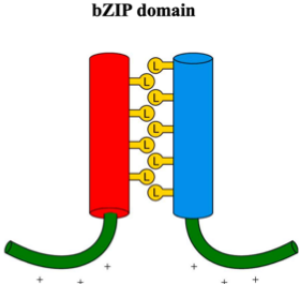
They are very common in eukaryotes but not as common as the zinc finger.

- **bZIP:** LEUCINE ZIPPER:

It is made by two long α-helixes that come together to form a sort of x. The first portion of amino acids make possible to contact the DNA, the other part is the so-called **leucine zipper**.







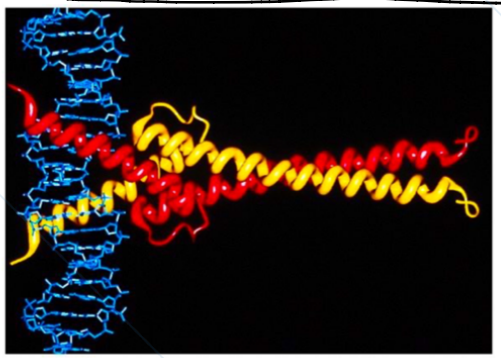
The leucines make a hydrophobic interaction to create the zip. Thanks to the interactions of the leucines the first part of the zip will be positive so that it can accept the negatively charged DNA.

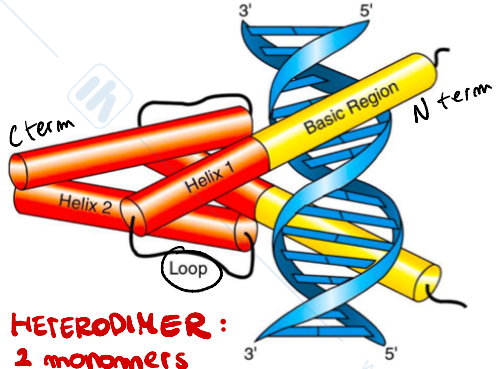
one leucine every 7 residues, every 2 rounds

Le proteine che contengono il motivo elica-ansa-elica (helix-loop-helix) agiscono sempre in forma dimerica, di norma come eterodimeri. Ogni monomero contiene un motivo lungo circa 60 amminoacidi, formato da due α-eliche connesse da un'ansa. l'α- elica H1, più lunga, posizionata all'estremità N-terminale del dominio, contiene residui basici che favoriscono l'interazione con un tratto di 6 nucleotidi del DNA chiamato E- box; quella più corta H2 al C-terminale serve per la dimerizzazione, in modo analogo a quanto abbiamo visto per le cerniere di leucine.

- **bHLH:**

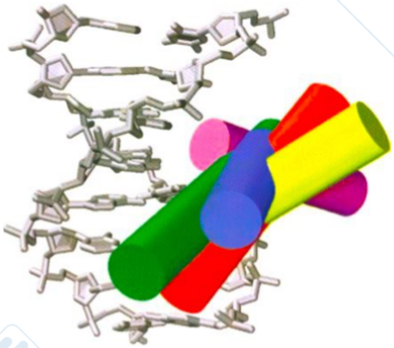
Very similar to the basic zip situation, there is a positive portion that will accept the DNA and in this case, we have a loop and 2 α-helixes. One of the helixes and the loop will interact in order to make enter the other helix inside the major groove.



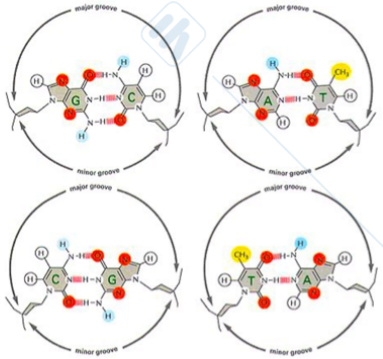


HETERODIMER: 2 monomers composed of 2 α helices connected by a loop

the longer one interacts with DNA



All the three DNA binding domains will put an α-helix inside the major groove. The helix is often the structure interacting with the major groove. But why always the major groove? Because in it we have higher diversity of the chemical functional groups.

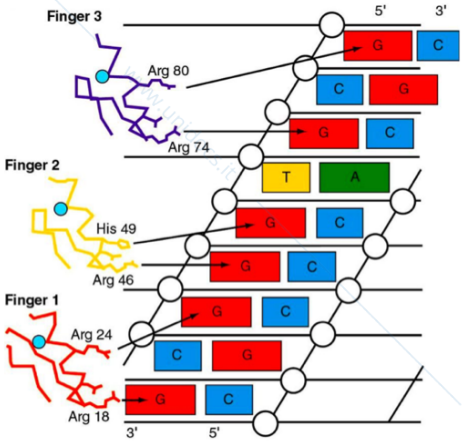


Specific interactions: Zif268 the murine TFIIIA

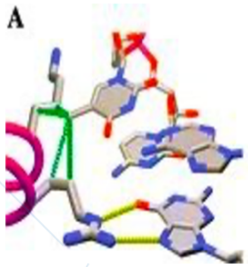
These interactions have been determined in the crystal form of DNA complex.

The specificity is not simple, there are only some specific 1:1 contact.

For example: arginine many times in the crystal complex interacts with guanines, but they can also be recognized by histidine. So it is not possible to tell with certainty what will the amino acids bind to a specific base



TFIIA has a zinc finer domain so histidines and arginines are present.



The interaction involves 2 hydrogen bonds, only arginine can binds to guanine with a double bond.

Sometimes instead is the adenine that interacts with the asparagine.

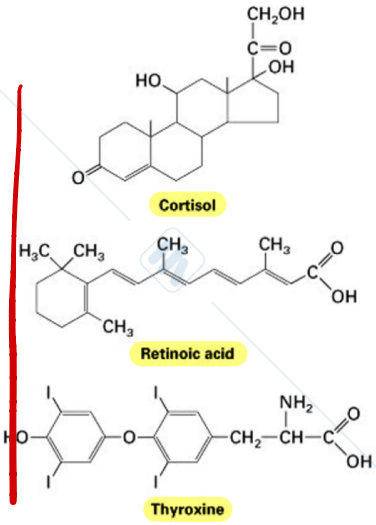
Nuclear receptor – inducible transcription factors

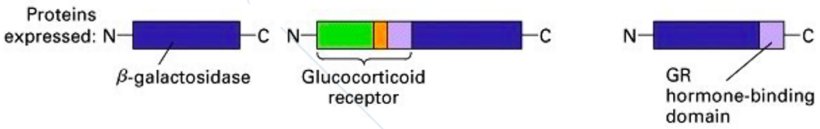
There are many nuclear receptors in a cell, they constitute a very large family and a super family of proteins that can be divided in several sub families that have all their domain organization.

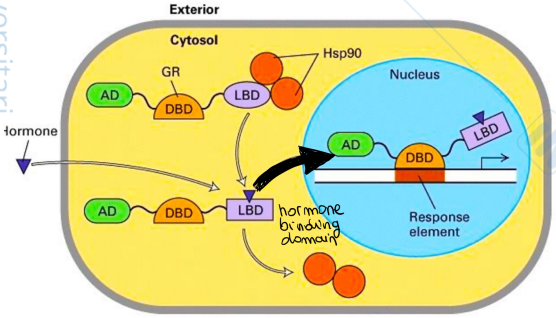
The proteins have quite all the same domain organization, the DNA binding domain is always in the middle, in the direction of the CTD there is the **hormone binding domain** (the hormone will contacts the transcriptor which will be activated so it is inducible) and at last there are the variable region that are the activation domain which are different for each family.

How they work? Firstly there is the recognition of the motives that are specific for each family. The sequences can be **palindromic** or **tandem repeat**.

Some of the hormones that induce the activity of these transcription factors:







Control cell in the first line while on the second the cells that have been exposed to a specific hormone.

A, B and C where transfected with different substances that are shown below the images of the cell.

The action of the hormone is: by binding to the hormone binding domain the protein will be translocated from the cytosol to the nucleus.

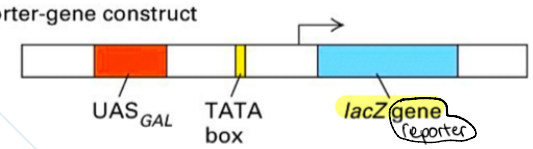
4

CAT assay:

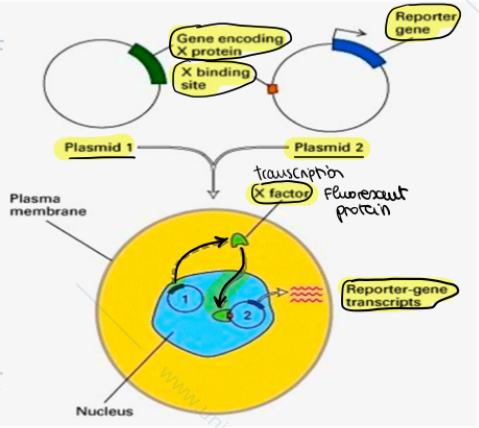
Is a technique used to establish the target function in this case transcription. The assay is made in vivo.

The gene reporter firstly was **lacZ** but now is used a fluorescent protein, in the case the promoter works and the gene is transcribed we can be sure of it if a fluorescence is perceived.

In vivo



To study a transcription factor, we do not change the promoter but we do something else.



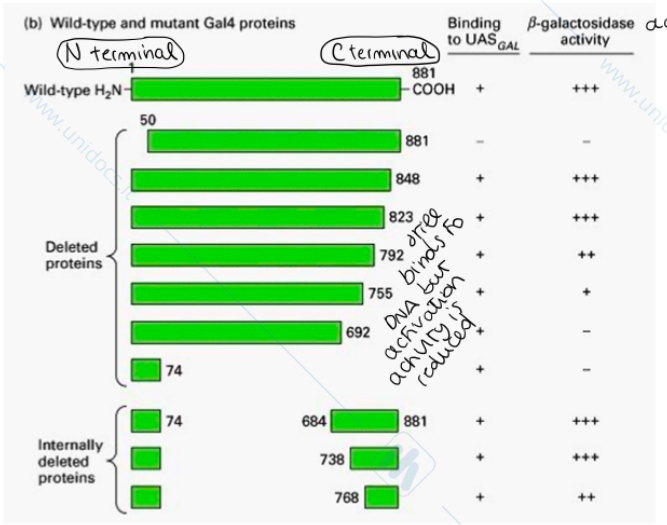
The promoter is a weak promoter meaning that does not read a high level of transcript unless it does not bind to a transcription factor.
If we transfect the cell with another plasmid that express the transcription factor that will be transcribed and the protein that will translocate the transcription factor will also activate the reporter gene.

By changing the protein we can study the transcription factor.

We can also change the transcription factor and so the effect on its modification on the cell.

To study protein-protein interactions, we can use a combination of two analyses: the CAT assay and the yeast two-hybrid system.

The CAT (chloroamphenicol acetyltransferase) assay is a method that uses a reporter gene to study gene expression. The principle behind the basic CAT assay is that CAT is responsible for acetylation of CAM, an antibiotic which causes decrease in bacterial growth (so it does not affect Eukaryotic cell growth). The promoter of our gene of interest is cloned onto a plasmid containing the *cat* gene. The plasmid is then inserted into a Eukaryotic cell and put in a sample with CAM. The solution is let sit to allow the reaction to occur and, after a while, the concentration of acetylated CAM will be measured because it corresponds to the amount of CAT produced, therefore giving us the hypothetical concentration of protein produced by our gene of interest. Another good reported gene used for this kind of assay is *lacZ*, which is also a bacterial gene and can be labelled with a GFP.



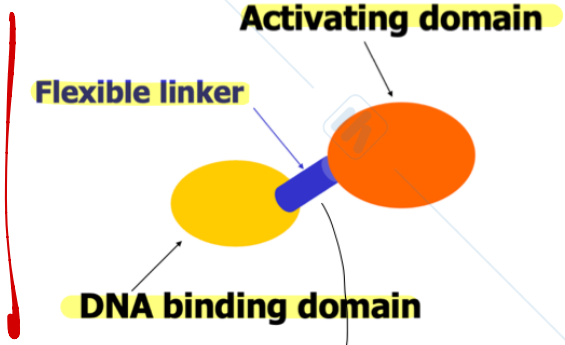
The second-last column on the right will show us if the gal4 binds to its motives using a DNA footprint assay. The last column shows the activation activity of the protein.

If we delete the N terminal domain which possesses the zinc finger, the entire protein does not bind to the DNA and does not activate, it does not work as a transcription factor.

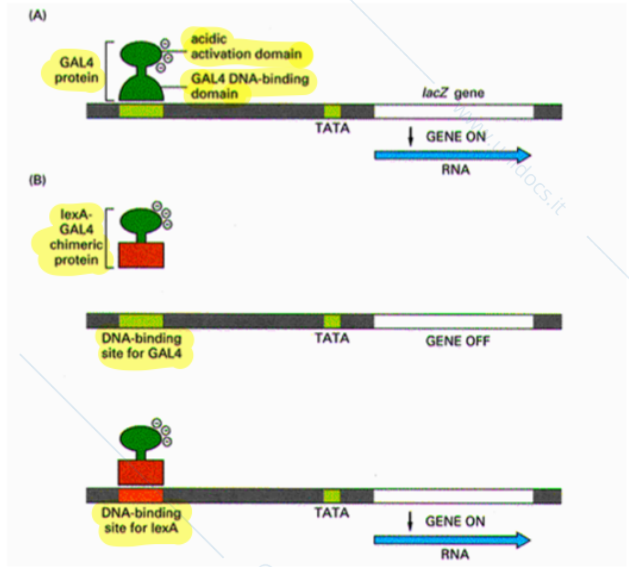
Then we progressively truncate form the CTD end. They will still bind to DNA but the activation activity is reduced to zero until 200 amino acids are lost.

Then another experiment where the deletions were internal and in this there is no change in the binding of DNA or the activation activity.

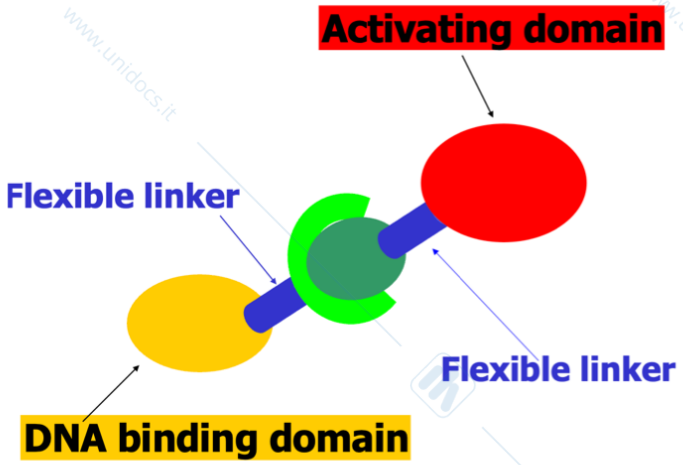
So we understand that the last part and the first one are essential for the protein activity.



The flexible linker keeps the two domains covalently linked to each other.



If A and B interact forming a complex with high affinity then I can get a new transcription factor composed of the 2 initial domains.

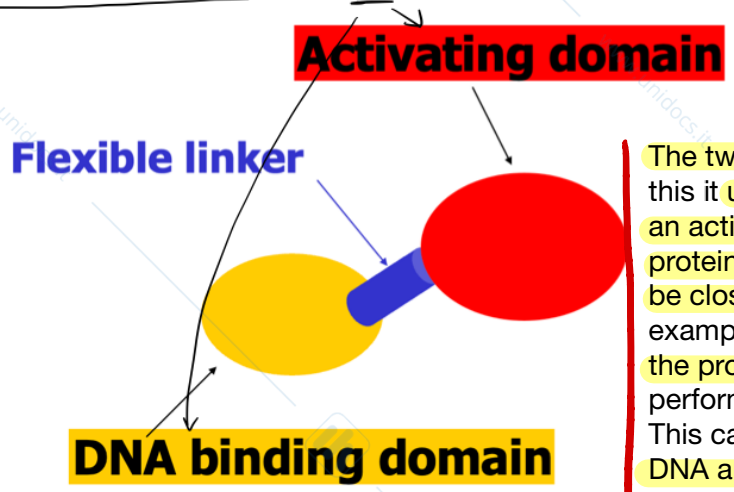


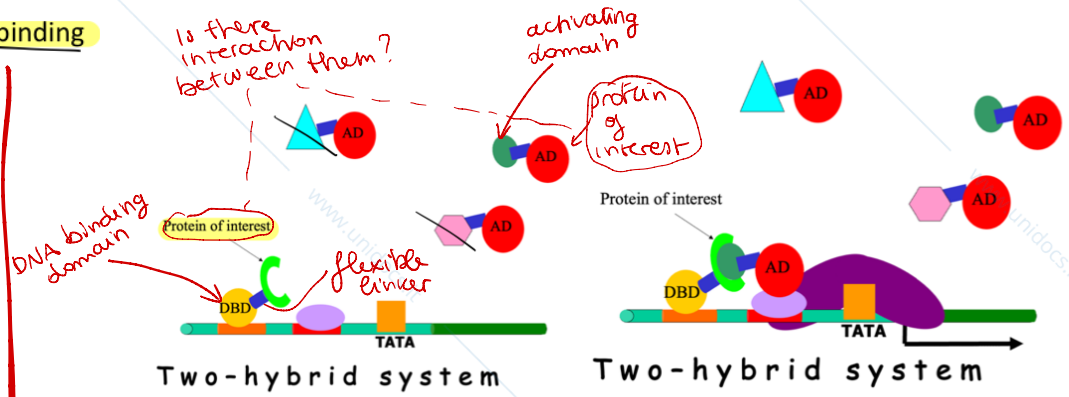
We can therefore change the protein by fusing the activation domain to the DNA binding domain of another.

The motives that constitute the promoter can be considered as domains.

Two-hybrid system:

A transcription factor is constituted mainly by two domains.





The two-hybrid system is the analysis that allows us to study protein- protein interactions, and to do this it uses transcription factors: inducible transcription factors have a DNA binding domain (bait) and an activating domain (prey) and in this technique, the two are separated and each bound to one of the proteins we are studying. This means that if there is interaction between the two, the two domains will be close enough to bind the promoter region and transcribe the gene (which can be the cat gene, for example); otherwise, if the two proteins do not interact, the two domains will be too far apart to bind the promoter of the gene, therefore not activating transcription. We can use the GAL4/UAS system to perform this analysis (GAL4 binding to UAS sequences activates gene expression).
This can make us think of transcription factors more as proteins with two main domains, one binding DNA and the other responsible for activation, linked by a flexible polypeptide.
The approaches resulted in the detection of 957 interactions involving 1,004 S. cerevisiae proteins. These data reveal interactions that place functionally unclassified proteins in a biological context, interactions between proteins involved in the same biological function, and interactions that link biological functions together into larger cellular processes.

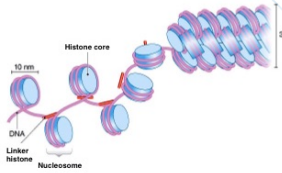
7. Architectural genome organisation

There are 3.3 x 10^9 bp in the human genome and each base is 0.34 nm in B-DNA, therefore the total length of the human DNA molecule is 2 m, which must be packaged into a nucleus with a diameter of 5 µm. This means DNA must be condensed by a factor higher than 100 000x. What we have just described represents the 'packaging problem'. It is not unique to human beings.

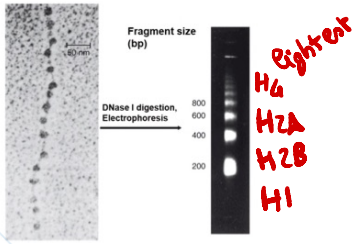
There are different levels of compaction generally divided into euchromatin and heterochromatin:

- euchromatin is less densely compacted than mitotic chromosomes and is dispersed throughout the nucleus. It is a more accessible part of the chromosome;
- heterochromatin is highly compacted, more similarly to mitotic chromosomes, which means it is a less accessible part of the molecule. We can distinguish between constitutive and facultative heterochromatin: constitutive heterochromatin is composed of genomic regions that are never expressed and usually have a structural role; facultative heterochromatin is composed of regions of the genome that are expressed or repressed depending on cell identity or tissue (i.e. the X-chromosome in mammals).

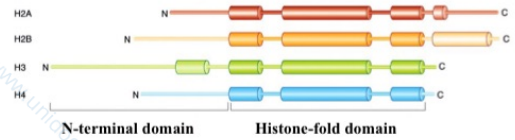
The basic structural unit of chromatin is the nucleosome, consisting of a portion of DNA coiled around a core of histones in a leftward manner (10 nm structure, 260 kDa).



To determine the amount of bases wrapped around the nucleosome core, DNase I digestion was used (DNase I footprinting assay). DNase I mainly performs its endonuclease activity on linker DNA, the one that connects one histone another. We can appreciate from electrophoresis results that a nucleosome contains about 200 bp of negatively supercoiled DNA.



Histones are the main protein component of chromatin. There are five different histones in eukaryotic cells (from largest to smallest): H1 (21.5 kDa), H3, H2B, H2A and H4 (11 kDa). These proteins are very abundant, have highly conserved structures (especially H3 and H4, least conserved is H1, which can also be H1° and H5) and are positively charged at neutral pH, in order to interact with the negatively charged DNA sugar backbone. They can be also divided into core and linker histones: the histone core is an octamer composed of 2xH2A, 2xH2B, 2xH3 and 2xH4. H1 are the linker histones.



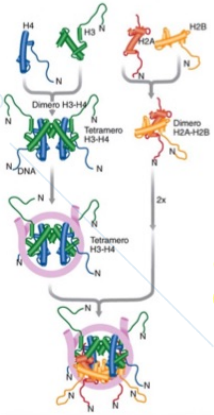
There will be a total of 1 tetramer of H3-H4 and 2 dimers of H2A and H2B, one of the two dimers of H2A and H2B will be placed behind the tetramer while the other in front of it.

The tetramer in the meantime will contact the DNA forming a plane structure.

Histone fold:

The histone is mainly composed by α helixes, the number of helixes is different for each type of histone, but the central helix is always bigger than the others.

The helixes form the **histone fold domain** and this domain determine the 3dimensional structure of the histone. This domain interacts with the domain of another histone so it is a **protein-protein interaction domain** and it can also be found in other proteins.



N-tails branch out from the nucleosome: these structures can interact directly with other proteins and can be modified during the cell cycle in order to change the chromatin's condensation level.

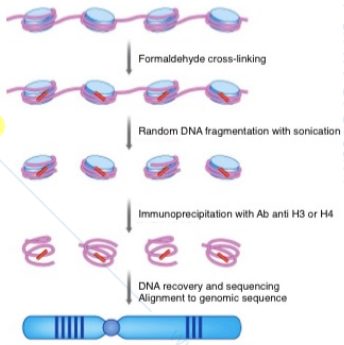
H1 histones belong to linker histones. They are not part of the nucleosome core structure, but can increase DNA compaction in two ways: they allow the DNA strand to wrap more tightly to the nucleosome core (globular domain), but they can also interact with other H1 (extended domain), shortening the linker DNA. The latter interaction is structurally fundamental for the creation of the 30 nm fibre, in which several histones positioned close to each other in an helicoidal manner. When compacting the DNA wrapped around the nucleosome core, H1 increases the amount of bases interacting with the nucleosome: with no linker histone, 146 bp of DNA interact with the core (1.75 turns), but when H1 is involved, the number of nucleotides increases to 200 and the DNA performs 2 whole turns around the nucleosome.

Histones can be post-translationally modified in several ways: ubiquitylation, phosphorylation of serines and threonines and methylation or acetylation of lysines.

Nucleosome mapping by chromatin-immunoprecipitation (ChIP) is a type of immunoprecipitation technique used to study the interaction between proteins and DNA in the cell. It aims to determine whether specific proteins are associated with specific genomic regions, such as transcription factors on promoters or other DNA binding sites. ChIP also determines the specific location in the genome that various histone modifications are associated with, indicating the target of the histone modifiers.

The general method follows these steps:

- DNA and associated proteins on chromatin in living cells or tissues are cross-linked (formaldehyde or UV light);
- The DNA-protein complexes (chromatin-protein) are then sheared into ~500 bp DNA fragments by sonication or nuclease digestion;
- cross-linked DNA fragments associated with the protein(s) of interest are selectively immunoprecipitated using an appropriate protein-specific antibody;
- the associated DNA fragments are purified. Their sequence is then determined and mapped to the genome.



Nucleosome positioning is an important aspect of chromatin architecture, the application of next-generation sequencing (NGS) to ChIP (ChIP-Seq) has revealed insights into gene regulation events that play a role in various diseases and biological pathways, such as development and cancer progression.

Nucleosomes are dynamic elements, so the mapping of their protein structure is essential to understand the active and continuous genetic architectural remodelling, especially in transcribed or regulatory regions of the genome.

The highest level of organisation of the genetic material is the chromosome itself. Euchromatin and heterochromatin can coexist on the same chromosome.

The simplest DNA organisational structure is composed of the DNA molecule and its nucleosomes ("beads on a string", which can be considered euchromatin). This structure, as mentioned before, is the 10 nm fibre.

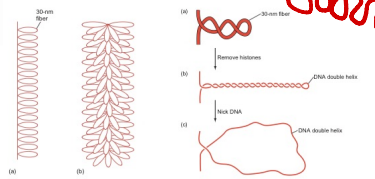
A helical disposition of the 10 nm fibre gives rise to the 30 nm fibre (solenoid), which can loop and arrange itself in different ways, forming 300 nm scaffolds. The 30 nm chromatin loops can be relaxed by removing the histones, releasing the supercoiling.

RADIAL LOOPS

chromo scaffold

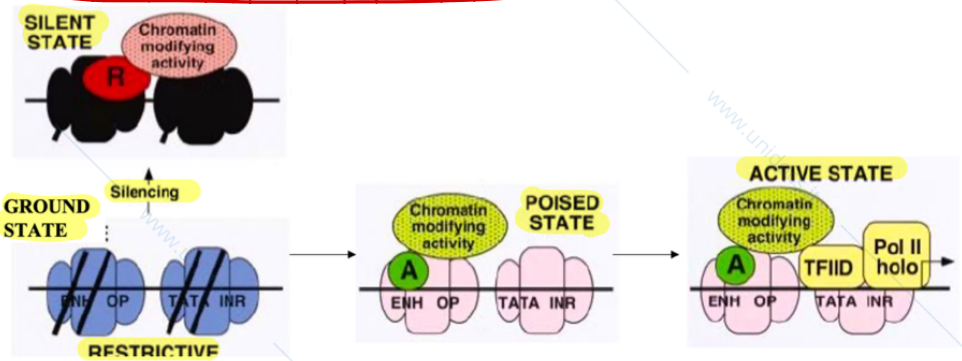
The next level of chromatin structure are the **radial loops** (the loops are stocked along an axis).

The chromatin forms loops of DNA. At the base of the loop there are sequences of DNA that have the function to bind to the protein skeleton, they are called **scaffold binding sequence**.



Those fibres can then be organised in a more complex manner establishing a 700 nm condensed structure (highest organisational level right after the chromosome, which is 1400 nm).

Ground state is restrictive silent state is more restrictive:



If there a nucleosome wrap around itself the DNA sequence containing the transcription start site (tata box and promoter), that part of DNA will become inaccessible therefore the gene will become **silent**. But if we breakdown the chromatin structure then the DNA will become accessible to transcription factors.

Not only the higher-level structures of chromatin can repress transcription, but even the single nucleosome alone is a stronger inhibitor.

So to map the position of a nucleosome is very important to understand the regulation of genes expression.

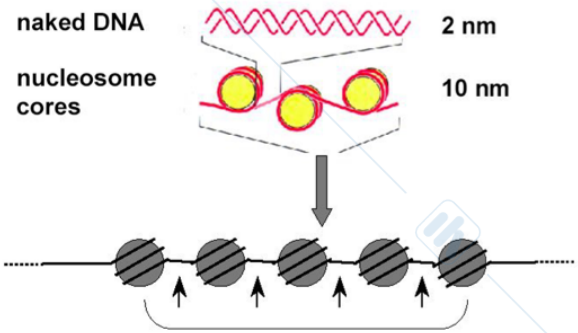
To map the position of the nucleosome we can use again the chromatin immunoprecipitation method or even the enzyme activity of **DNase I**, it can bind to the duplex and then cut a strand of it to insert ss cut.

The DNase can bind only to DNA and not to tRNA.

The chromatin structure is necessary but not sufficient for gene transcription, elongation, and synthesis of the transcript. True that a gene is transcribed if it is in a opened structure, **but not all the gene in a opened chromatin structure are transcribed**.

ANALIZE CHROM STRUCT.

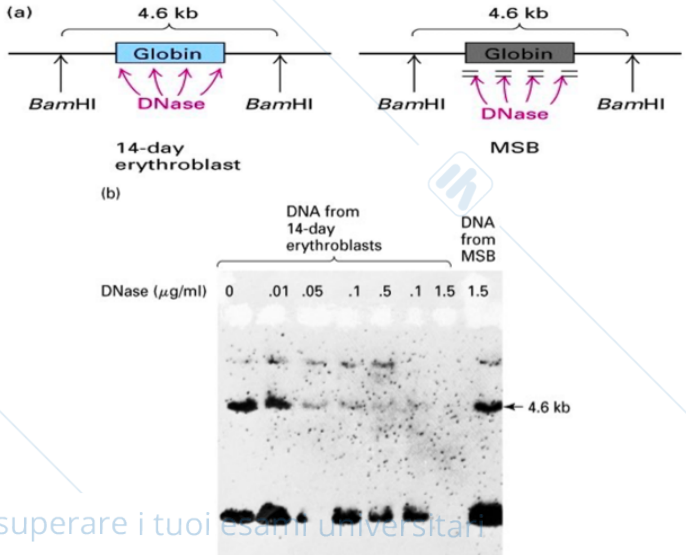
The DNase I is able to recognize the accessible and inaccessible chromatin conformations.



1. Isolate nuclei
2. Add increasing amounts MNase I for fixed time
3. Purify DNA after deproteinization
4. Separate DNA on agarose gel
5. Stain w/EthBr or blot and hybridize w/specific probe

Common method for analyzing chromatin structure

Experiment to distinguish euchromatin and heterochromatin:



12

The nucleosome can have two types of position:

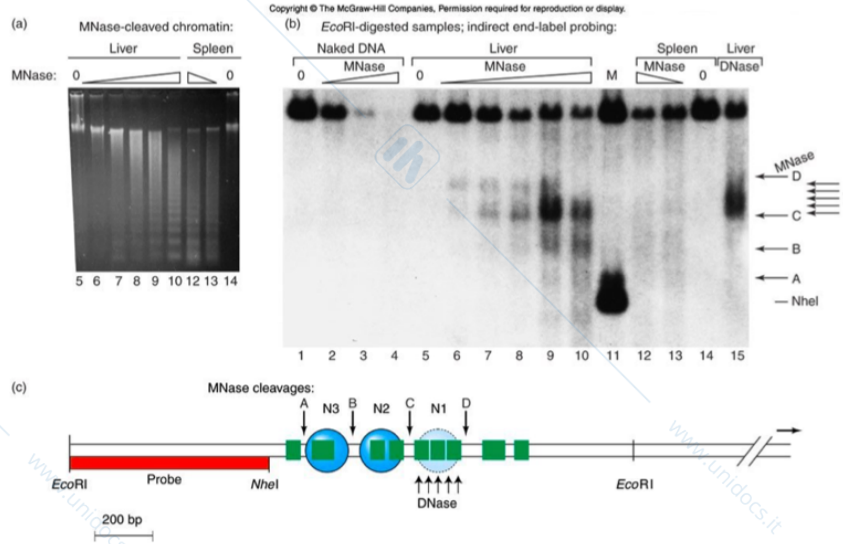
- **Translational position**
- **Rational position**

Translational position of nucleosomes:

Is the relative position to the primary sequence, if we move a nucleosome along the primary sequence then we are going to change the translational position of a nucleosome. The translational position can be study rough the activity of nuclease. Shown in the experiment below.

You can map the position of the nucleosome if you can identify the sites when the nuclease cuts the DNA (corresponding to linker DNA).

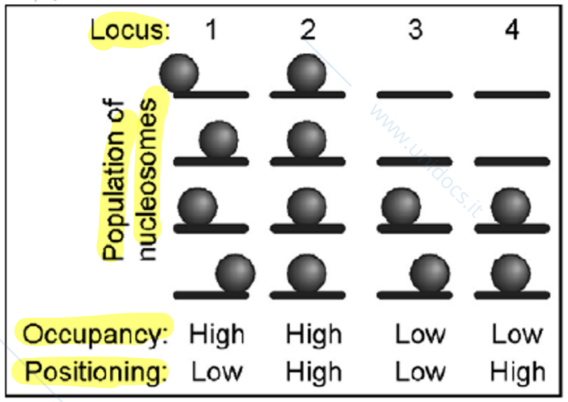
It's possible to do that with a **southern blot** .



The principle is that by using the enzyme activitiy of the **MNase** we will cut firstly t DNA and under this condition we can map the nulceosomal DNA by sequencing t cleavage. We are going to sequence the linker DNA because it is the part most exposed to the activity of the nuclease. We can do the map at a level of a single gene or genome wide.

Fragments defined by restriction enzyme's sites of cleavage. When we expose purified DNA to endonuclease activity it is completely degraded and on a gel we don't see any bands, this means that the DNA is cut at random site.

But with chromatin we see bands, this means that the endonuclease activity is not more random but at specific sites. The nucleosomes impede the activity of endonucleases at nucleosomal DNA (it is not bent at linker DNA).



The distance of the bands is 180-200 bp, meaning that this is the distance between linker DNA.

Rotational position of nucleosomes:

This position refers to the side of the duplex DNA that contacts the histone core.

The duplex will touch the histone octamer (yellow) but the contacts is made by only one side of the duplex while the other side is free. The free side can be contact by other proteins while the one attach to the histone core is inaccessible.

We can so rotate the duplex to change the contact.