MODULE 2: CONTROL OF GENE EXPRESSION AS THE BASIS OF METABOLIC REGULATION

LESSON 2. Regulation of transcription

Elena Ortiz Zapater

elena.ortiz-zapater@uv.es

Department of Biochemistry and Molecular Biology

VNIVERSITAT DE VALÈNCIA

Facultat de Medicina i Odontologia

Aquest material ha estat corregit pel Servei de Llengües i Política Lingüística en el marc dels Premis Fernando Sapiña 2024.

Premis Fernando Sapiña a materials docents en valencià i anglès

LESSON 2. Regulation of transcription

Learning objectives: upon completion of this lesson you will be able to...

Understand that transcriptional regulation is another way to achieve metabolic control in both prokaryotes and eukaryotes.

- Understand the complexity of transcription regulation in eukaryotes compared to prokaryotes.

- Understand the concept of operon and its organisation in metabolic pathways. E.g. the lac operon as an example of catabolic regulation by glucose and lactose.

- Examine the subunit composition of eukaryotic RNA polymerases and understand the initiation of transcription by RNA pol II, including general transcription factors and regulators that control the frequency of initiation.

- Recognise the importance of initiation in the regulation of transcription and analyse the common structural motifs of transcription factors and their relevance to gene expression.

LESSON 2. Regulation of transcription

Table of contents:

1. Regulation of gene expression: Introduction

- There are genes that code for essential products at all times, such as enzymes of central metabolic pathways. These genes are constantly expressed in all cells and are referred to as *housekeeping genes*. The constant expression of a gene is called *constitutive gene expression*.
- Other genes have levels that vary in response to molecular signals, which is referred to as the ***regulation of gene expression***.
- Transcription is mediated and regulated by protein-DNA interactions, especially with RNA polymerase components.
- RNA polymerase initiates transcription at promoters, whose sequences affect the frequency of initiation.
- In eukaryotes, general transcription factors and proteins such as TBP act as specificity factors.
- Activators in eukaryotes can bind to distant sites called enhancers, affecting long-distance transcription.
- The distance between activators/repressors and promoters is bridged by the DNA loop. The interaction between activators and RNA polymerase is often mediated by coactivators.

1. Regulation of gene expression: Introduction

Prokaryotes	**Eukaryotes**
Single chromosome	Multiple chromosomes (23 pairs in humans)
Organised genes in operons	Scattered genes across the genome
Polycistronic mRNAs	Not common
Transcription/Translation coupled	Transcription/Translation separately
Naked DNA	Protein-associated DNA (chromatin)
Single cell type	Multiple specialised cell types

Regulation of transcription

Can you name any transcription factors in eukaryotes?

What is the *lac* Z operon?

2. Operons in prokaryotes

- In prokaryotes, genes for enzymes of the same metabolic pathway are clustered together to form **operons**.
- Operons enable the coordinated expression of their genes.
- An operon consists of a **regulatory gene**, a **control centre** (operator DNA + promoter DNA) and a set of **structural genes**.
- The regulatory gene encodes a protein (**repressor or activator**) that interacts with the control centre operator to inhibit or stimulate transcription.

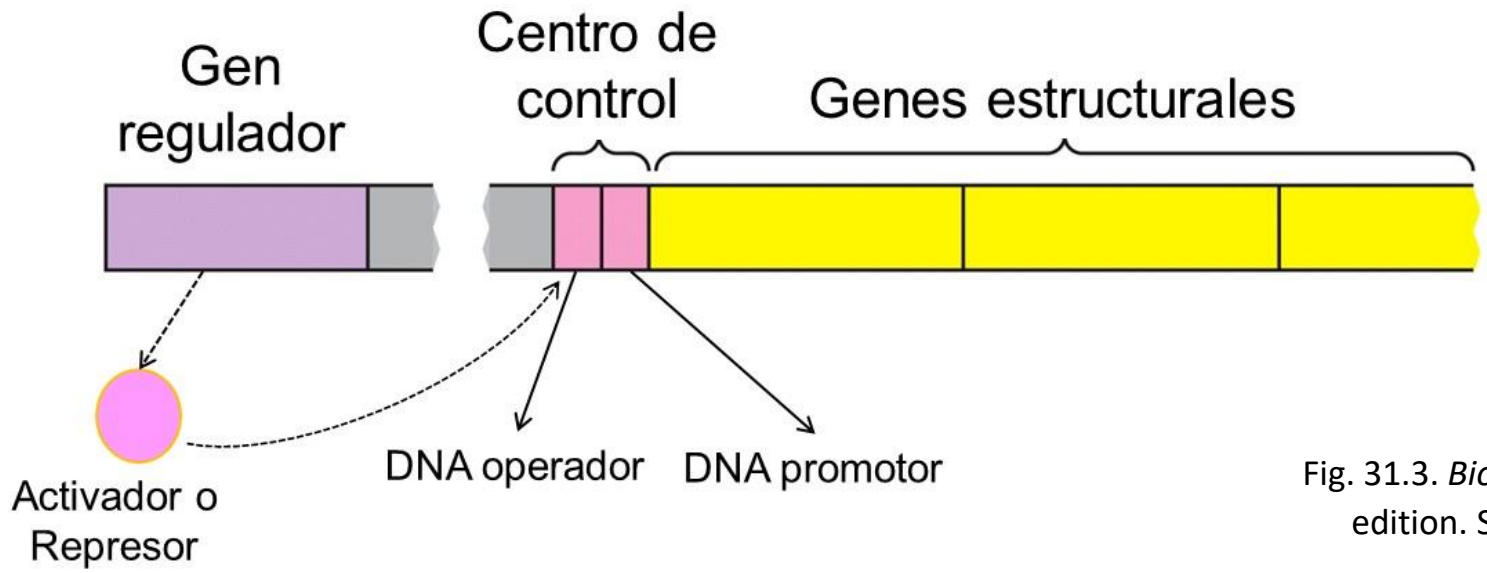


Fig. 31.3. *Biochemistry*, 4th edition. Stryer et al.

2. Operons in prokaryotes

2.1. The lactose operon

- The lactose operon is a catabolic operon. Bacteria such as *E. coli* depend on glucose; but when glucose is scarce, they can adapt to another carbon and energy source such as lactose.
- To use lactose as an energy source, bacteria must express the enzyme **β-galactosidase** and two additional proteins encoded in the lactose operon.
- These enzymes are rapidly induced (about 1000-fold) when the bacteria switch from growing in the presence of glucose to growing in the presence of lactose.

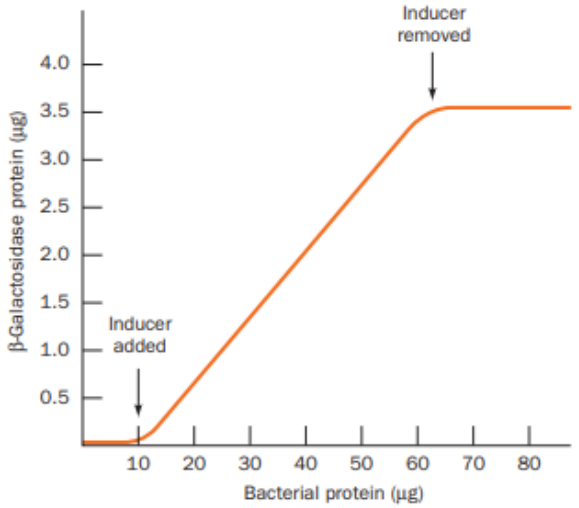
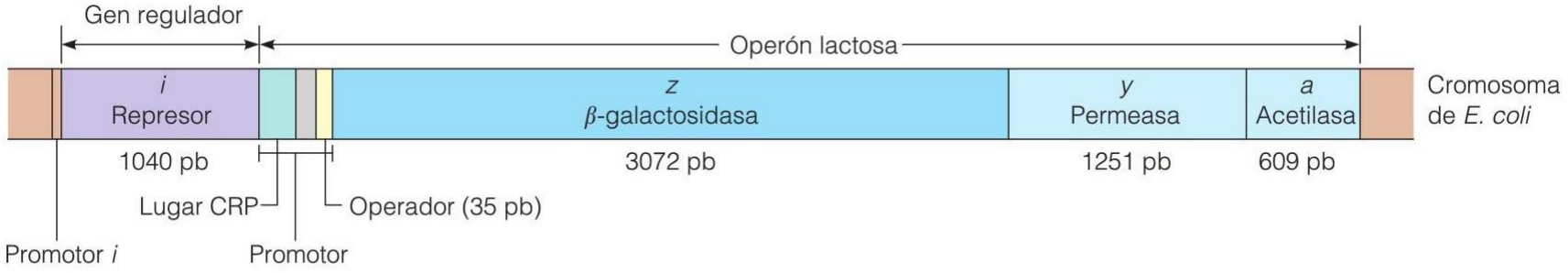


Figure 31-1 **The induction kinetics of β-galactosidase in *E. coli*.** [After Cohn, M., *Bacteriol. Rev.* **21,** 156 (1957).]

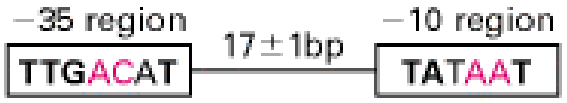


Biochemistry, 4th edition. Matthews et al.

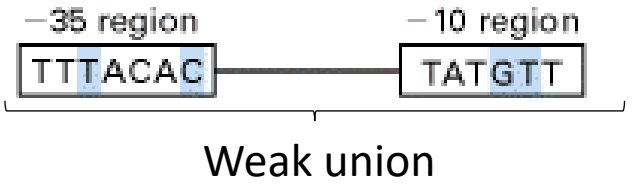
2. Operons in prokaryotes

2.1. The lactose operon

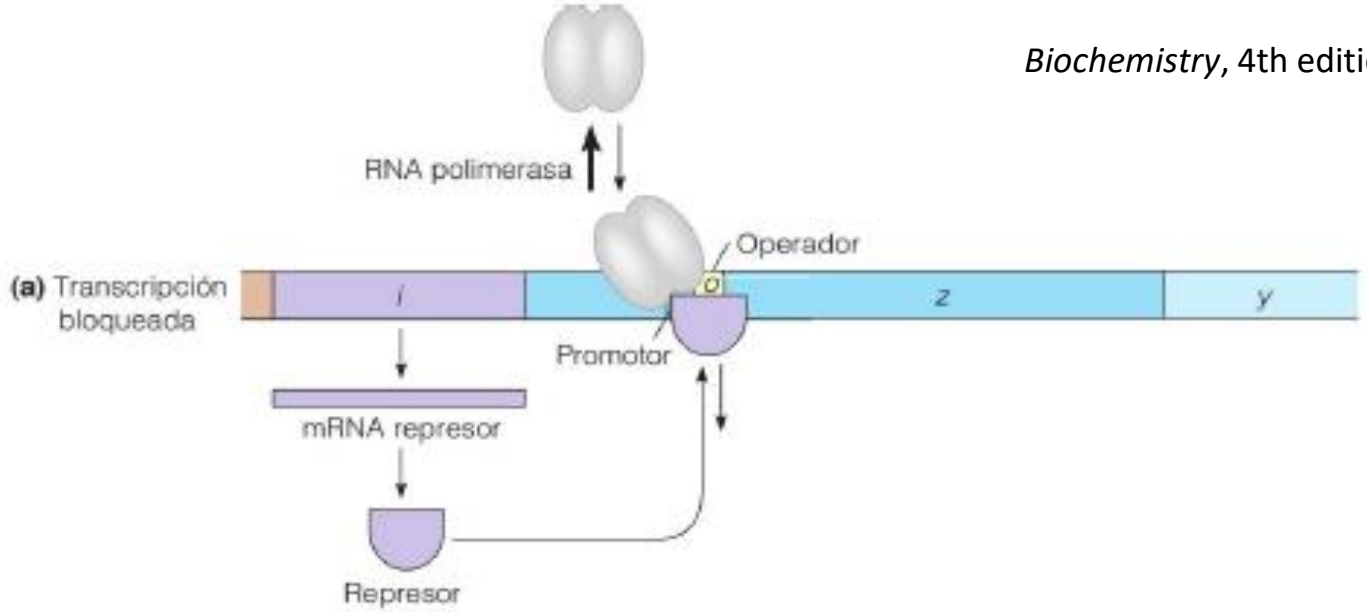
Consensus sequence of prokaryotic promoters



Lac promoter sequence



In the **absence of lactose**, the ***lac* repressor protein** binds to the DNA operator in dimeric form, and blocks transcription. Prokaryotic RNApol is hindered by the *lac* repressor.



Biochemistry, 4th edition. Matthews et al.

2. Operons in prokaryotes

2.2. Regulation of the lactose operon (repression)

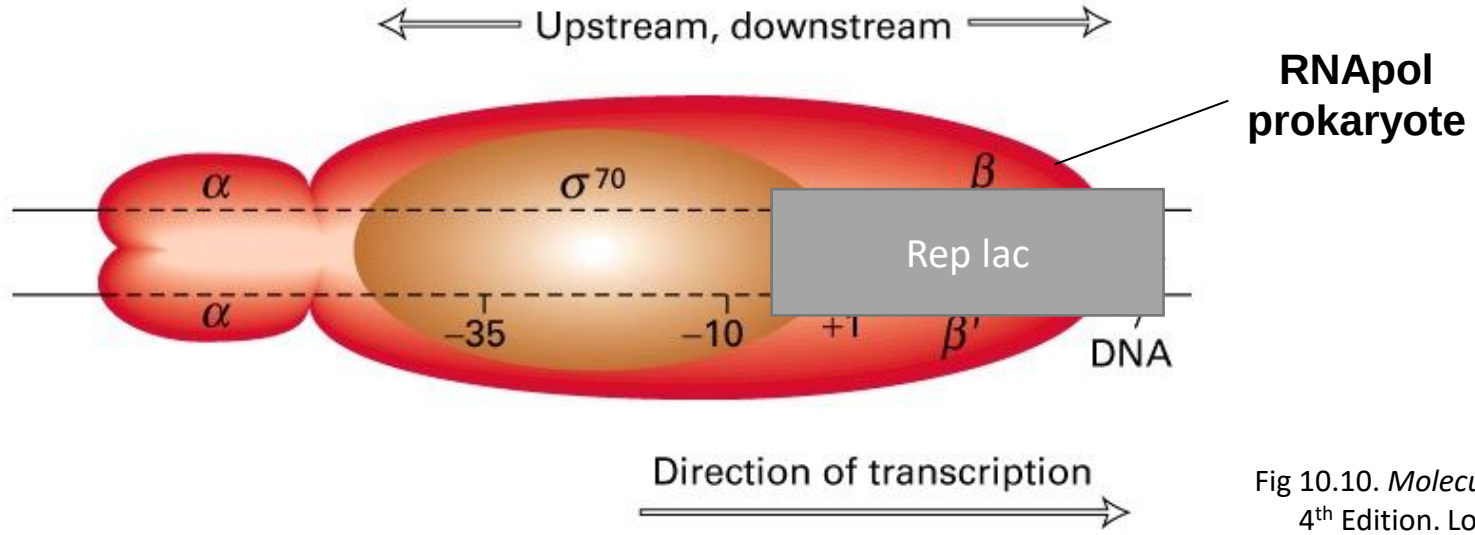


Fig 10.10. *Molecular Cell Biology*. 4th Edition. Lodish, H et al.

2. Operons in prokaryotes

2.2. Regulation of the lactose operon (transcription)

In the **presence of lactose**, the **inducer allolactose** (a lactose derivative) is generated, which binds to the *lac* repressor and causes it to change conformation, causing it to cleave the DNA. RNA polymerase can then move along the operon and transcribe the structural genes.

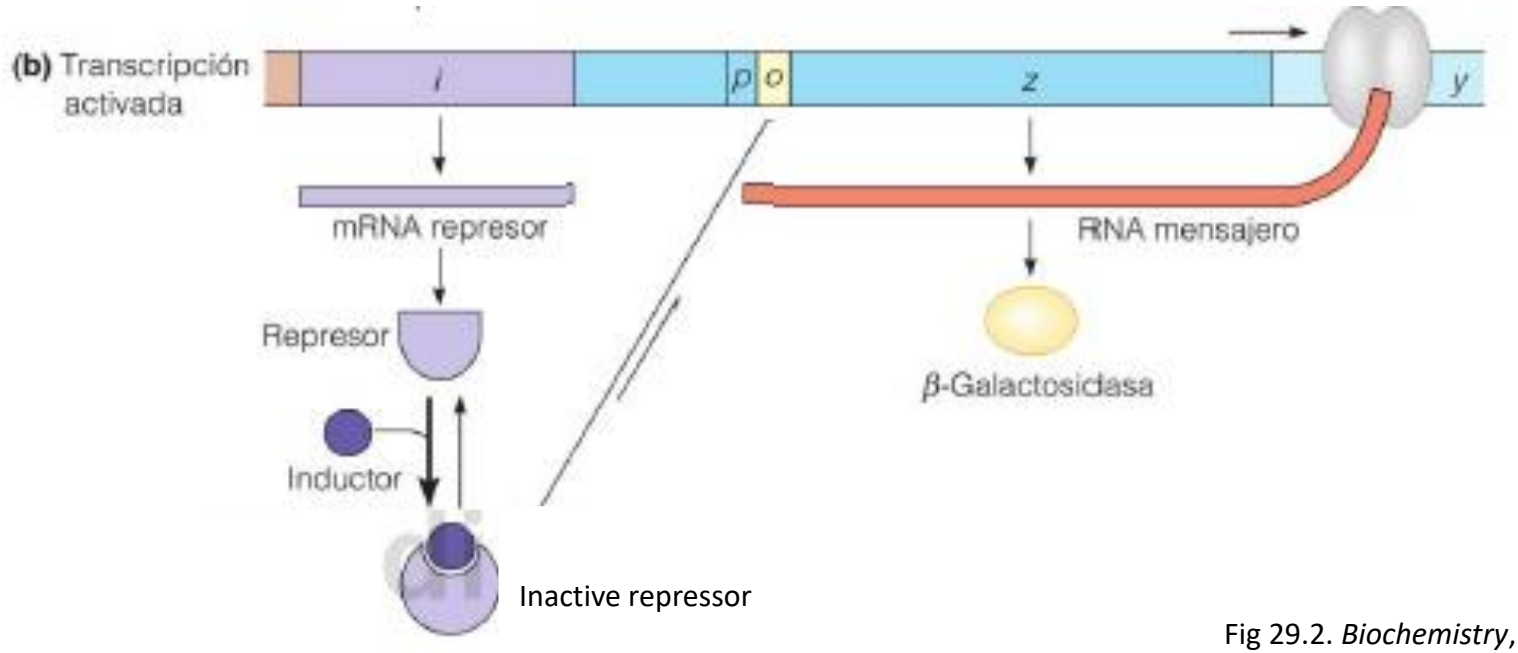
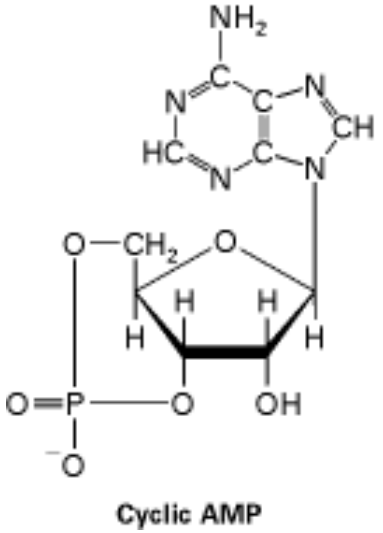


Fig 29.2. *Biochemistry*, 4th edition. Matthews et al.

2. Operons in prokaryotes

2.2. Regulation of the lactose operon (activation)



Cyclic AMP

- **Cyclic AMP** (**cAMP**), a "hunger/fasting" signal, stimulates transcription of many catabolic operons by binding to **catabolite activator protein** (**CAP**).

- Binding of the cAMP-CAP complex to a DNA operator of the catabolic operon increases the affinity of RNA polymerase for the promoter DNA and transcriptional activity.

Control of the *lac* operon by cAMP: a hunger signal

(a) If glucose is abundant, the enzyme adenylate cyclase **does not** convert ATP to cAMP.

(b) If glucose is scarce, adenylate cyclase is active and cAMP is formed.

(c) When cAMP is present, it binds to the CAP protein and confers affinity for DNA.

(d) The cAMP-CAP complex binds next to the promoter and activates transcription of the *lac* operon.

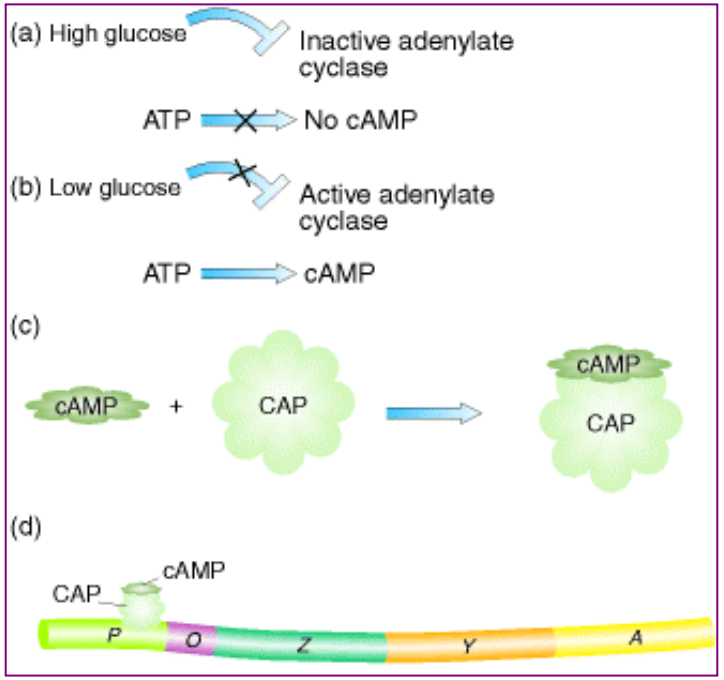


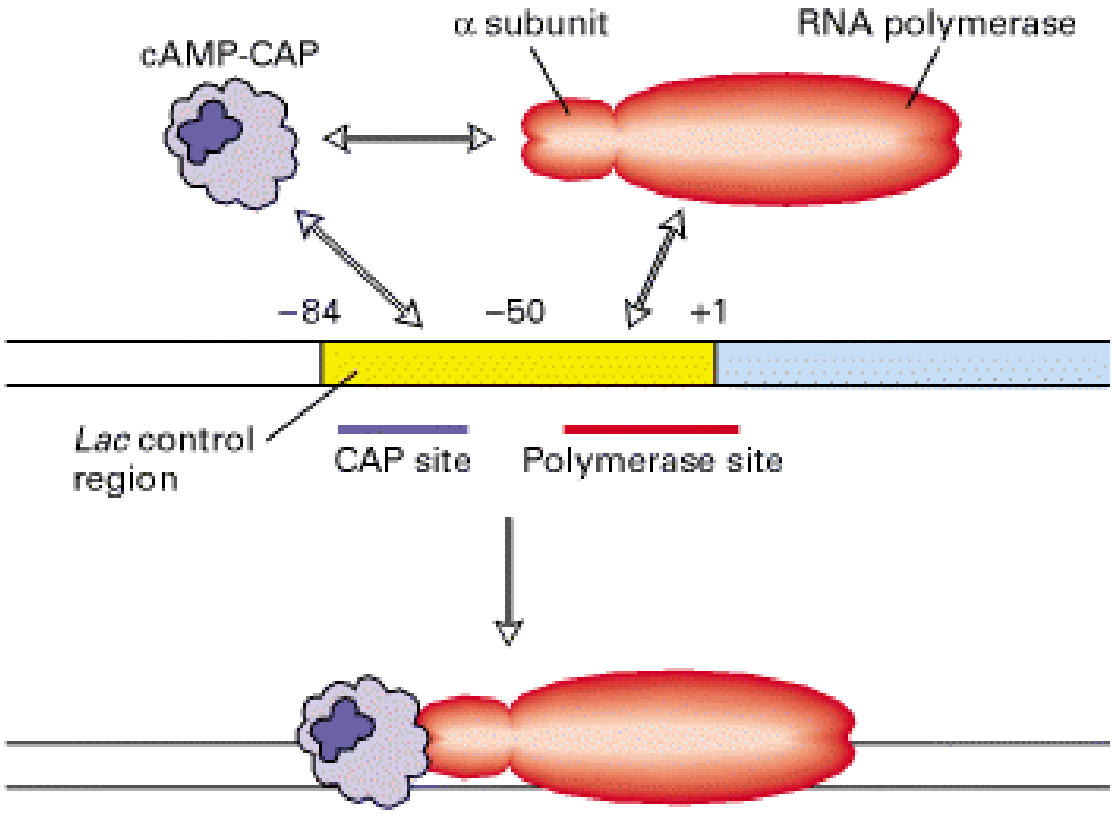
Fig 31.10. *Biochemistry*, 4th edition. Stryer et al.

2. Operons in prokaryotes

2.2. Regulation of the lactose operon (activation in the absence of glucose)

Why does RNA polymerase need help to activate the *lac* operon?

The cAMP-CAP complex cooperates with RNApol to bind to the promoter DNA of the **"weak"** *lac* operon.



2. Operons in prokaryotes

2.2. Regulation of the lactose operon (summary)

Diagram of the transcriptional control centre of the *lac* operon with the promoter DNA and the operator DNAs.

The activator cAMP-CAP facilitates the binding of RNA pol to promoter DNA and enhances transcription initiation.

The *lac* repressor hinders the binding of RNA pol to the promoter DNA and prevents transcription from starting.

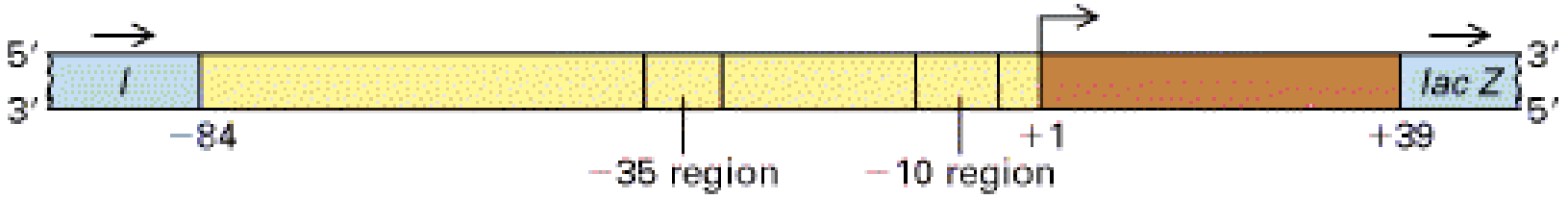
Adjacent in 5'.
Facilitates RNA pol binding

cAMP-CAP

Overlapping in 3'.
Impedes RNA pol binding

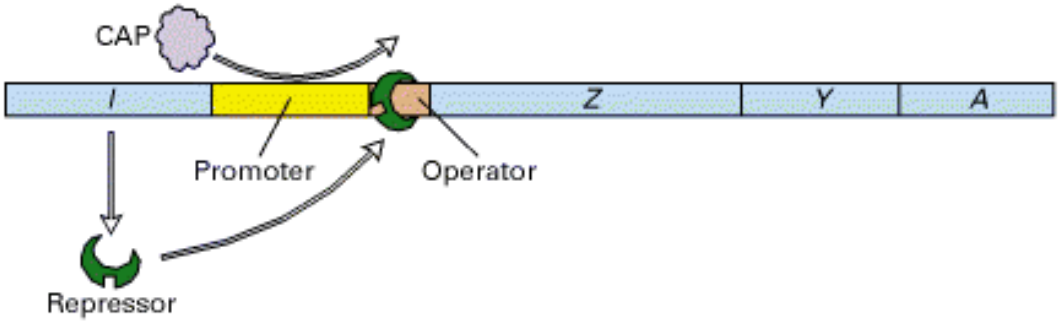
Repressor

RNA polymerase

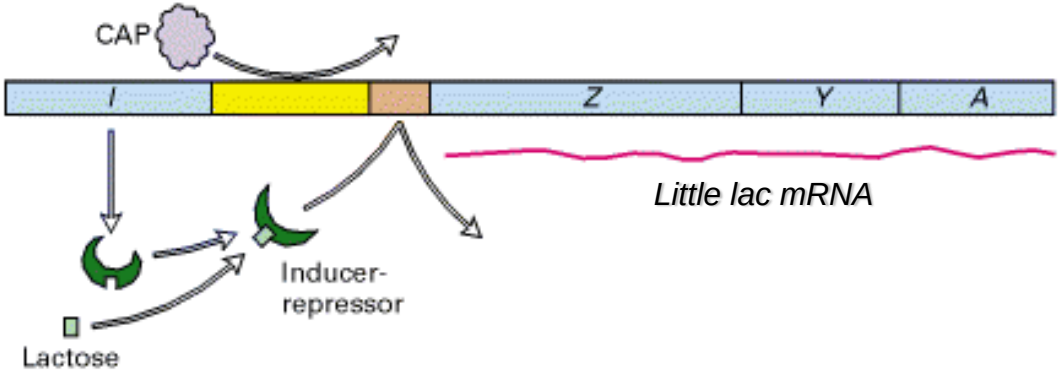


Co-ordinated regulation of the *lac* operon by its *lac* **repressor** and by the CAP **activator**

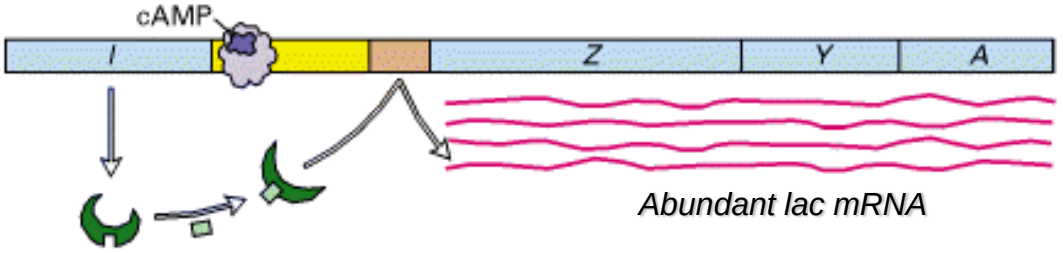
A) Glucose present (low cAMP level). No lactose.



B) Glucose present (low cAMP level). Lactose present.



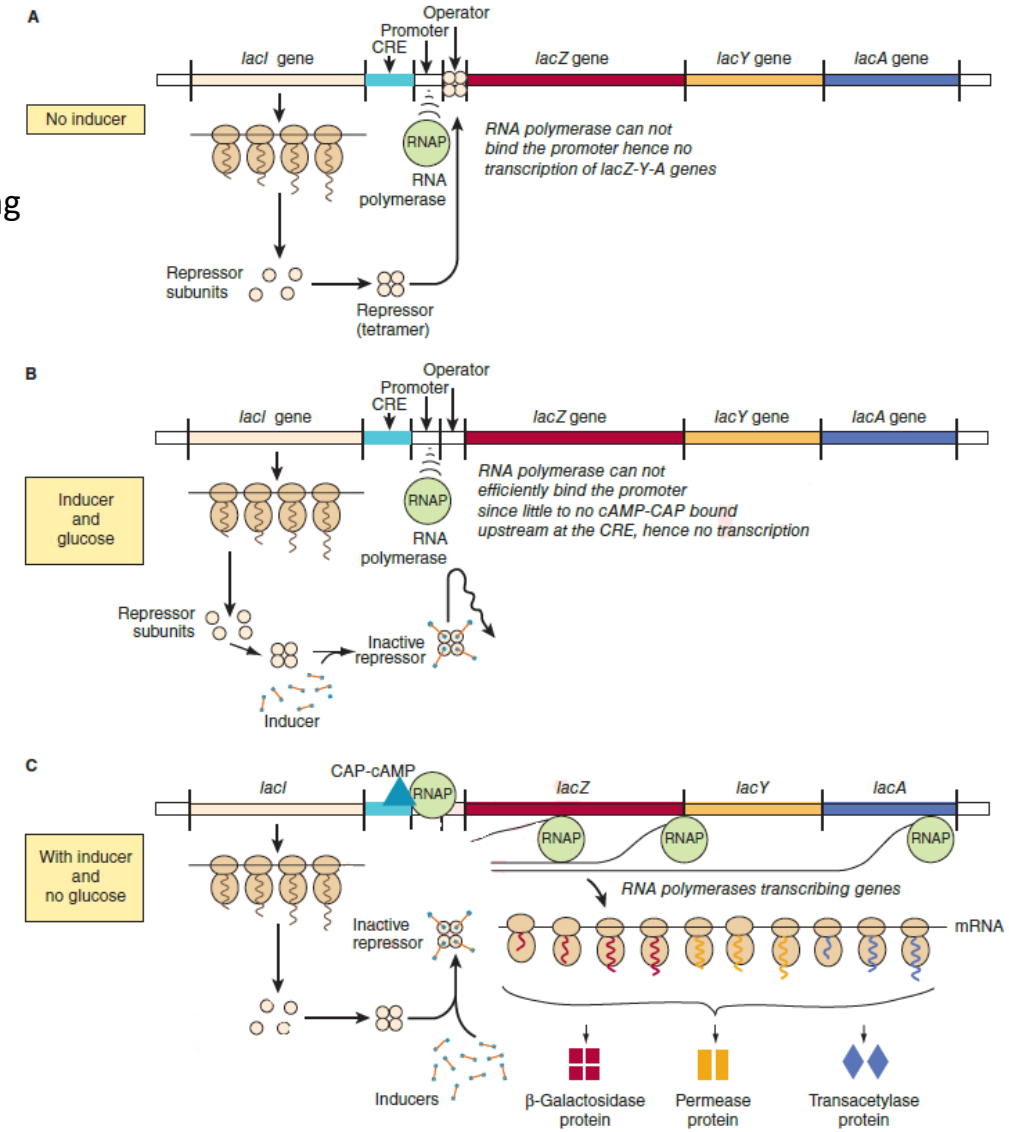
C) No glucose (high cAMP level). Lactose present.



2. Operons in prokaryotes

2.2. Regulation of the lactose operon (summary)

CRE = CAP protein-binding site

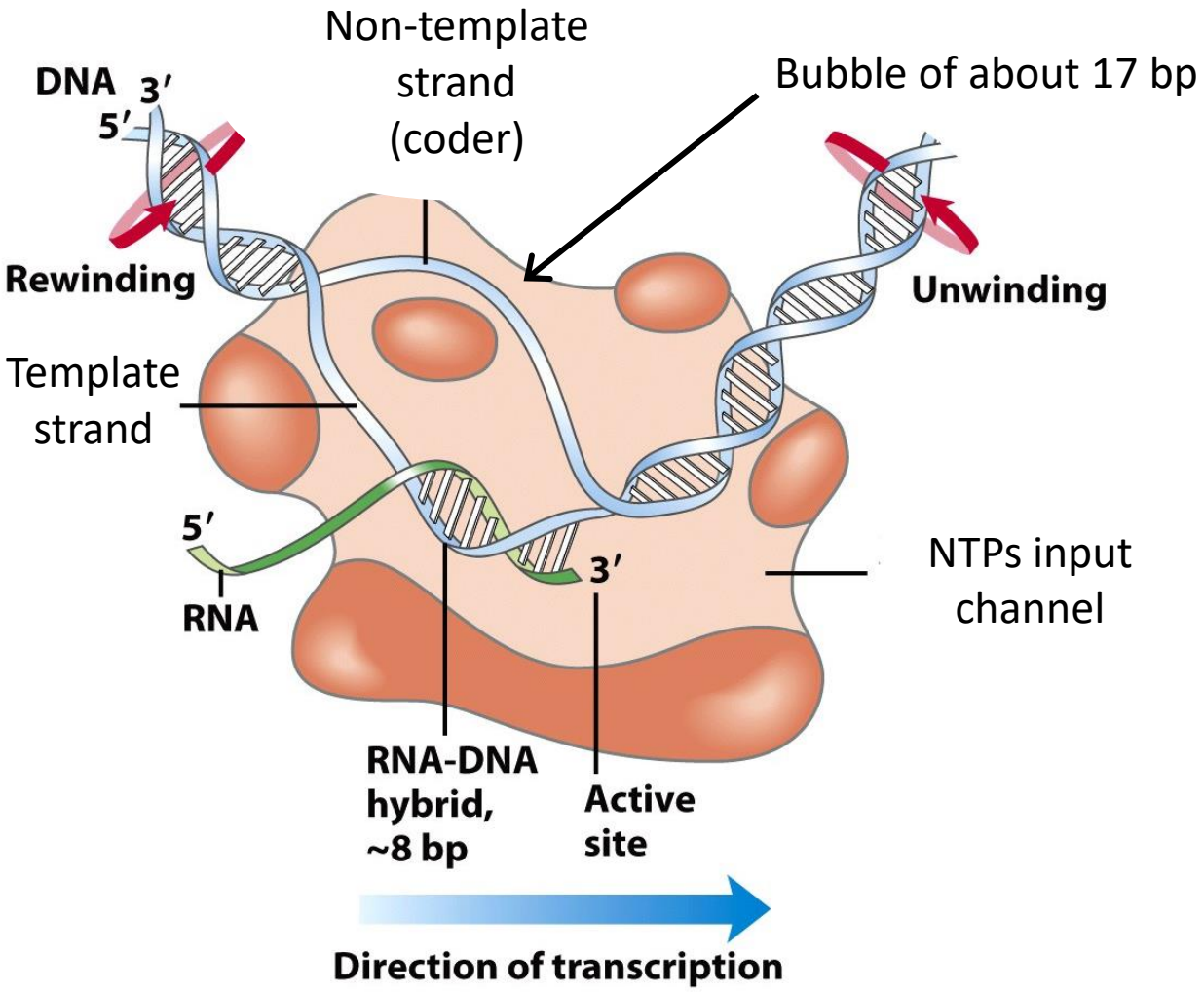


Source: *Haper's Illustrated Biochemistry*, 30th edition

3. Regulation of transcription in eukaryotes

The expression and regulation of eukaryotic genes is **more complex** than that of prokaryotes.

Transcription by RNA polymerase: a **simplified view**



3. Regulation of transcription in eukaryotes

Transcription by RNA polymerase in eukaryotes: a **more realistic picture**

It is a much more complex process with many regulatory mechanisms. The promoter regions of eukaryotic genes are much MORE EXTENSIVE and with MORE CONTROL ELEMENTS, in which we can distinguish:

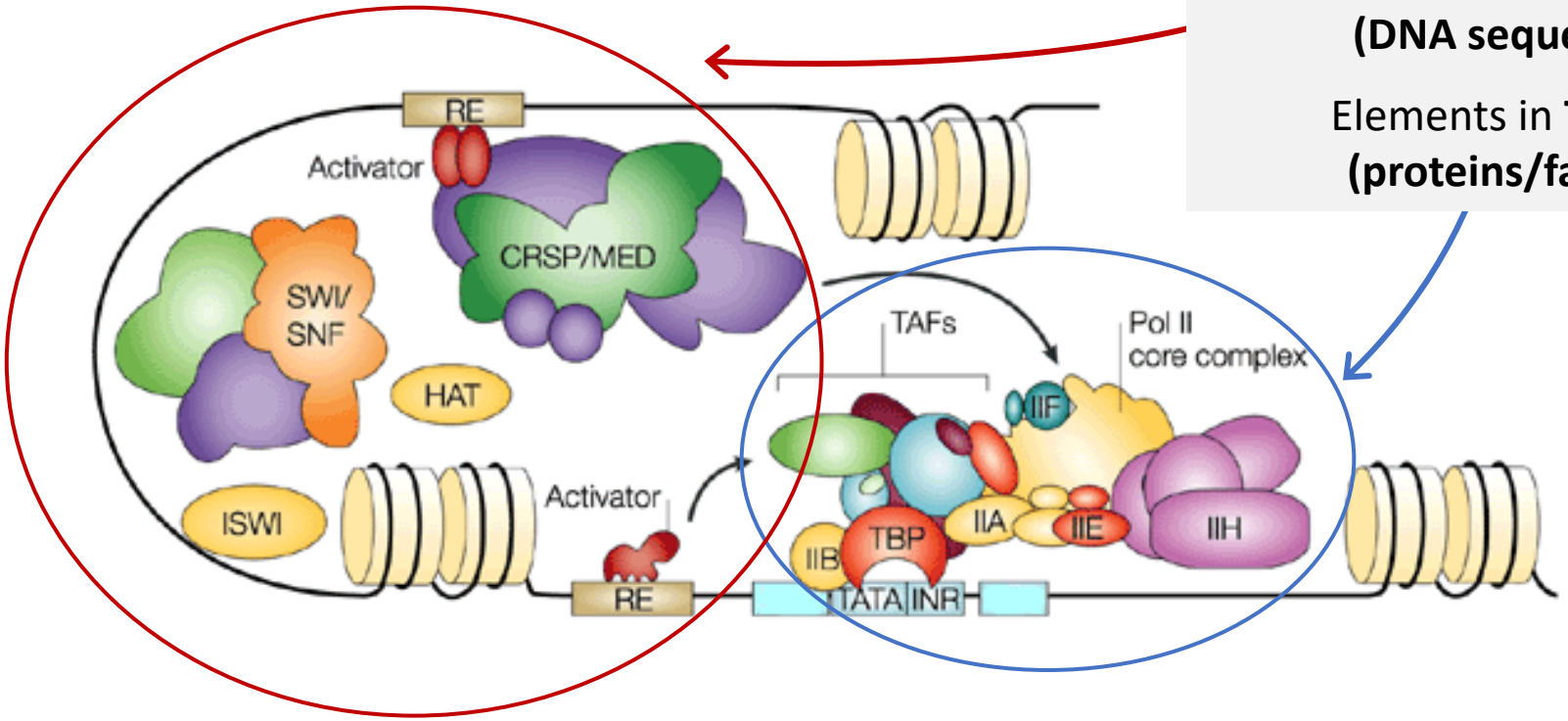
a) Basic mechanism of transcription: basal factors

b) Control and regulatory mechanisms

It requires the interaction of

Elements in **CIS:**
(DNA sequences)

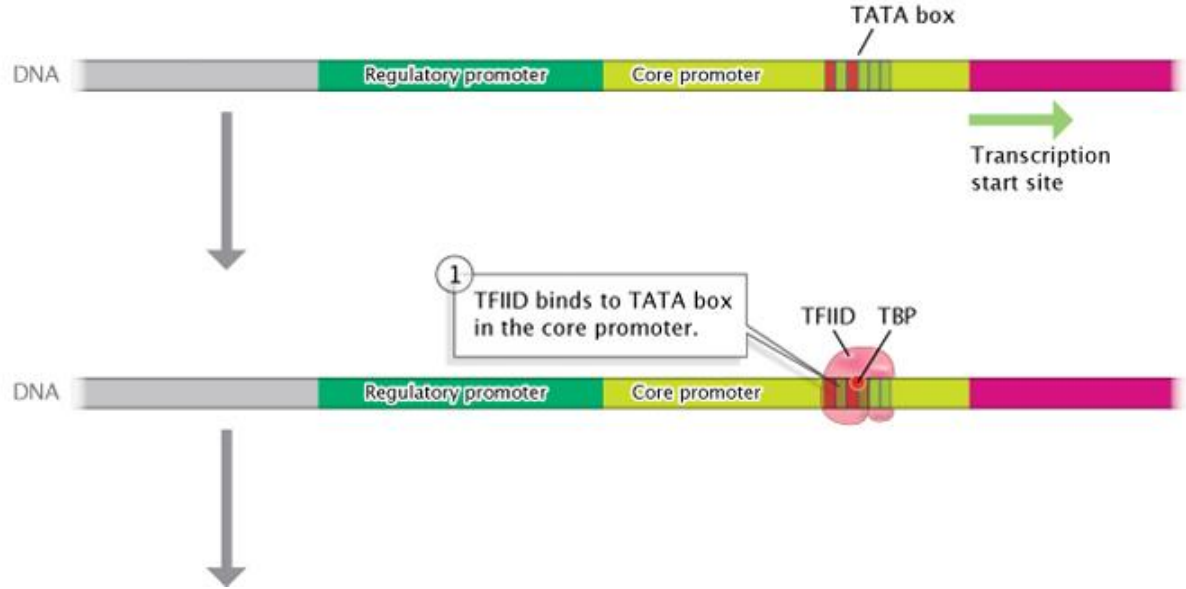
Elements in **TRANS:**
(proteins/factors)



3. Regulation of transcription in eukaryotes

3.1. Transcription initiation: RNA polymerase II and general transcription factors

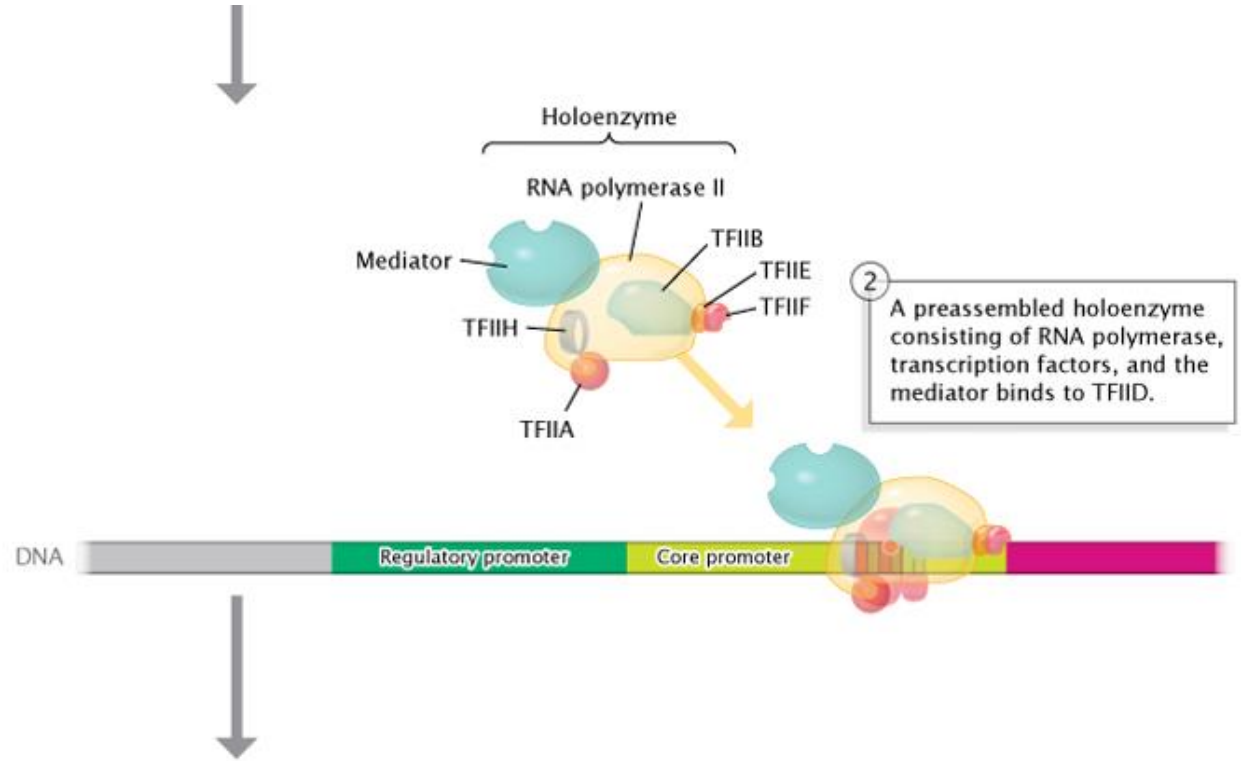
- RNA polymerase II (RNA Pol II): The enzyme responsible for the synthesis of mRNA from DNA in eukaryotes.
- Transcription begins at specific regions of DNA known as promoters, which usually contain the TATA box sequence (at position -25 to -35). These promoters are essential for the assembly of the transcription machinery.
- General transcription factors (GTFs): These are proteins that help RNA Pol II bind to the promoter and form the pre-initiation complex (PIC). The main GTFs include: TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH.
- Formation of the pre-initiation complex (PIC): The initiation process begins when TFIID, which contains the TBP (TATA-binding protein) subunit, binds to the TATA box, allowing the recruitment of other GTFs and RNA Pol II, forming the pre-initiation complex.



3. Regulation of transcription in eukaryotes

3.1. Transcription initiation: RNA polymerase II and general transcription factors

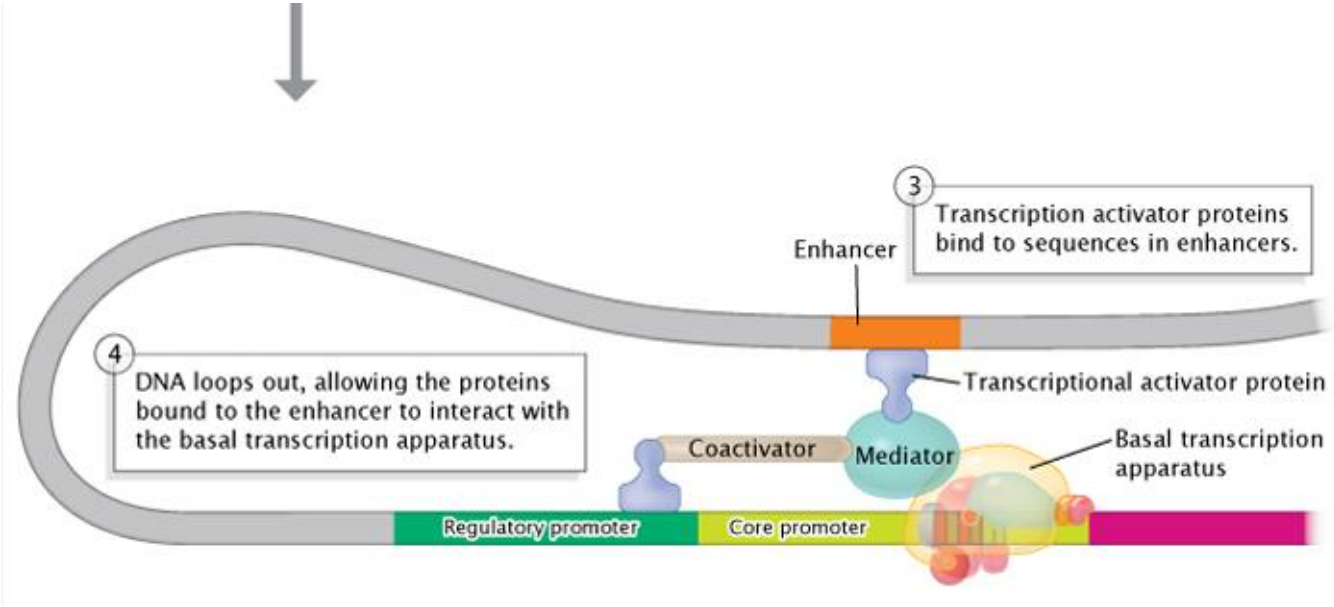
- Function of TFIIH: This factor is essential for transcription initiation, as it has helicase activity that unwinds DNA and kinase activity that phosphorylates the C-terminal end of RNA Pol II, triggering the initiation of RNA synthesis.
- Phosphorylation of the C-terminal tail (CTD) of RNA Pol II: CTD phosphorylation is a critical step that marks the transition from the initiation to the elongation phase of transcription, allowing RNA Pol II to move along the DNA.



3. Regulation of transcription in eukaryotes

3.1. Transcription initiation: RNA polymerase II and general transcription factors

- Initiation and escape from the promoter: Once RNA Pol II has initiated RNA synthesis, it must overcome a series of initial blocks to escape from the promoter and enter the elongation phase, which involves the release of some transcription factors.
- Control of gene initiation and expression: Specific transcription factors and coactivators regulate the efficiency with which the pre-initiation complex is formed, modulating the rate of gene transcription and thus gene expression in different cellular contexts.



3. Regulation of transcription in eukaryotes

3.2. DNA sequences important for eukaryotic gene transcription

1. The promoter

Promoters are DNA sequences that determine the binding site of RNA polymerase and the initiation of transcription.

In the promoters of eukaryotic genes, **several types of consensus DNA sequences** have been identified

1. The TATA box
2. The initiator element

} are very frequent in tissue-specific genes.

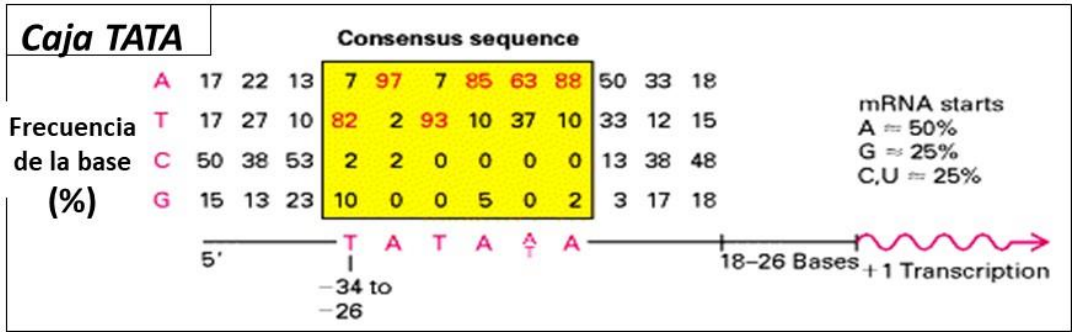
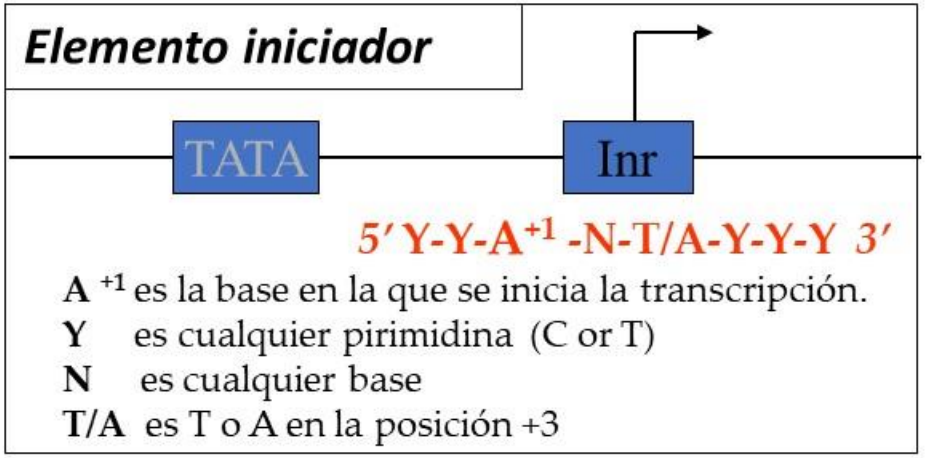


Fig 11.9. *Molecular Cell Biology*. 5th Edition. Lodish, H et al.



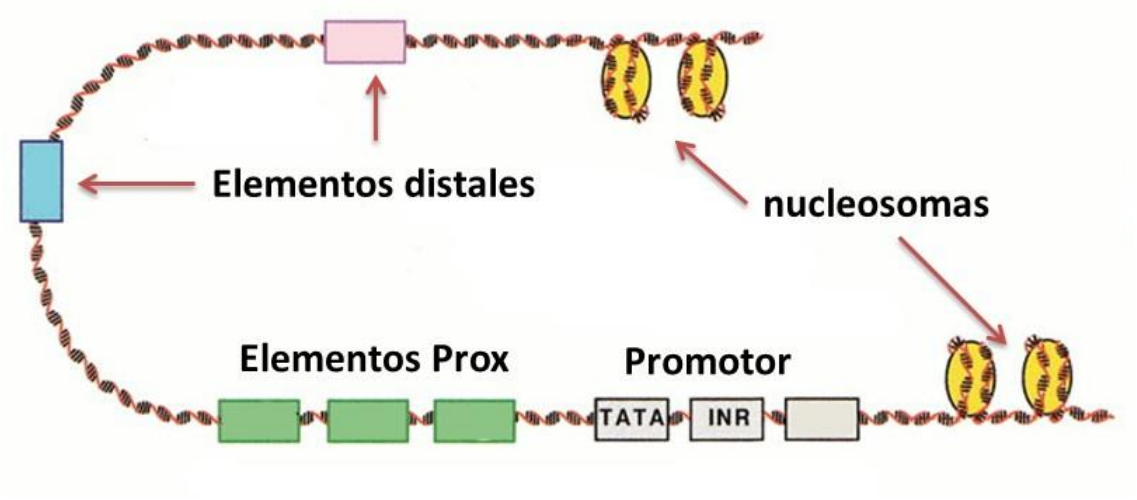
3. Regulation of transcription in eukaryotes

3.2. DNA sequences important for eukaryotic gene transcription

2. Proximal and distal control elements (DNA sequences)

In eukaryotes, transcription of protein-coding genes is regulated through multiple DNA sequences called transcription **control elements**.

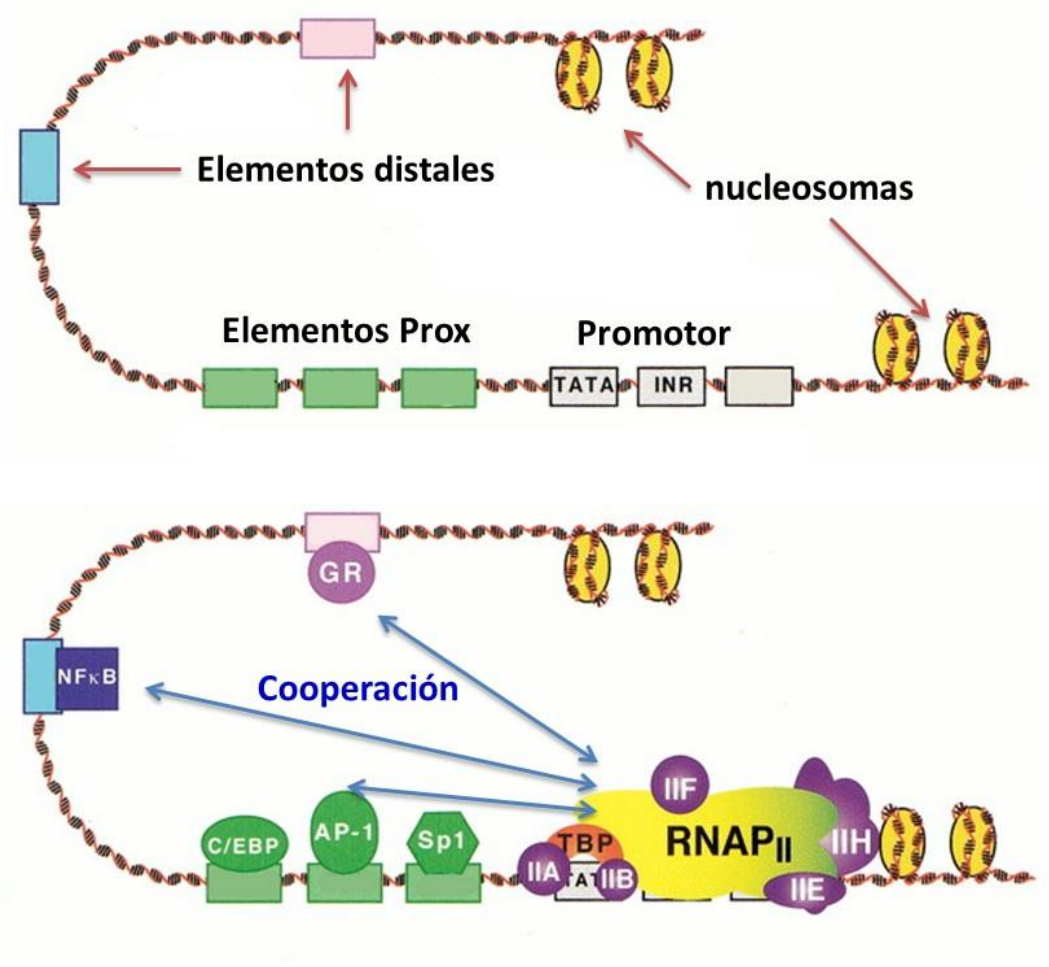
- **Proximal regulatory sequences** (8 to 20 bp) are located no more than 200 bp from the start site.
- The **distal amplifying sequences** contain multiple control elements (8-20 bp each element) clustered in ~200 bp in total. They may be located up to 50 kbp 5' or 3' from the start site, or in an intron.
- In general, several proximal and distal control elements contribute to the regulation of each particular gene.

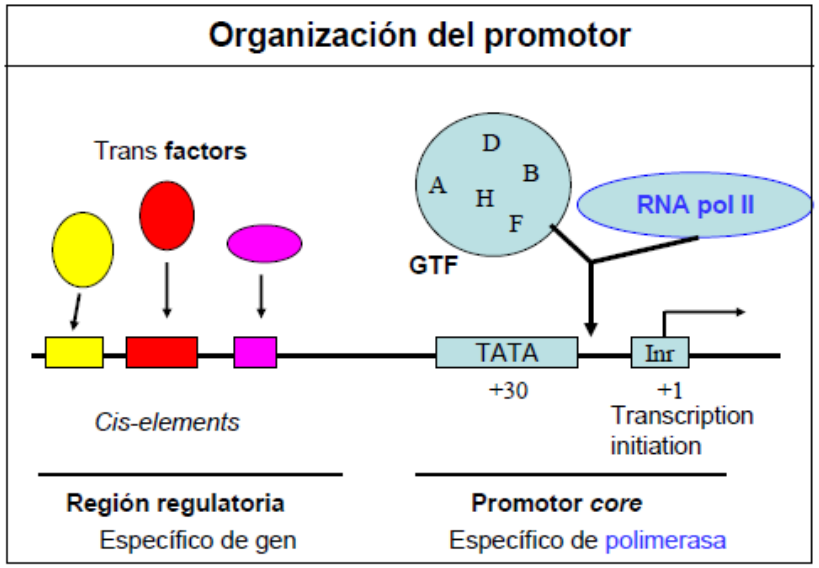


3. Regulation of transcription in eukaryotes

3.2. DNA sequences important for eukaryotic gene transcription

Multiple activators attached to control elements cooperate to attract and bind RNA pol to the promoter and facilitate transcription initiation.

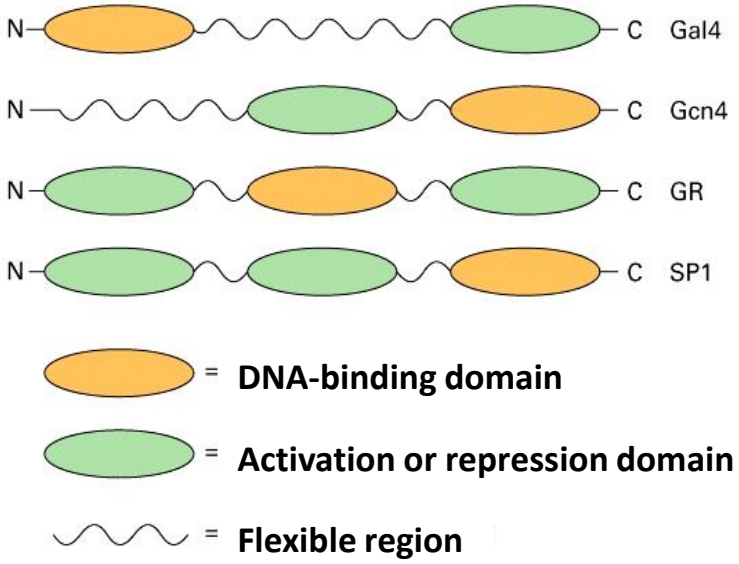




3. Regulation of transcription in eukaryotes

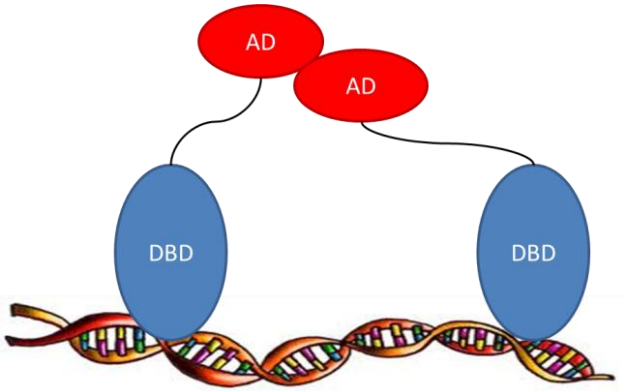
3.3. DNA response elements and transcription factors

Transcription factors are usually modular proteins containing a single **DNA-binding domain** and one or more **activation or repression domains**; the different domains are often linked through flexible polypeptide regions.



The flexible regions allow the interaction of the activation domains of different factors, even when their DNA-binding domains are far apart.

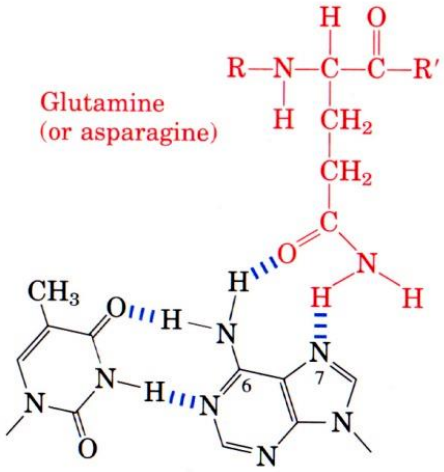
→ **Another form of regulation!**



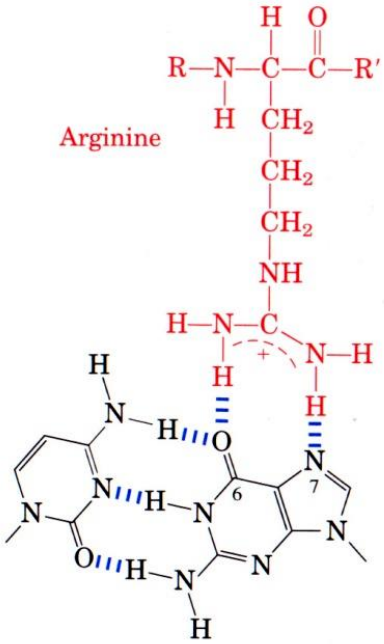
3. Regulation of transcription in eukaryotes

3.4. DNA-binding domains

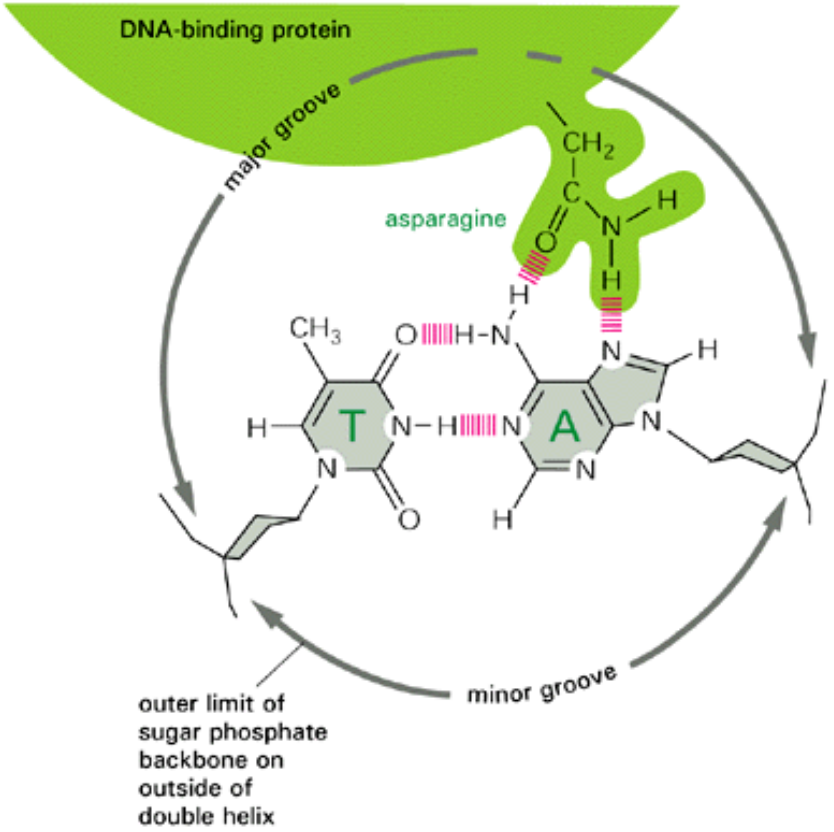
Recognition and binding to DNA by a transcription factor: specificity in recognition is achieved by establishing a protein-DNA interface where 10-20 weak contacts are established between different amino acids and nitrogenous bases. Each weak interaction contributes to a high total binding energy.



Thymine: Adenine



Cytosine: Guanine

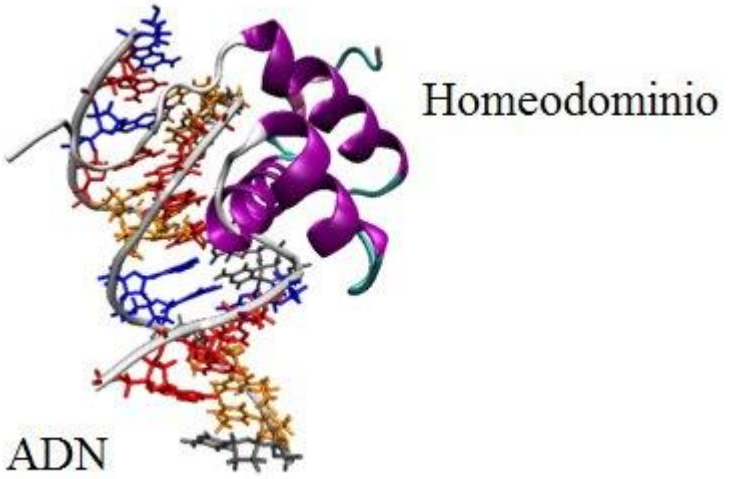


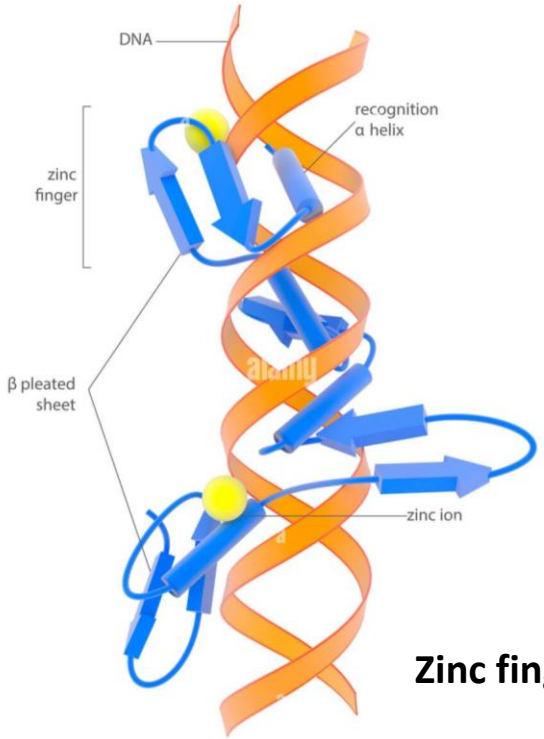
3. Regulation of transcription in eukaryotes

3.5. Common structural motifs of transcription factors

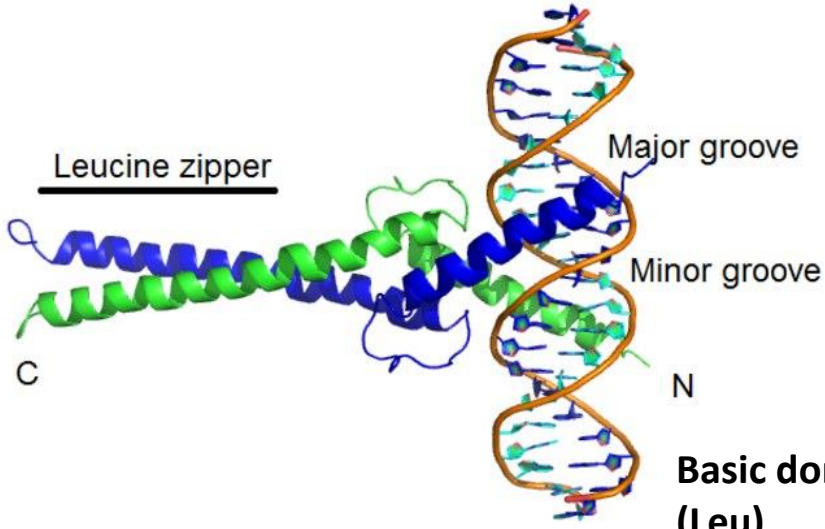
The **DNA-binding domains** have characteristic structures:

- The zinc finger
- Homeodomain
- The basic domain + hydrophobic zip (Leu)
- The winged propeller
- The helix-turn-helix motif





Zinc finger



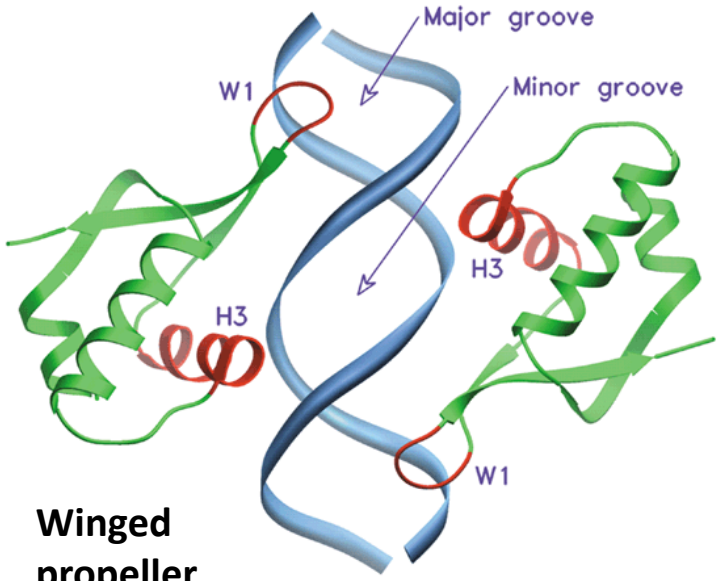
Basic domain + hydrophobic zip (Leu)

3. Regulation of transcription in eukaryotes

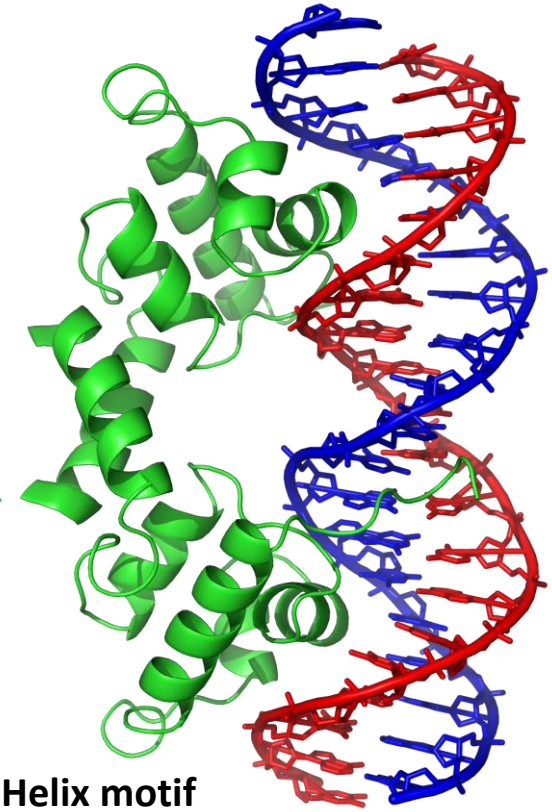
3.5. Common structural motifs of transcription factors

The **DNA-binding domains** have characteristic structures:

- The zinc finger
- Homeodomain
- The basic domain + hydrophobic zip (Leu)
- The winged propeller
- The helix-turn-helix motif



Winged propeller



Helix-Turn-Helix motif

The **activation domains** are structurally very diverse:

- Acid activation domains (Asp & Glu) (GR)
- Glutamine-rich activation domains (SP1)
- Proline-rich activation domains (c-jun)
- Serine- and threonine-rich activation domains (-OH)
- Ligand-dependent activation domains

4. Diseases associated with transcription failures (mutations)

Diseases associated with mutations in **DNA regulatory elements**			
Regulatory element	**Disease**	**Mutation (bound factor)**	**Affected gene**
Core promoter	β-thalassemia	TATA box, CACCC box, DCE	*β-globin*
Proximal promoter	Bernard-Soulier Syndrome	133 bp upstream of TSS (GATA-1)	GpIbβ
	Charcot-Marie-Tooth disease	215 bp upstream of TSS	*connexin-32*
	Congenital erythropoietic porphyria	70, 90 bp upstream of TSS (GATA-1, CP2)	uroporphyrinogen III synthase
	Familial hypercholesterolemia	43 bp upstream of TSS (Sp1)	*low density lipoprotein receptor*
	Familial combined hyperlipidaemia	39 bp upstream of TSS (Oct-1)	*lipoprotein lipase*
	Haemophilia	CCAAT box (C/EBP)	*factor IX*
	Hereditary persistence of foetal haemoglobin	~175 bp upstream of TSS (Oct-1, GATA-1)	*Aγ-globin*
	Progressive myoclonus epilepsy	Expansion ~70 bp upstream of TSS	*cystatin B*
	Pyruvate kinase deficiency anaemia	72 bp upstream of TSS (GATA-1)	*PKLR*
	β-thalassemia	CACCC box (EKLF)	*β-globin*
	δ-thalassemia	77 bp upstream of TSS (GATA-1)	*δ-globin*
	Treacher-Collins syndrome	346 bp upstream of TSS (YY1)	*TCOF1*
Enhancer	Preaxial polydactyly	1 Mb upstream of gene	*SHH*
	Van Buchem disease	Deletion ~35 kb downstream of gene	*sclerostin*
	X-linked deafness	Microdeletions 900 kb upstream	*POU3F4*

4. Diseases associated with transcription failures (mutations)

Diseases associated with mutations in **general transcription factors and regulators**		
Component	**Disease**	**Mutated factor**
General transcription factors	Xeroderma pigmentosum, Cockayne syndrome, trichothiodystrophy	TFIIH
Activators	Aniridia	PAX6
	Campomelic dysplasia	SOX9
	Congenital central hypoventilation syndrome	PHOX2B
	Congenital heart disease	Nkx2-5
	Down syndrome with acute megakaryoblastic leukaemia	GATA-1
	Nail-patella syndrome	LMX1B
	Prostate cancer	ATBF1
	X-linked deafness	POU3F4
	X-linked dyserythropoietic anaemia and thrombocytopenia	GATA-1
	X-linked thrombocytopenia	GATA-1
Repressors	X-linked autoimmunity-allergic dysregulation syndrome	FOXP3
Coactivators	Parkinson's disease	DJ-1
	Type II diabetes mellitus	PGC-1
Chromatin-remodelling factors	Cancer	BRG1/BRM
	Retinal degeneration	ataxin-7
	Rett syndrome	MeCP2
	Rubinstein-Taybi syndrome	CREB-binding protein
	α-thalassemia myelodysplasia syndrome	ATRX

5. Take-home messages

- There are many differences between transcription in prokaryotes and eukaryotes (much more complex).
- You must know the structure and regulation of the lactose catabolic operon (*lac*) and the organisation of the prokaryotic operon and its advantages.
- The *lac* operon is under dual regulation: positive and negative. In the first case, ligand binding (cAMP) facilitates and in the second case, ligand binding (allolactose) hinders the binding of the positive (CAP) and negative (*lac* repressor) regulatory protein, respectively.
- You must know the two *trans* elements involved in mRNA transcription: RNA polymerase II, general transcription factors and regulatory transcription factors.
- You must know the *cis* elements involved in mRNA transcription: promoter, proximal regulatory elements and distal regulatory elements.
- You must know the types and structure and mechanism of action of the main regulatory transcription factors.