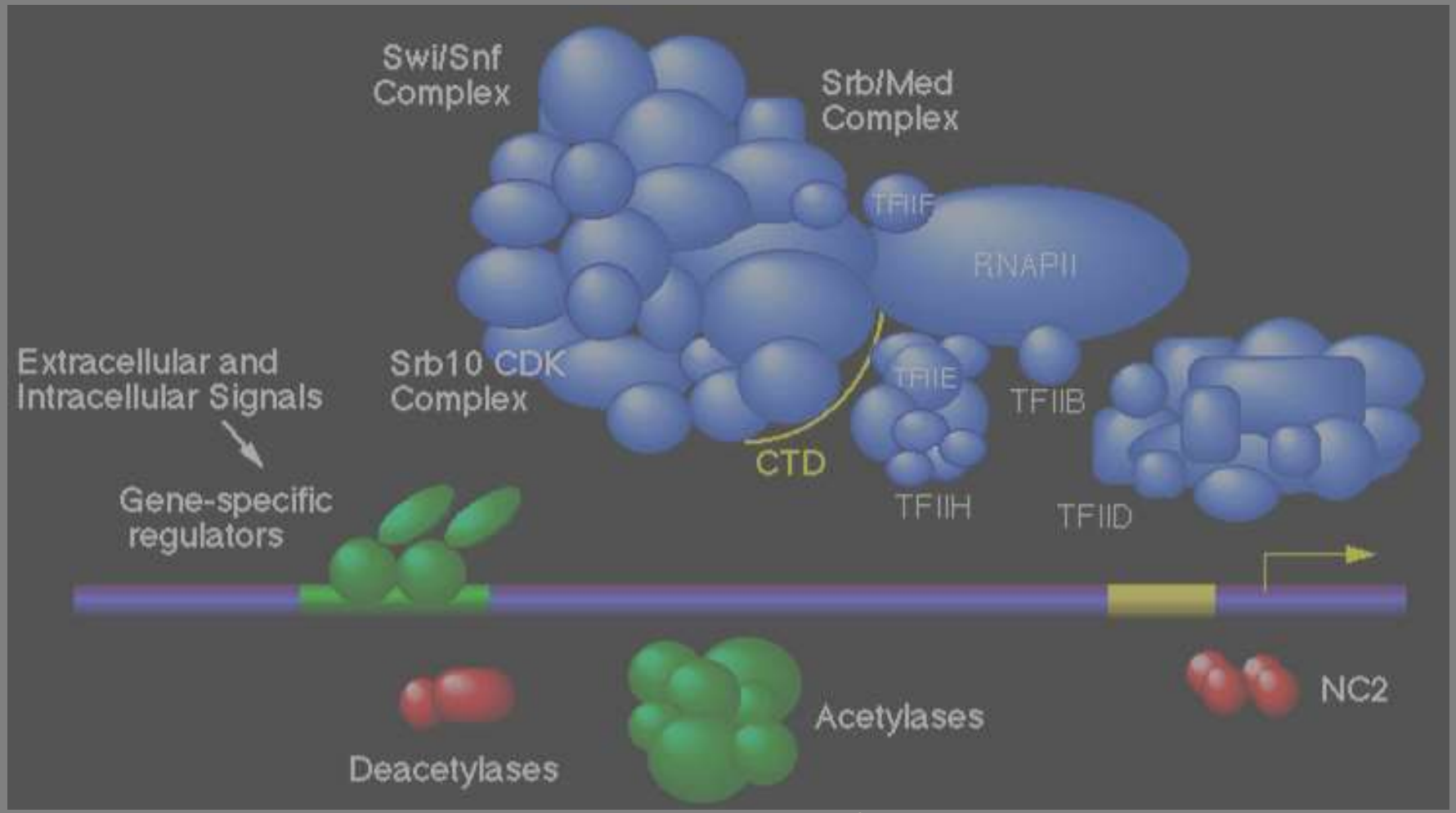
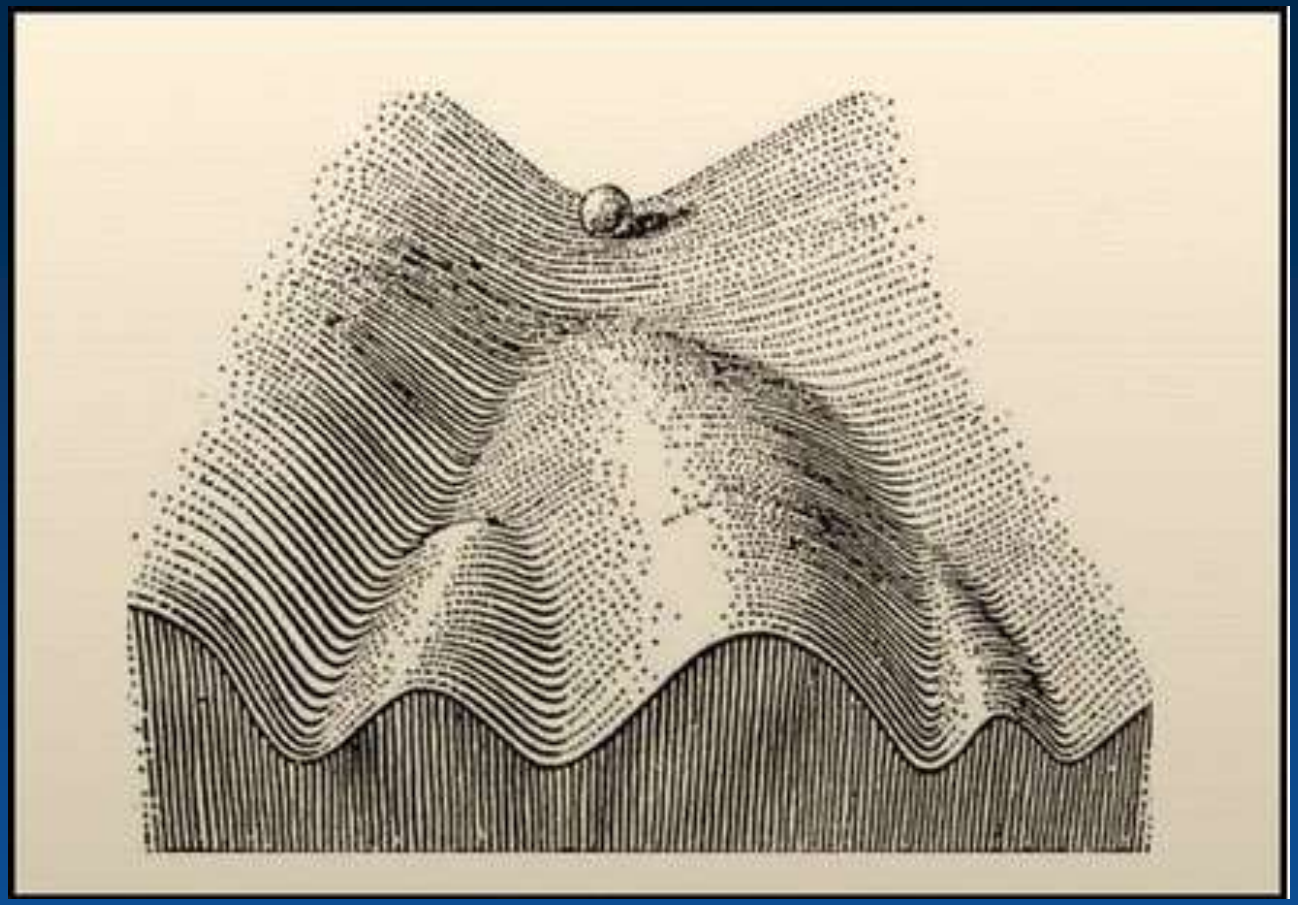
# Regulatory mechanisms of ontogenesis I.

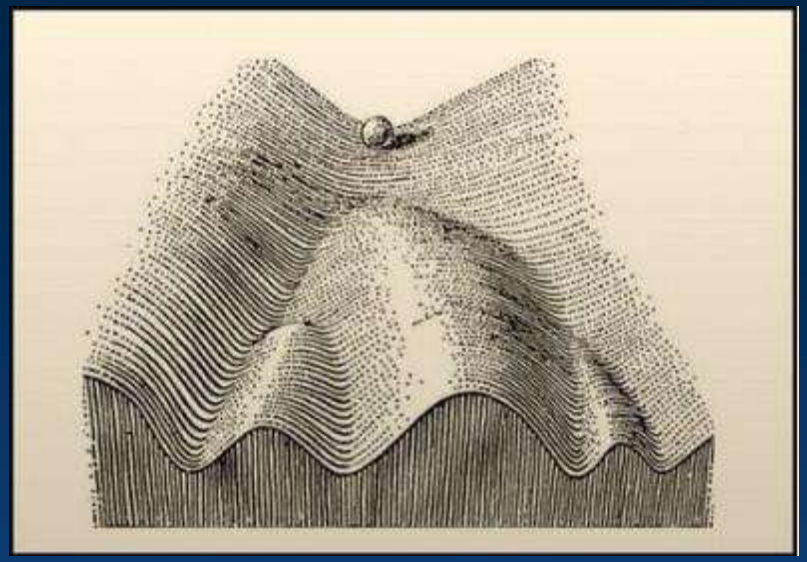



by Krisztina H.-Minkó
Semmelweis University
Developmental biology lectures
2018



Conrad Hal Waddington (November 8, 1905 – September 26, 1975) postulated that it was not just the genes that shaped development but also the environment that shape the genes.

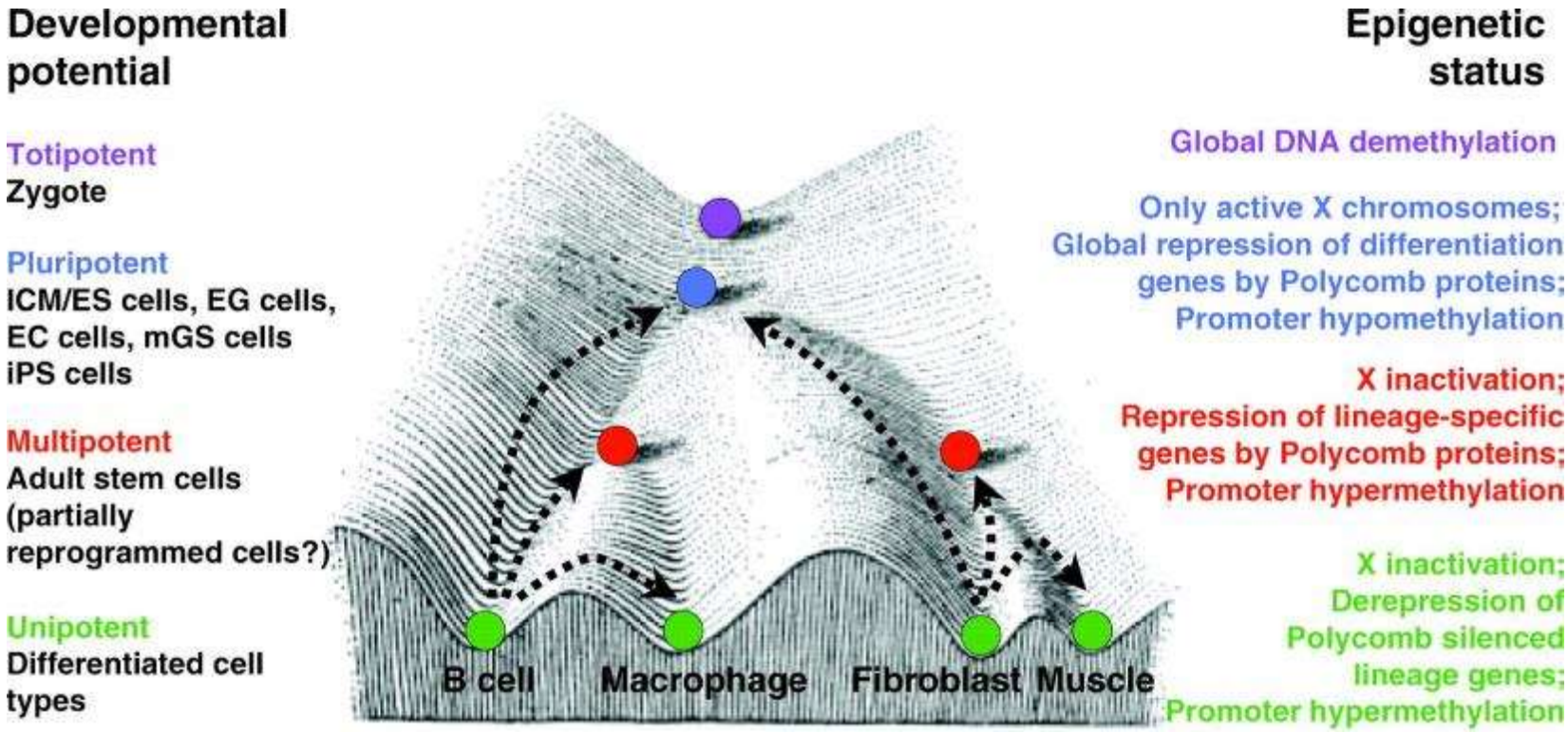
EPIGENETICS: external manifestation of genetic activity



Epigenetic landscape

Waddington's epigenetic landscape is a metaphor for how gene regulation modulates development.[8] One is asked to imagine a number of marbles rolling down a hill towards a wall. The marbles will compete for the grooves on the slope, and come to rest at the lowest points. These points represent the eventual cell fates, that is, tissue types. Waddington coined the term Chreode to represent this cellular developmental process. This idea was actually based on experiment: Waddington found that one effect of mutation (which could modulate the epigenetic landscape) was to affect how cells differentiated. He also showed how mutation could affect the landscape and used this metaphor in his discussions on evolution—he was the first person to emphasise that evolution mainly occurred through mutations that affected developmental anatomy.

**The developmental potential and epigenetic states of cells at different stages of development.**




The developmental potential and epigenetic states of cells at different stages of development. A modification of C. H. Waddington's epigenetic landscape model, showing cell populations with **different developmental potentials** (left) **and their respective epigenetic states** (right). Developmental restrictions can be illustrated as marbles rolling down a landscape into one of several valleys (cell fates). Colored marbles correspond to different differentiation states (purple, totipotent; blue, pluripotent; red, multipotent; green, unipotent). Examples of reprogramming processes are shown by dashed arrows. Adapted, with permission, from Waddington (Waddington, 1957).


THE COMPANY OF Biologists
Published by The Company of Biologists Limited
Development
Hochedlinger K , Plath K Development 2009;136:509-523
dev.biologists.org

Bioessays. 1990 Sep;12(9):441-6.

**Metaphors and the role of genes in development.**

Nijhout HF.

**Source**

Department of Zoology, Duke University, Durham, North Carolina 27706.

**Abstract**

In describing the flawless regularity of developmental processes and the correlation between changes at certain genetic loci and changes in morphology, biologists frequently employ two metaphors: that **genes 'control' development, and that genomes embody 'programs' for development.** Although these metaphors have an admirable sharpness and punch, they lead, when taken literally, to highly distorted pictures of developmental processes. **A more balanced, and useful, view of the role of genes in development is that they act as suppliers of the material needs of development and, in some instances, as context-dependent catalysts of cellular changes, rather than as 'controllers' of developmental progress and direction.** The consequences of adopting this alternative view of development are discussed.

PMID:1979486

DNA is only 50% of the cell nucleus!

50% is made from proteins

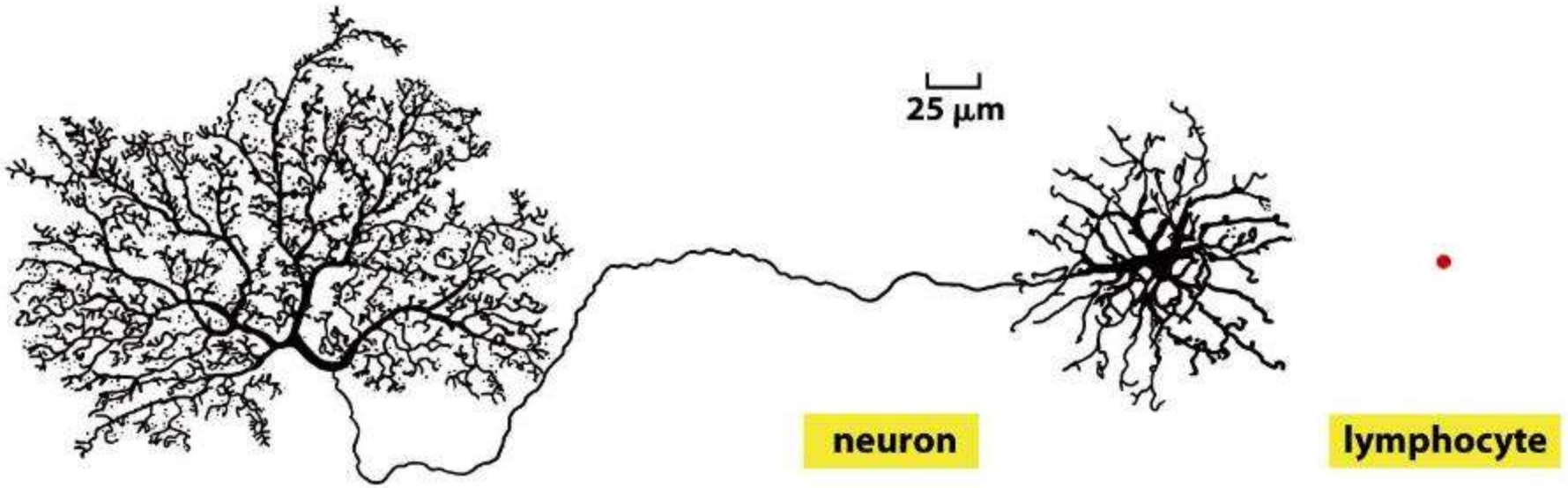
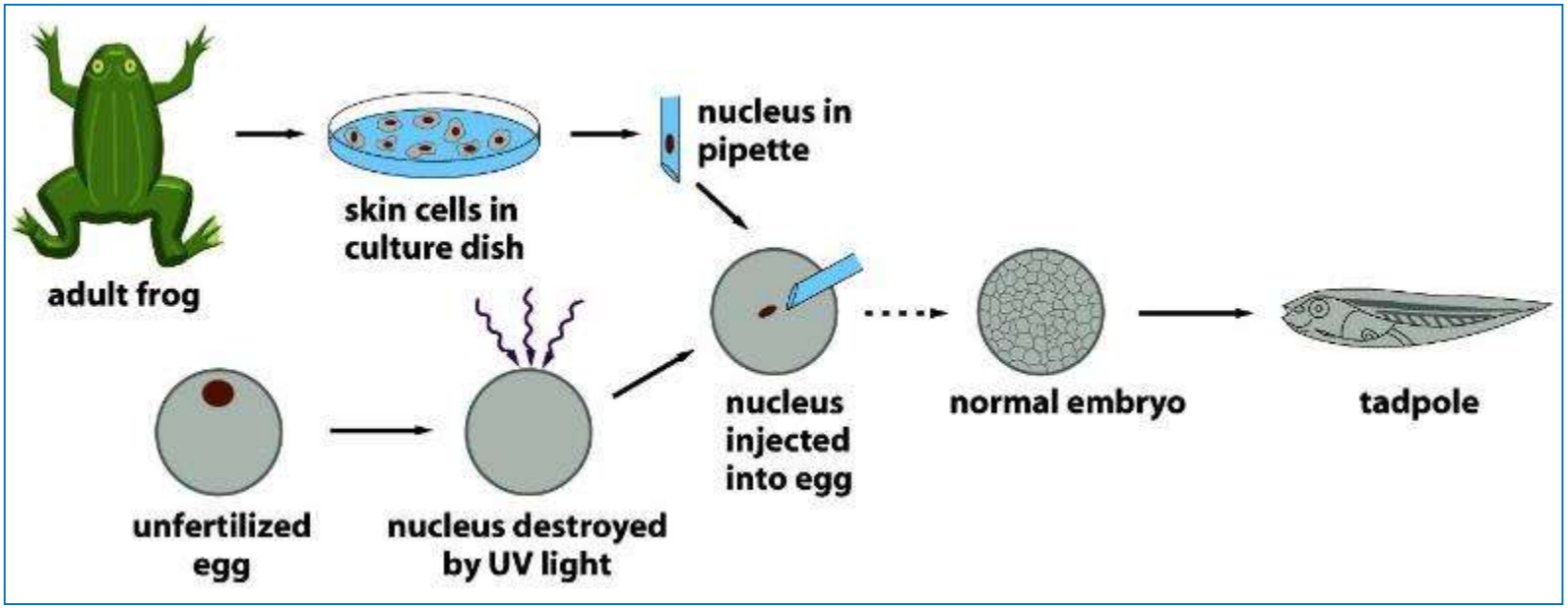
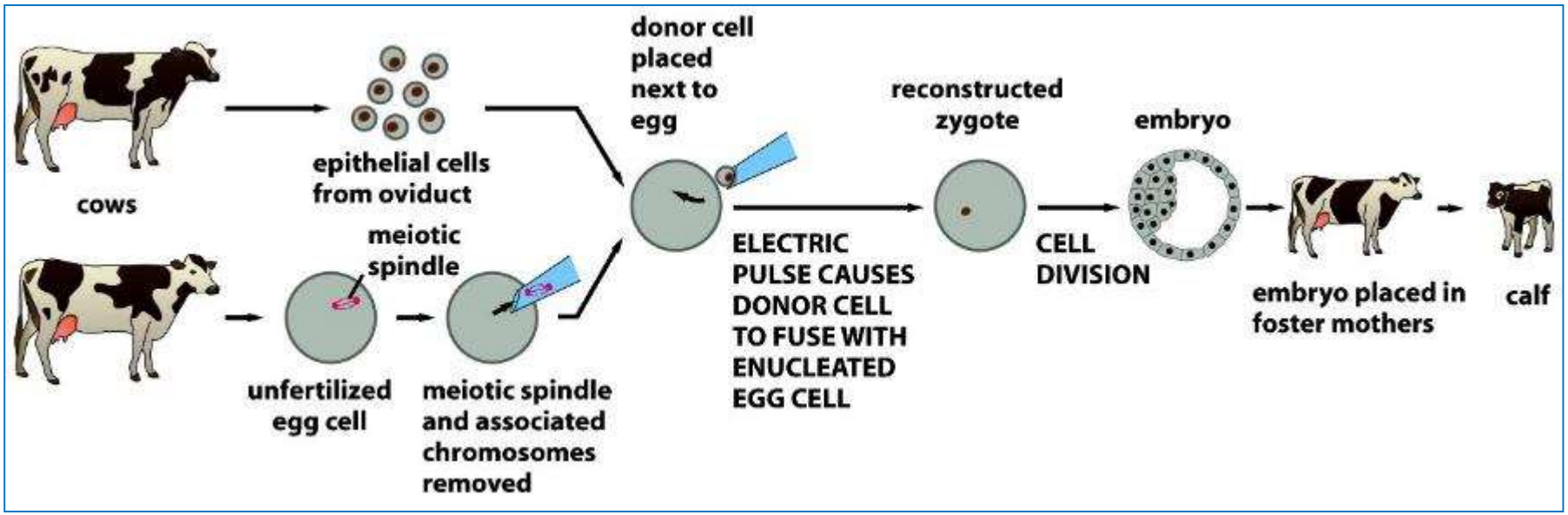



Figure 7-1 *Molecular Biology of the Cell* (© Garland Science 2008)

# Evidence that differentiated cells contains all genetic instructions necessary to direct the formation of a complete organism



Different cell types make different sets of proteins, even through their genomes are identical. Each human beings has roughly 150 000 genes in each nucleus, but each cell uses only a small subset of these genes. Moreover different cell types use different subsets of these genes.

Developmental genetics is the discipline that examines how the genotype is transformed into the phenotype, and **the major paradigm of developmental genetics is differential gene expression from the same nuclear repertoire.**

What determines the particular pattern of gene activity in a differentiated cell?

1. Structural alterations in the genetic material itself during differentiation (cells of the immun system, irreversible differentiation)

2. No irreversible structural changes to the genome:

Differential gene expression

The regulation of gene expression can be accomplished at several levels:

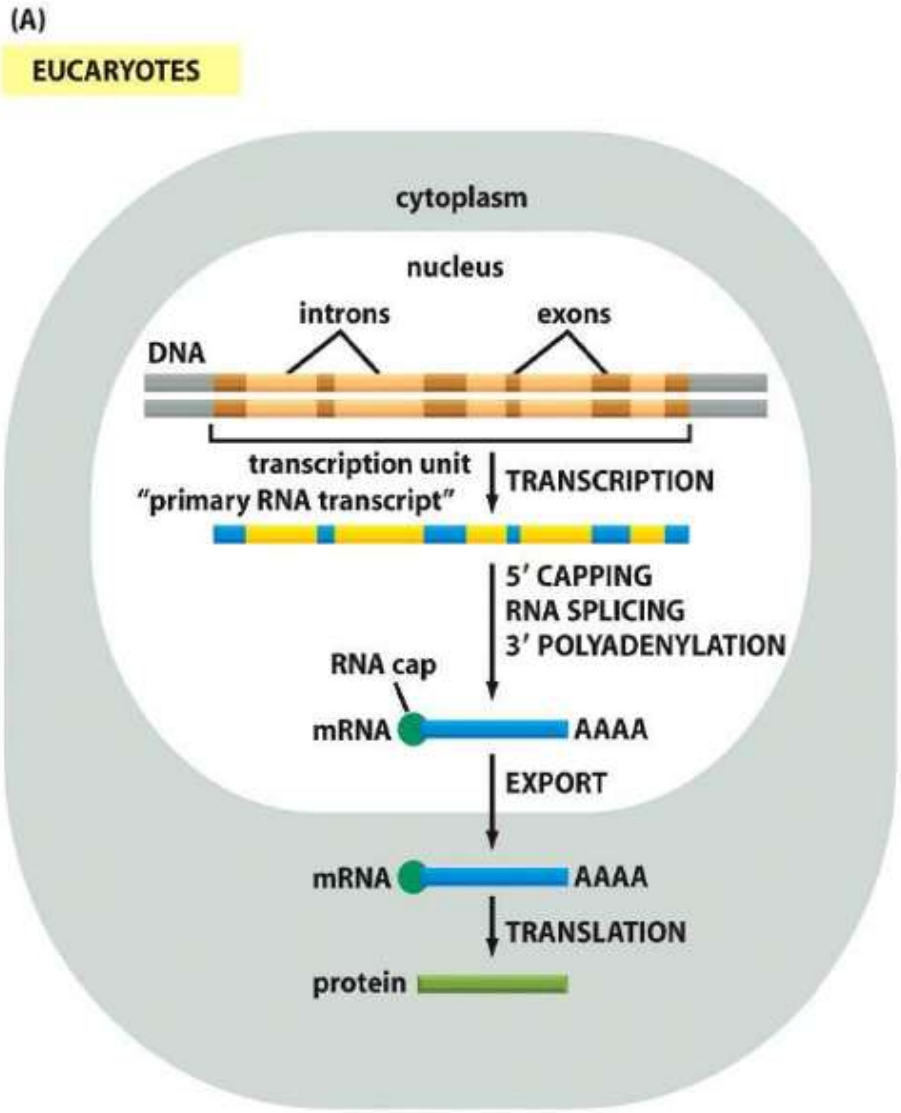
-differential gene transcription

-selective nuclear RNA processing

-selective mRNA translation

-posttranslational modification of the proteins

# From gene to protein




transcription

5’ cap

3’ polyadenilation

RNS splicing

} in the nucleus

RNA transport

translation

} citoplasm

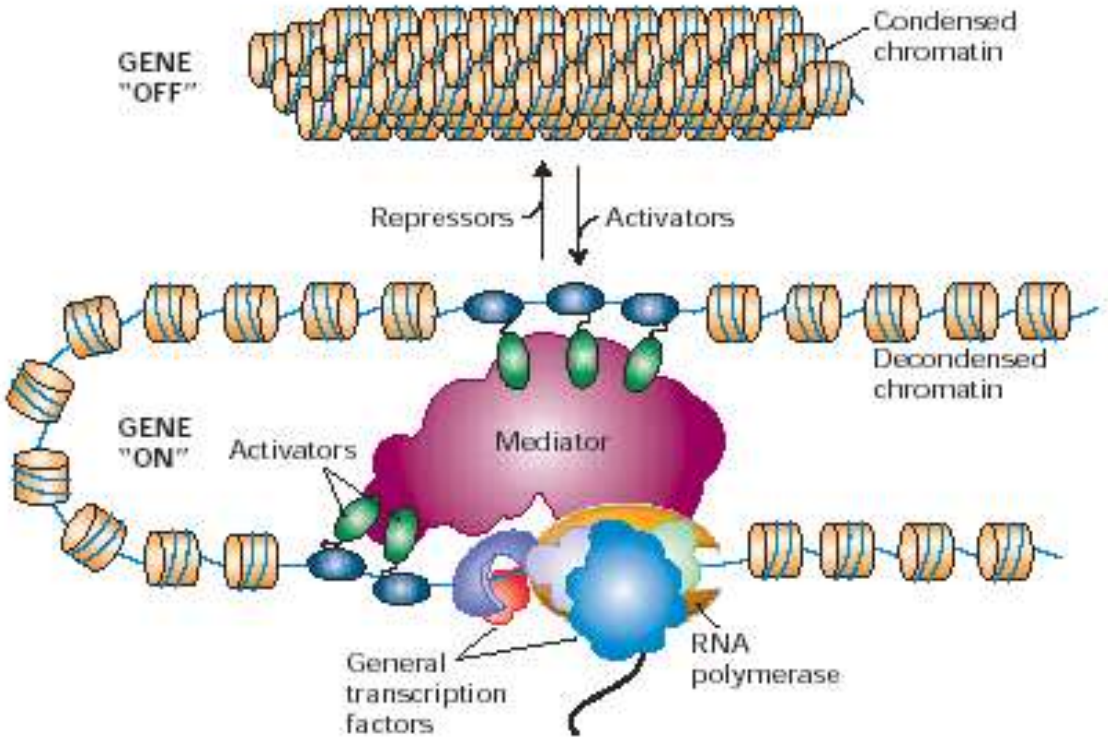
The regulation of gene expression can be accomplished at several levels:

-differential gene transcription

-selective nuclear RNA processing

-selective mRNA translation

-posttranslational modification of the proteins




▶ **FIGURE 11-1 Overview of transcription control in multicellular eukaryotes.** Activator proteins bind to specific DNA control elements in chromatin and interact with multiprotein co-activator machines, such as mediator, to decondense chromatin and assemble RNA polymerase and general transcription factors on promoters. Inactive genes are assembled into regions of condensed chromatin that inhibit RNA polymerases and their associated general transcription factors (GTFs) from interacting with promoters. Alternatively, repressor proteins bind to other control elements to inhibit initiation by RNA polymerase and interact with multiprotein co-repressor complexes to condense chromatin.

# Cryptic transcription

**Cryptic Unstable Transcripts** (CUTs) are a subset of non-coding RNAs (ncRNAs) that are produced from intergenic and intragenic regions. CUTs were first observed in *S. cerevisiae* yeast models and are found in most eukaryotes.[1] Some basic characteristics of CUTs include a length of around 200-800 base pairs,[2] a 5' cap, poly-adenylated tail, and rapid degradation due to the combined activity of poly-adenylating polymerases and exosome complexes.[1][3] CUT transcription occurs through RNA Polymerase II and initiates from nucleosome-depleted regions, often in an antisense orientation.[2][4] To date, CUTs have a relatively uncharacterized function but have been implicated in a number of putative gene regulation and silencing pathways.[5][6][7][8] Thousands of loci leading to the generation of CUTs have been described in the yeast genome.[9] Additionally, Stable Uncharacterized Transcripts, or SUTs, have also been detected in cells and bear many similarities to CUTs but are not degraded through the same pathways. WIKIPEDIA
https://en.wikipedia.org/wiki/Cryptic_unstable_transcript

Transcription. 2013 May-Jun;4(3):97-101. Epub 2013 May 1.

**The great repression: chromatin and cryptic transcription.**

Hennig BP[1], Fischer T.
**Author information**

**Abstract**

The eukaryotic chromatin structure is essential in correctly defining transcription units. Impairing this structure can activate cryptic promoters, and lead to the accumulation of aberrant RNA transcripts. Here we discuss critical pathways that are responsible for the repression of cryptic transcription and the maintenance of genome integrity.

# The organisation of genes on a human chromosome

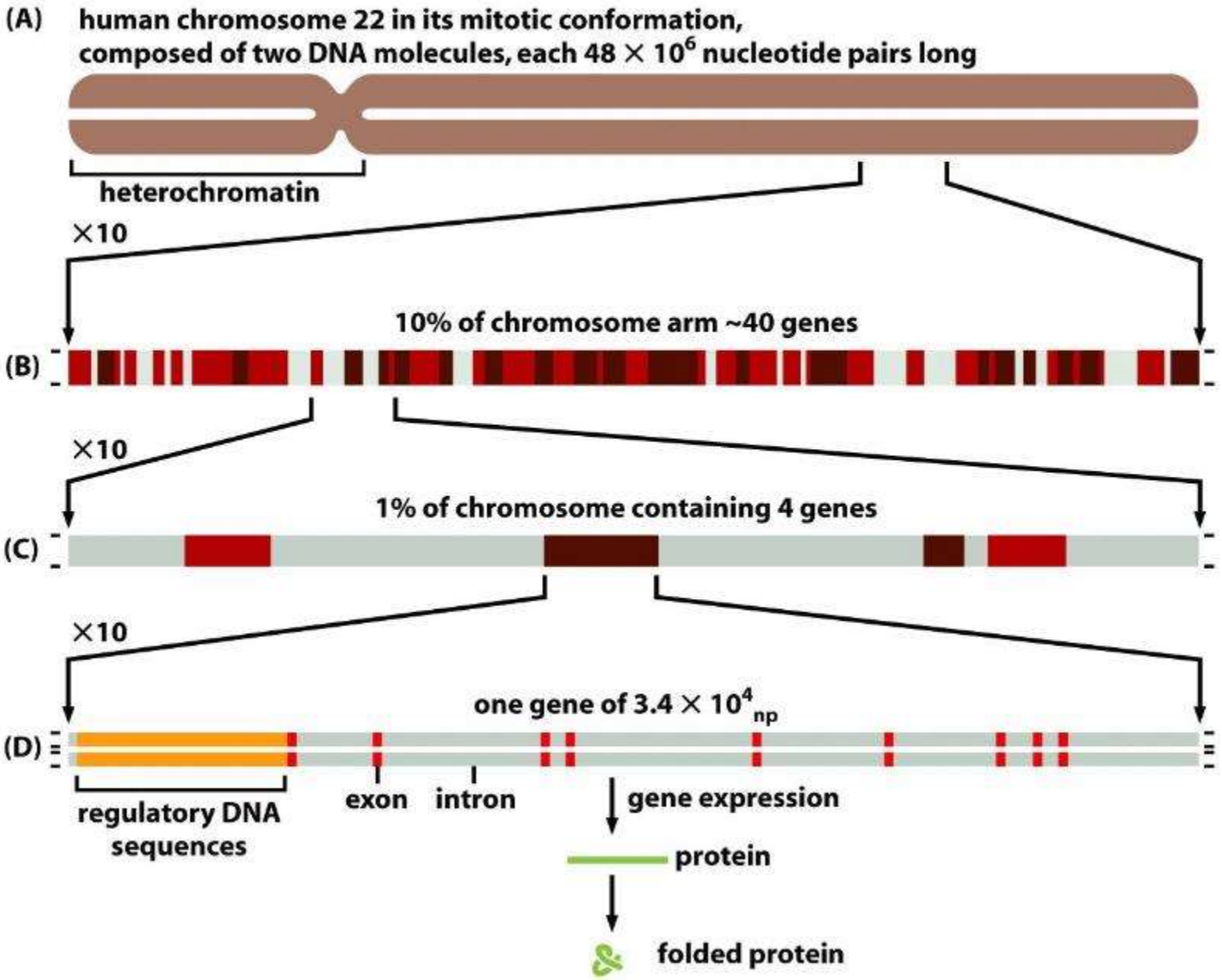
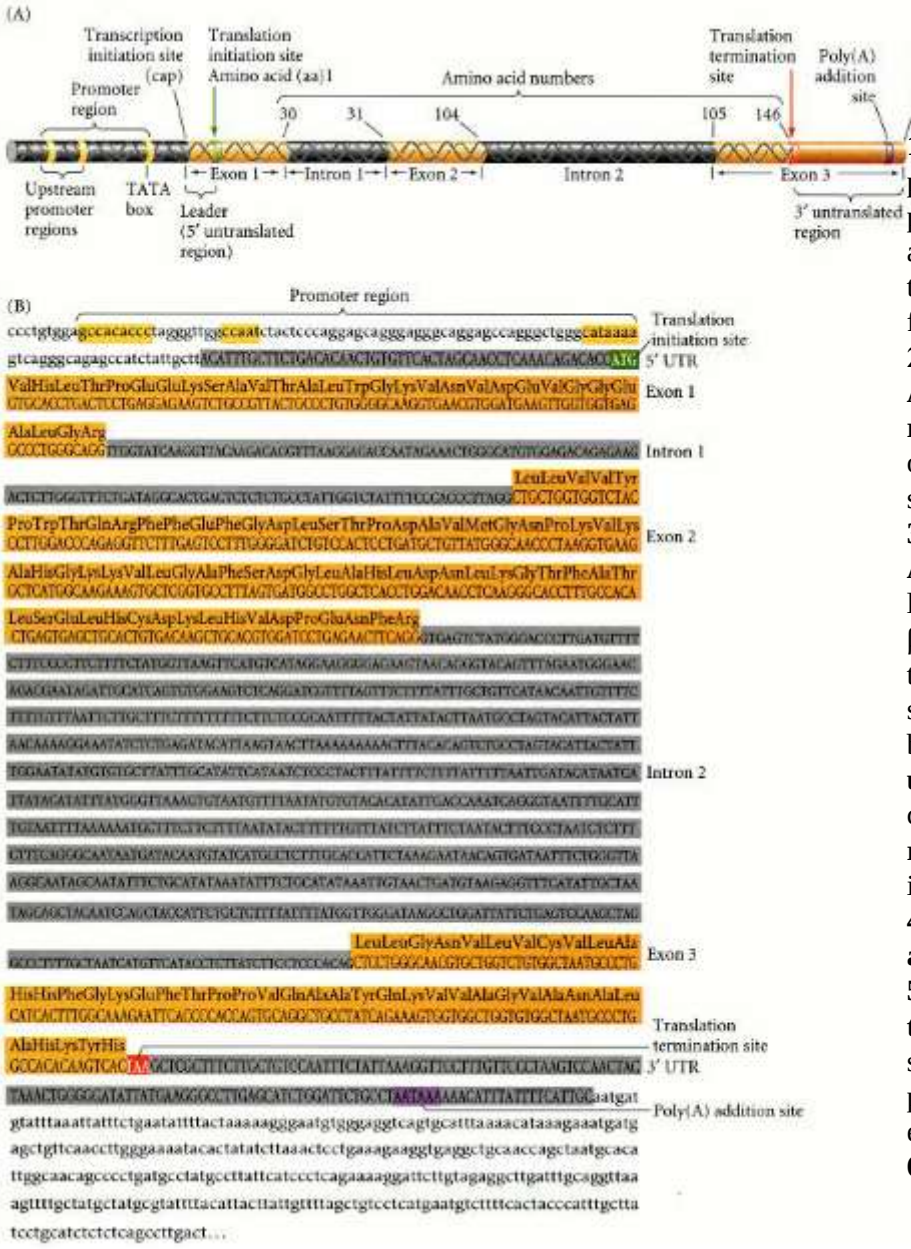



Figure 4-15 *Molecular Biology of the Cell* (© Garland Science 2008)

# Anatomy of the eukaryotic genes: the human β-globin gene as an example 1.




This gene consists of the following elements:

**1. A promoter region, which is** responsible for the binding of RNA polymerase and for the subsequent initiation of transcription. The promoter region of the human β-globingene has three distinct units andextends from 95 to 26 base pairs before ("upstream from") the transcription initiation site (i.e., from -95 to -26).

**2. The transcription initiation site, which for human β-globin is ACATTTG. This site is often** called the **cap sequence because it represents the 5′ end of the RNA, which will receive a "cap"** of modified nucleotides soon after it is transcribed. The specific cap sequence varies among genes.

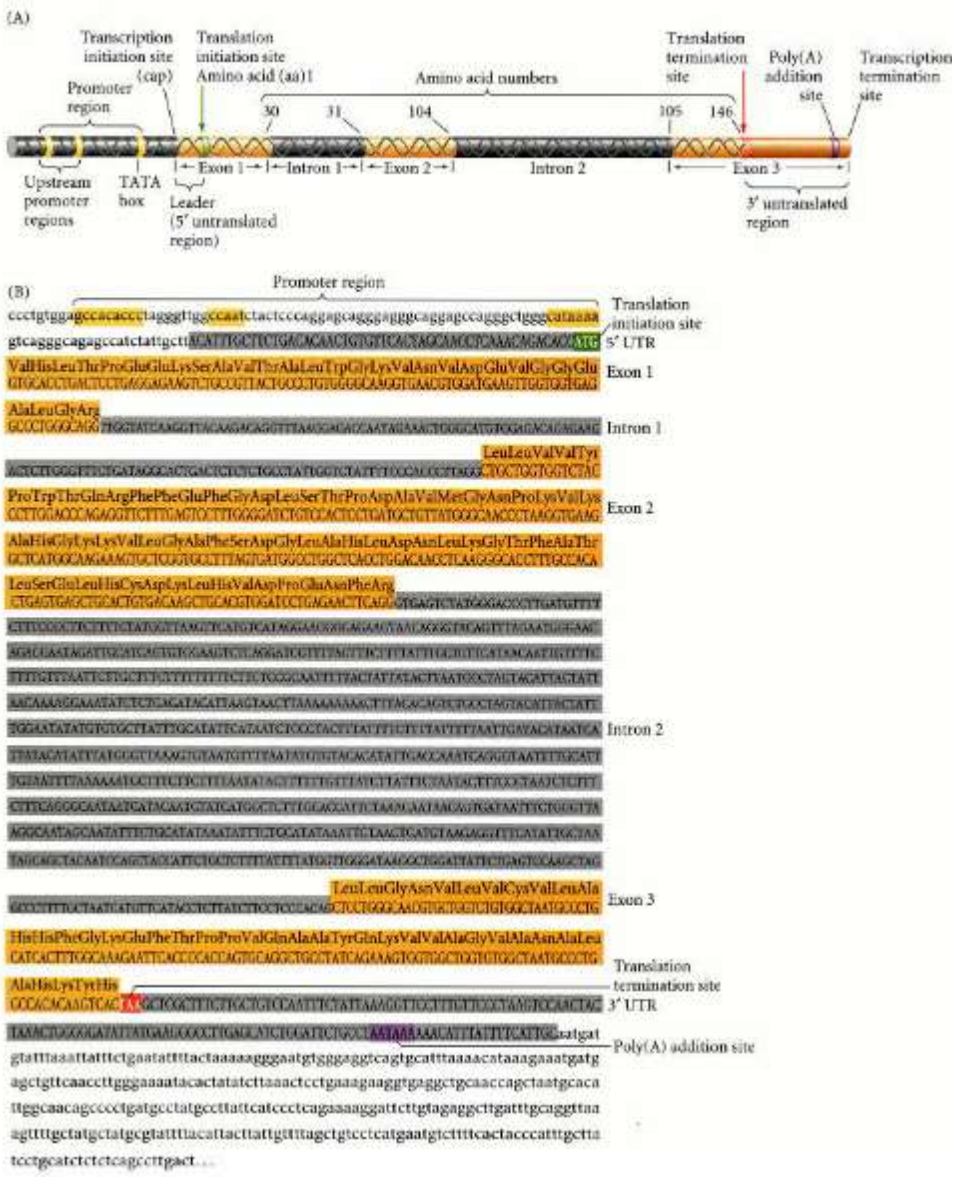
**3. The translation initiation site, ATG. This codon (which becomes AUG in the mRNA) is** located 50 base pairs after the transcription initiation site in the human β-globin gene (although this distance differs greatly among different genes). The intervening sequence of 50 base pairs between the initiation points of transcription and translation is the **5′ untranslated region, often** called the 5′ UTR or **leader sequence. The 5′ UTR can determine the rate at which translation is** initiated.

**4. The first exon, which contains 90 base pairs coding for amino acids 1 30 of human β-globin.**

**5. An intron containing 130 base pairs with no coding sequences for the globin protein. The** structure of this intron is important in enabling the RNA to be processed into messenger RNA and exit from the nucleus.

**6. An exon containing 222 base pairs coding for amino acids 31 104.**

# Anatomy of the eukaryotic genes: the human β-globin gene as an example 2.

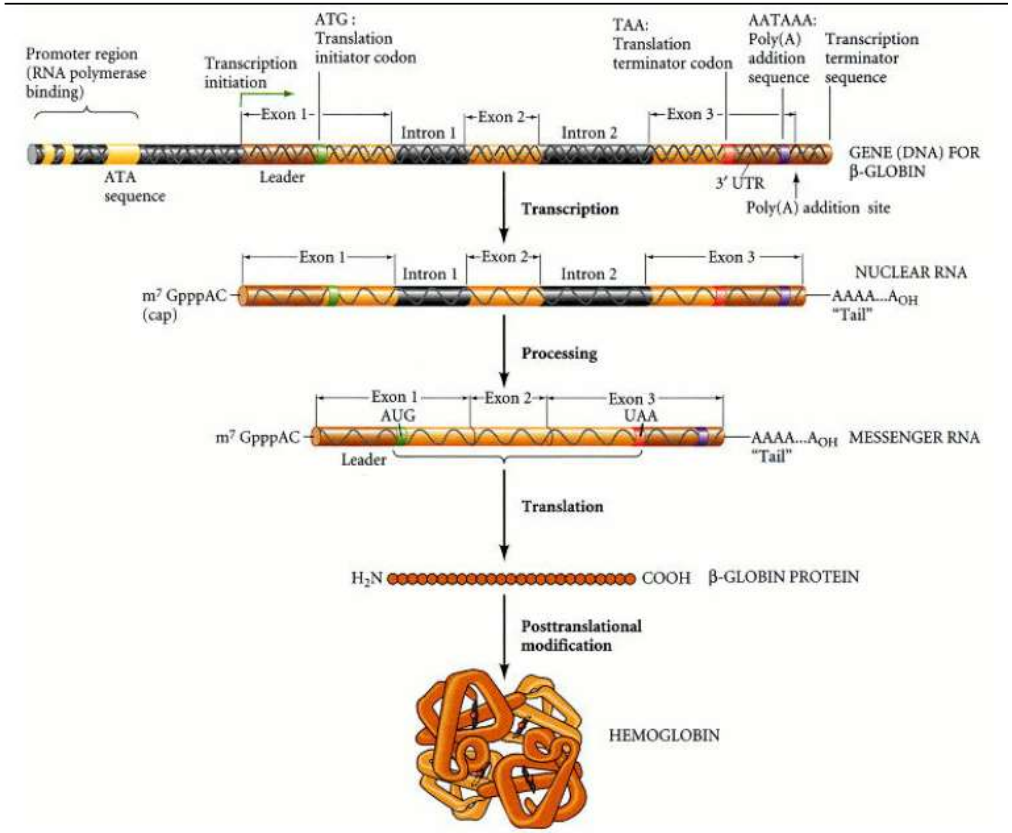



**7. A large intron 850 base pairs having nothing to do with the globin protein structure.**

**8. An exon containing 126 base pairs coding for amino acids 105 146.**

**9. A translation termination codon, TAA. This codon becomes UAA in the mRNA. The**
ribosome dissociates at this codon, and the protein is released.

**10. A 3´ untranslated region that, (3´ UTR) although transcribed, is not translated into protein.**
This region includes the sequence AATAAA, which is needed for **polyadenylation: the lacement**
of a "tail" of some 200 to 300 adenylate residues on the RNA transcript. This **poly(A) tail**
(1) confers stability on the mRNA, (2) allows the mRNA to exit the nucleus, and (3) permits the
mRNA to be translated into protein.The poly(A) tail is inserted into the RNA about 20 bases
downstream of the AAUAAA sequence. Transcription continues beyond the AATAAA site for
about 1000 nucleotides before being terminated.




The original nuclear RNA transcript for such a gene contains the capping sequence, the 5′ untranslated region, the exons, the introns, and the 3′ untranslated region. In addition, both its ends become modified. **A cap consisting of methylated guanosine is placed on the 5′ end of the RNA in opposite polarity to the RNA itself.** This means that there is no free 5′ phosphate group on the nuclear RNA. The 5′ cap is necessary for the binding of mRNA to the ribosome and for subsequent translation (Shatkin 1976). **The 3′ terminus is usually modified in the nucleus by the addition of a poly(A) tail.** These adenylate residues are put together enzymatically and are added to the transcript; they are not part of the gene sequence. Both the 5′ and 3′ modifications may protect the RNA from exonucleases that would otherwise digest the mRNA (Sheiness and Darnell 1973; Gedamu and Dixon 1978). The modifications thus stabilize the message and its precursor.

# Maintenance and inheritance of patterns of gene activity

# Structural organisation of the chromatin

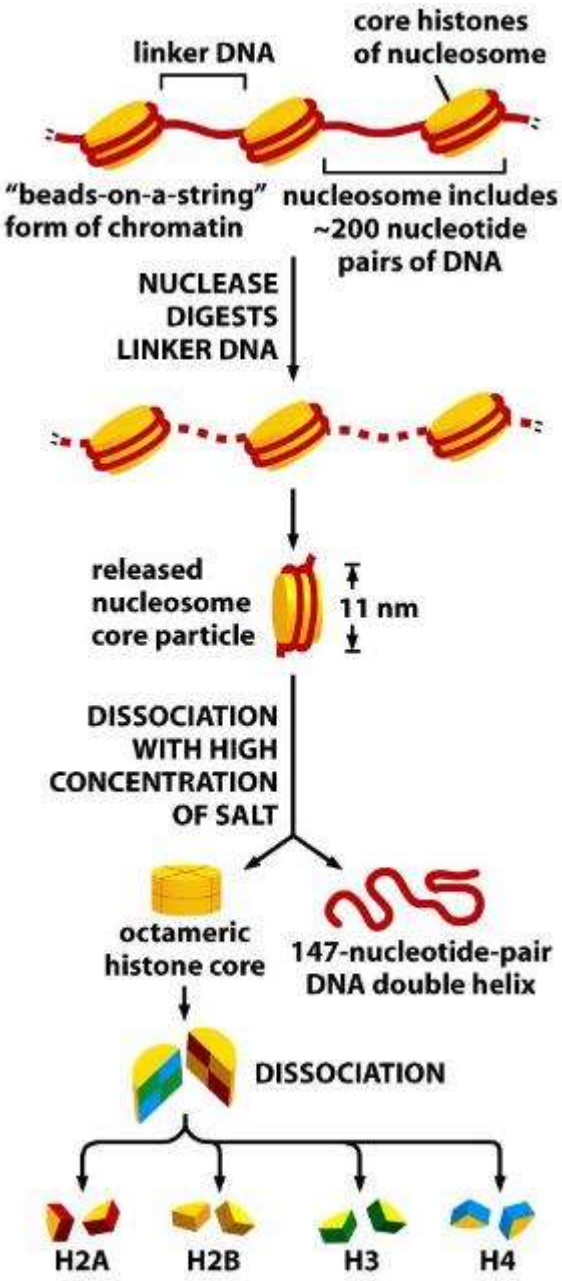



Figure 4-23 *Molecular Biology of the Cell* (© Garland Science 2008)

# Maintenance and inheritance of patterns of gene activity

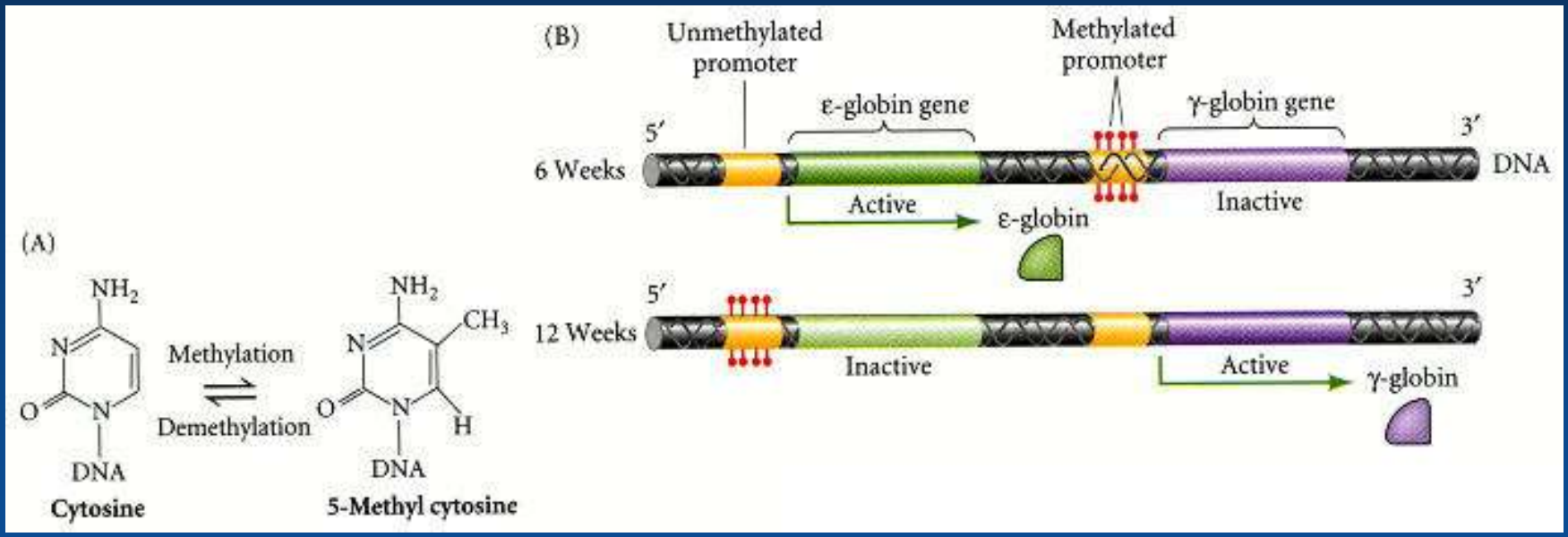
DNA methylation

In vertebrates, methylation of cytosine at certain sites in the DNA is correlated with the absence of transcription in those regions. The pattern of methylation can be faithfully inherited when DNA replicates.

Histone acetylation /deacetylation

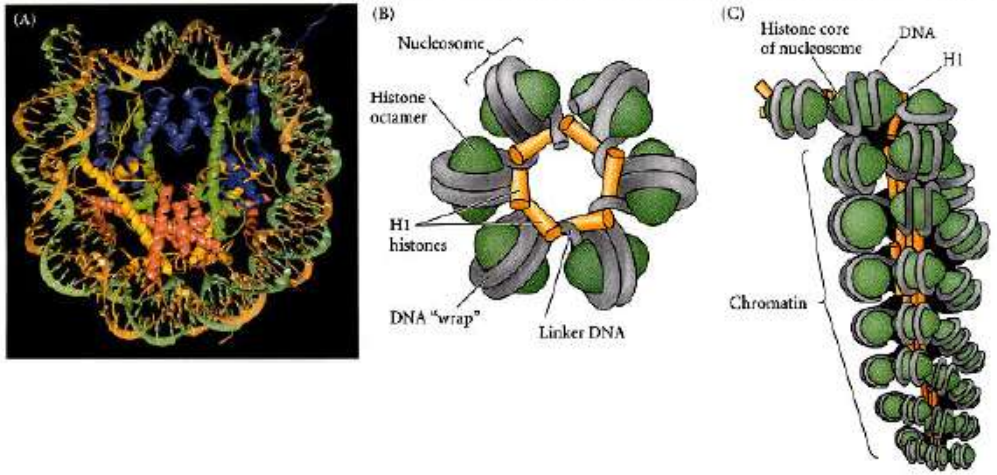
How is methylation involved in repressing genes? One hypothesis is that methylated DNA stabilizes nucleosomes. Here DNA methylation is linked to histone deacetylation. Whereas acetylated histones are relatively unstable and cause the nucleosomes to disperse, deacetylated histones form a stable nucleosome.

# Methylation of globin genes in human embryonic blood cells



# The „default” condition of chromatin is probably a repressed state

There are two fundamental differences distinguishing most eukaryotic genes from most prokaryotic genes. First, eukaryotic genes are contained within a complex of DNA and protein called **chromatin**. The protein component constitutes about half the weight of chromatin and is composed largely of **nucleosomes**. The nucleosome is the basic unit of chromatin structure. It is composed of an octamer of **histone** proteins (two molecules each of histones H2A-H2B and histones H3-H4) wrapped with two loops containing approximately 140 base pairs of DNA (Figure 5.1; Kornberg and Thomas 1974). Chromatin can thus be visualized as a string of



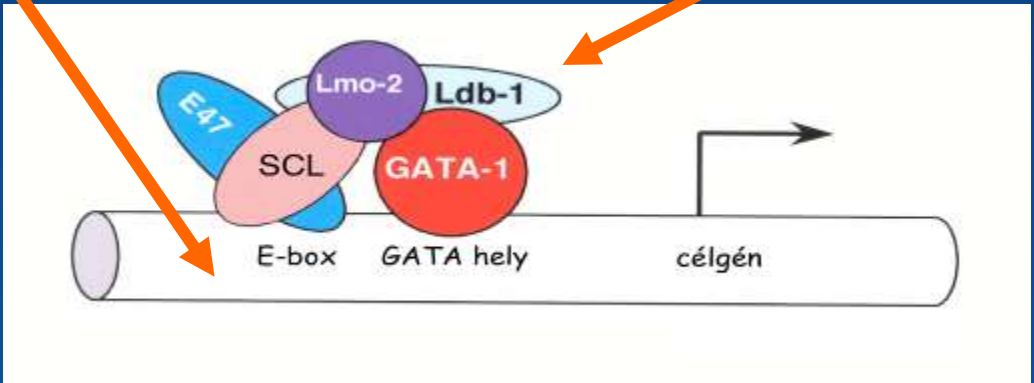

nucleosome beads linked by ribbons of DNA. While classic geneticists have likened genes to "beads on a string," molecular geneticists liken genes to "string on the beads." Most of the time, the nucleosomes are themselves wound into tight "solenoids" that are stabilized by histone H1.

Histone H1 is found in the 60 or so base pairs of "linker" DNA between the nucleosomes (Weintraub 1984). This H1-dependent conformation of nucleosomes inhibits the transcription of genes in somatic cells by packing adjacent nucleosomes together into tight arrays that prohibit the access of transcription factors and RNA polymerases to the genes (Thoma et al. 1979; Schlissel and Brown 1984). It is generally thought, then, that the "default" condition of chromatin is a repressed state, and that tissue-specific genes become activated by local interruption of this repression (Weintraub 1985).

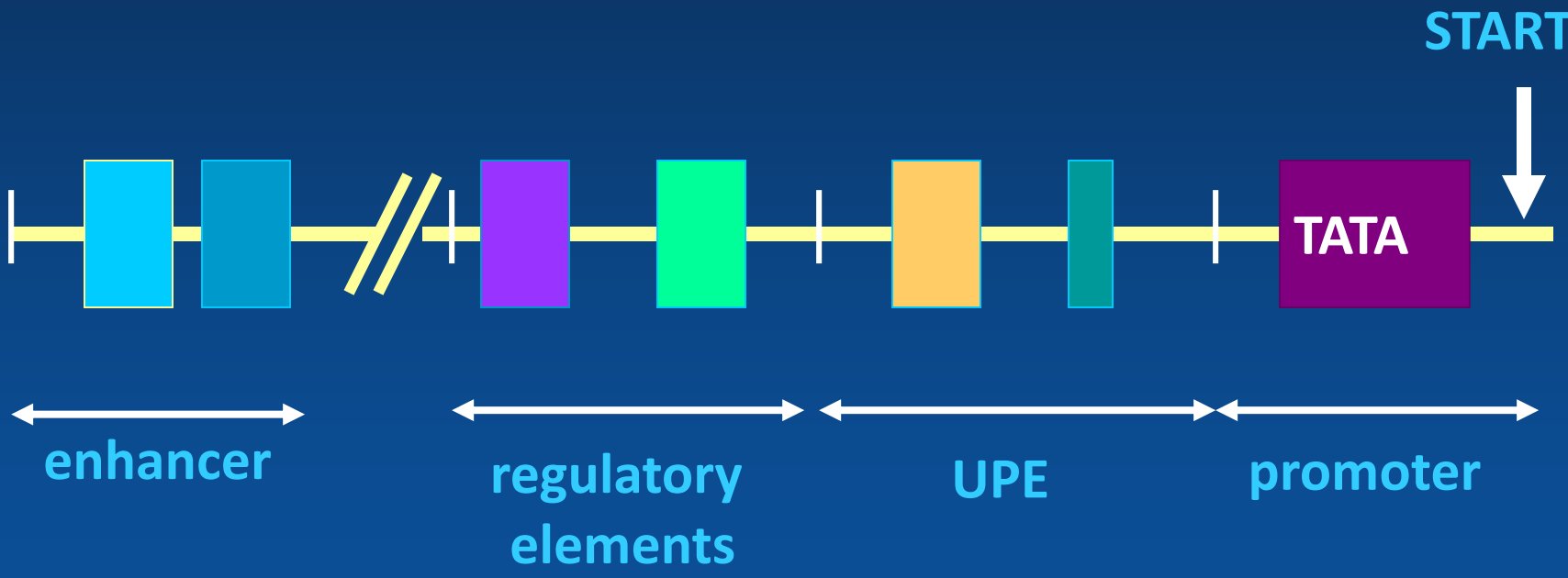
# Transcription regulation

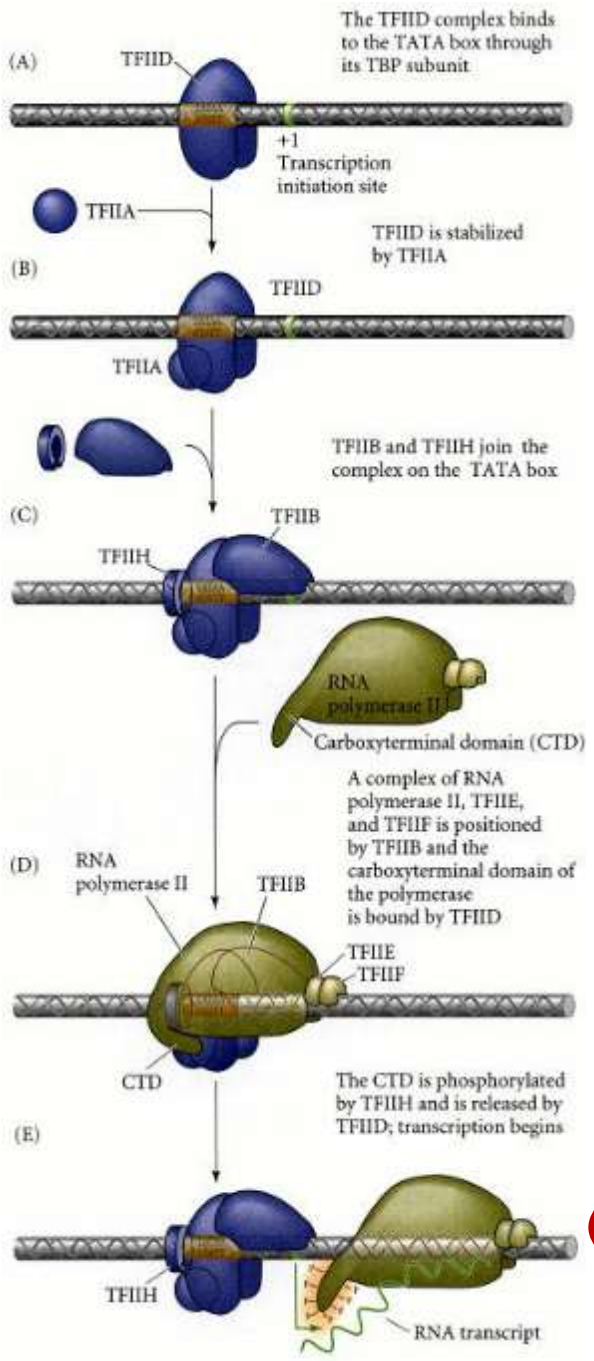
cis elements
DNA sequences

trans elements
Genes of regulatory proteins



# Gene regulatory elements in eukaryotes





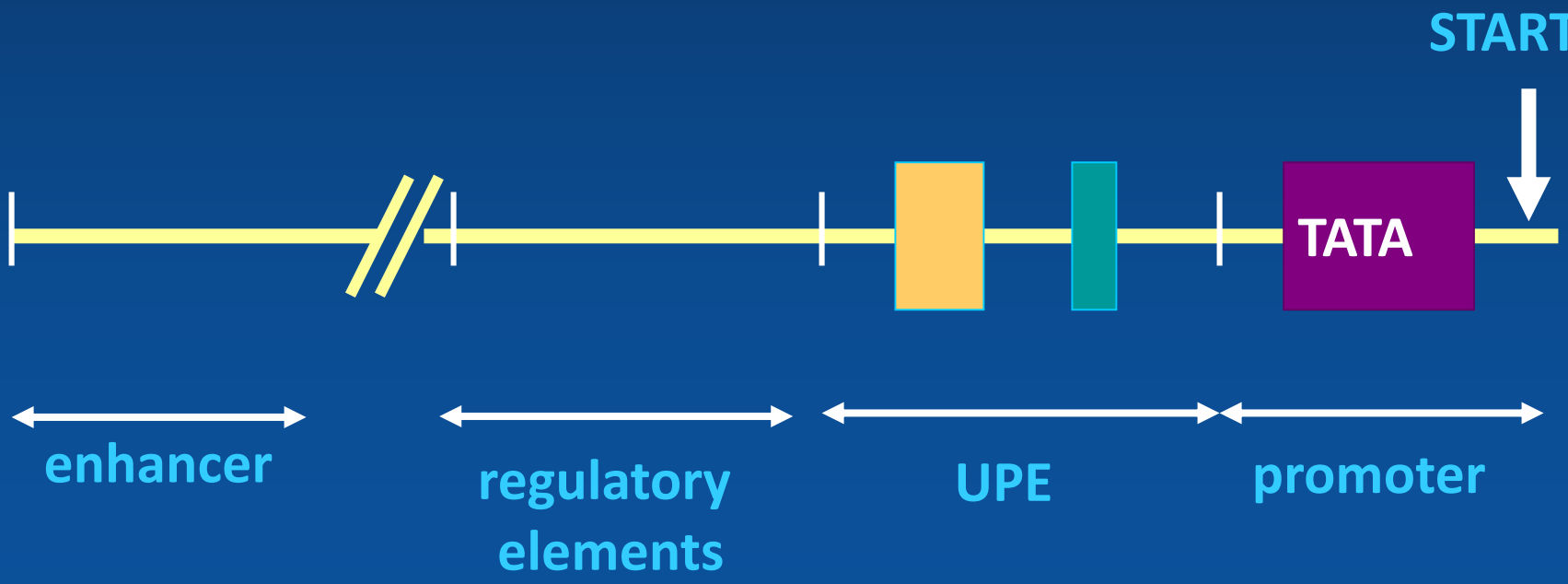

**Promoters are the sites where RNA polymerase binds to the DNA to initiate** transcription. Promoters of genes that synthesize messenger RNAs (i.e., genes that encodeproteins ) are typically located immediately upstream from the site where the RNA polymerase initiates transcription. Most of these promoters contain the sequence TATA, where RNA polymerase will be bound (Figure 5.4). This site, known as the **TATA box, is usually about 30** base pairs upstream from the site where the first base is transcribed. Eukaryotic RNA polymerases, however, will not bind to this naked DNA sequence. Rather, they require additional protein factors to bind efficiently to the promoter. The protein-encoding genes are transcribed by RNA polymerase II, and at least six nuclear proteins have been shown to be necessary for the proper initiation of transcription by RNA polymerase II (Buratowski et al. 1989; Sopta et al.1989). These proteins are called **basal transcription factors. The first of these, TFIID,** recognizes the TATA box through one of its subunits, **TATA-binding protein (TBP). TFIID** serves as the foundation of the transcription initiation complex, and it also serves to keep nucleosomes from forming in this region. Once TFIID is stabilized by TFIIA, it becomes able to bind **TFIIB. Once TFIIB is in place, RNA polymerase can bind to this complex. Other** transcription factors (TFIIE, F, and H) are then used to release RNA polymerase from the complex so that it can transcribe the gene, and to unwind the DNA helix so that the RNA polymerase will have a free template from which to transcribe.

In addition to these basal transcription factors, which are found in each nucleus, there is also a set of transcription factors called **TBP-associated factors, or TAFs (Figure 5.5; Buratowski 1997; Lee and Young** 1998), which can stabilize the TBP. This function is critical for gene transcription, for if the TBP is not stabilized, it can fall off the small TATA sequence. The TAFs are bound by **upstream promoter elements on the DNA. These DNA sequences are near** the TATA sequence, and usually upstream from it. These TAFs need not be in every cell of the body, however.

**Cell-specific transcription factors (such as the Pax6** and microphthalmia proteins can also activate the gene by stabilizing the transcription initiation complex. They can do so by binding to the TAFs, by binding directly to other factors such as TFIIB, or (as we will see soon) by destabilizing nucleosomes.

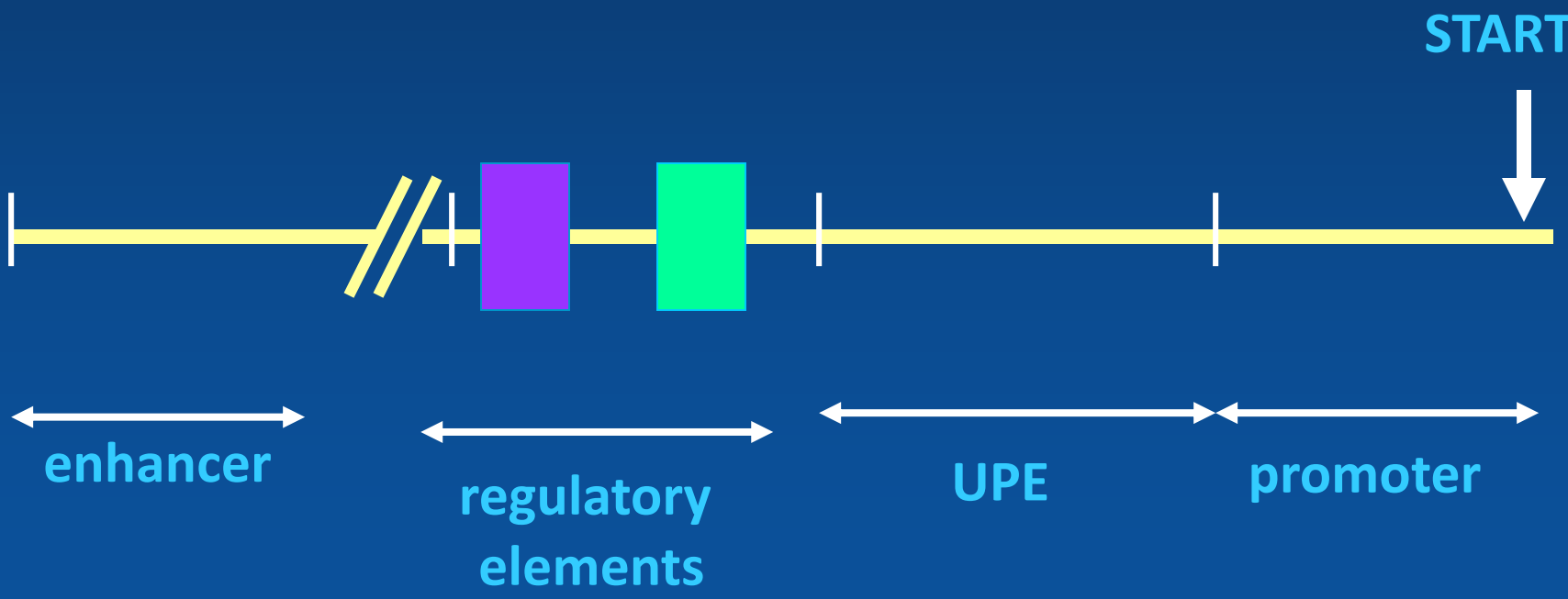
# For basic transcription you need:

| | |
|---|---|
| generally TATA box: | RNS polimerase II and other factors |
| UPE: | upstream promoter elements, enforce the promoter |



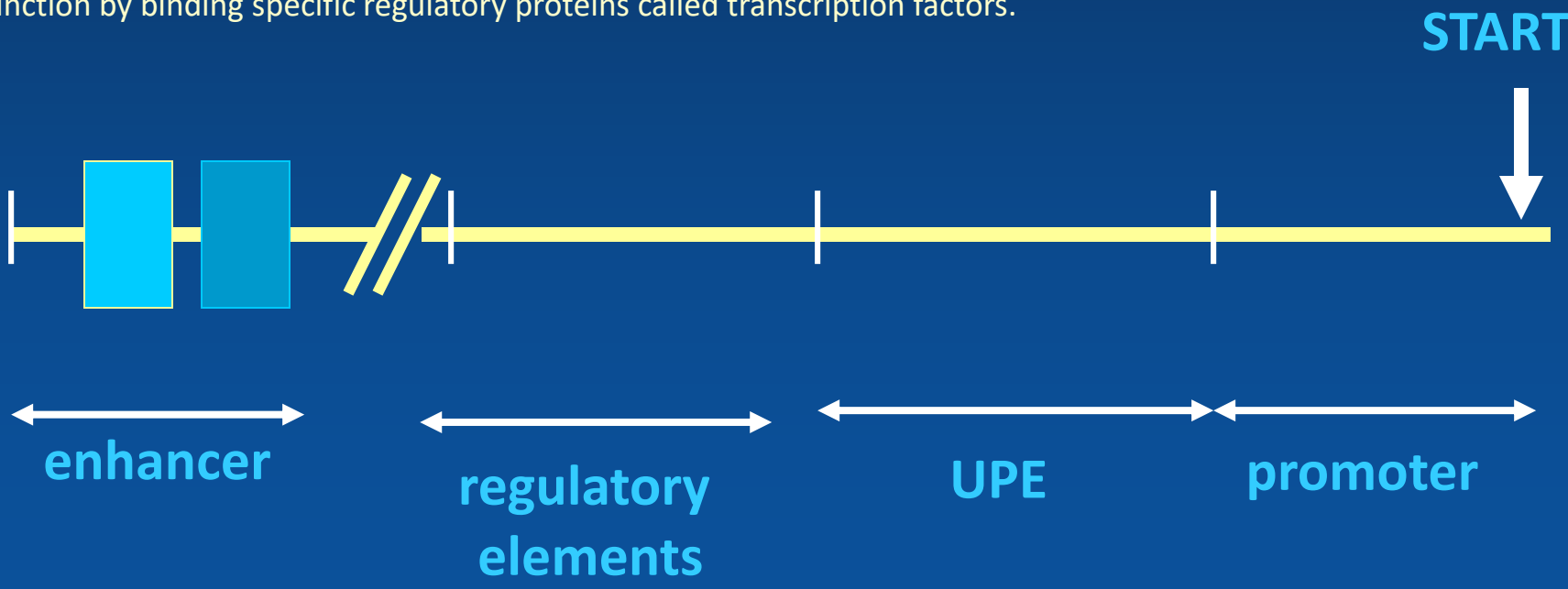
# For regulated transcription you need regulatory sequences

Gene specific sequences such as:
heat shock element, glucocorticoid responsive element, GATA site,
E box, N-box, etc.



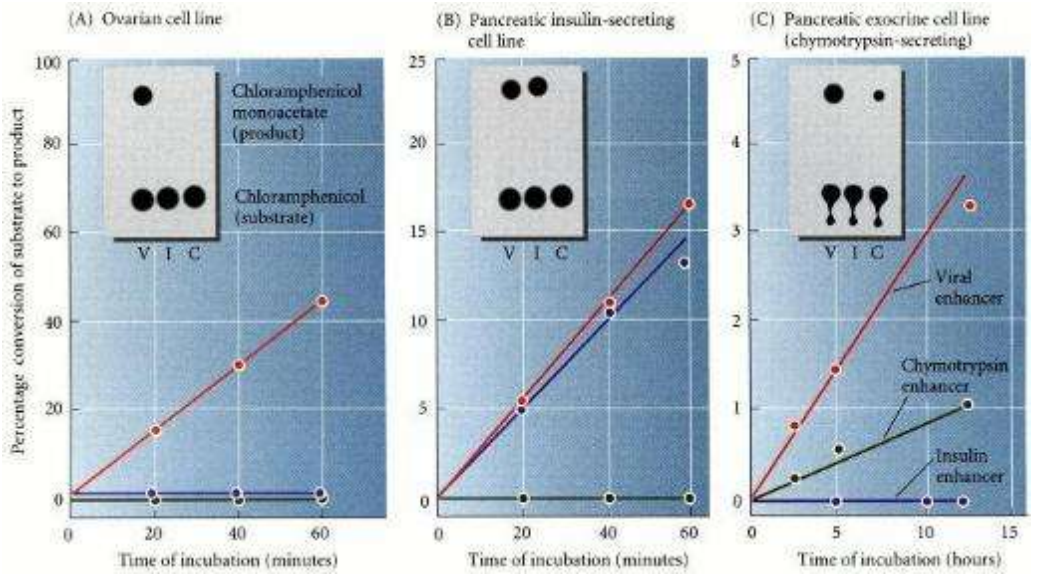
# Regulating sequences in distance: enhancers

An **enhancer is a DNA sequence that can activate the utilization of a promoter,** controlling the efficiency and rate of transcription from that particular promoter. Enhancers can activate only *cis-linked promoters (i.e., promoters on the same chromosome ), but they can do so* at great distances (some as great as 50 kilobases away from the promoter). Moreover, enhancers do not need to be on the 5′ (upstream) side of the gene. They can also be at the 3′ end, in the introns, or even on the complementary DNA strand (Maniatis et al. 1987). The human β-globin gene has an enhancer in its 3′ UTR, roughly 700 base pairs downstream from the AATAAA site. This sequence is necessary for the temporal- and tissue-specific expression of the β-globin gene in adult red blood cell precursors (Trudel and Constantini 1987). Like promoters, enhancers function by binding specific regulatory proteins called transcription factors.



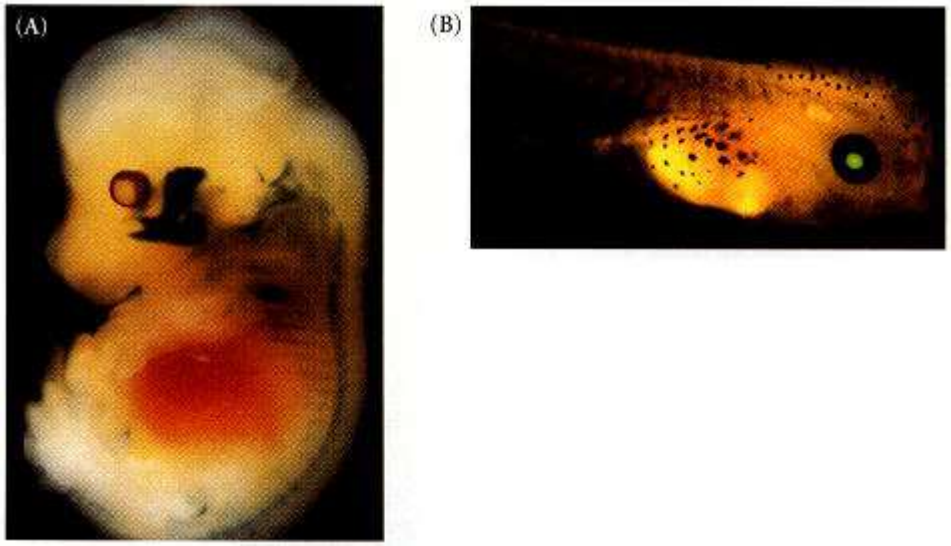
**Enhancers can regulate the temporal and tissue-specific expression of any differentially regulated gene, but different types of genes normally have different enhancers.**

In the pancreas, for instance, the exocrine protein genes (for the digestive proteins chymotrypsin, amylase, and trypsin) have enhancers different from that of the gene for the endocrine protein insulin. These enhancers both lie in the 5′ flanking sequences of their genes. Walker and colleagues et al. 1983 created transgenes by placing flanking regions from the genes for chymotrypsin and insulin onto the gene for bacterial chloramphenicol acetyltransferase (CAT), an enzyme that is not found in mammalian cells. CAT activity is easy to assay in mammalian cells, so the bacterial *CAT gene* can be used as a **reporter gene to tell investigators whether a particular enhancer is functioning.** The researchers then transfected the transgenes into (1) ovary cells (which do not secrete either insulin or chymotrypsin), (2) an insulin-secreting cell line, and (3) a chymotrypsin-secreting cell line, and measured the activity of CAT in each of these cells. As shown in Figure 5.6, neither enhancer sequence caused the enzyme to be made in the ovarian cells. In the insulin-secreting cell, however, the 5′ flanking region from the insulin gene enabled the *CAT gene to be expressed,* but the 5′ flanking region of the chymotrypsin gene did not. Conversely, when the clones were placed into the exocrine pancreatic cell line, the chymotrypsin 5′ flanking sequence allowed *CAT* expression, while the insulin enhancer did not. The enhancers for 10 different exocrine proteins share a 20-base-pair consensus sequence, suggesting that these similar sequences play a role in activating an entire set of genes specifically expressed in the exocrine cells of the pancreas (Boulet et al. 1986). Thus, **the expression of genes in exocrine and in endocrine cells of the pancreas appears to be controlled by different enhancers.**



# Enhancers activate any gene sequences in experimental conditions

By taking an enhancer from one gene and fusing it to another gene, it has been shown that enhancers can direct the expression of any gene sequence. For instance, the β-**galactosidase** gene from *E. coli (the lacZ gene) can be used as a* ***reporter gene*** *and placed onto an enhancer that* normally directs a particular mouse gene to become expressed in muscles. If the resulting transgene is injected into a newly fertilized mouse egg and gets incorporated into its DNA, the β-galactosidase gene will be expressed in the muscle cells. By staining for the presence of β-galactosidase, the expression pattern of that muscle-specific gene can be seen (Figure 5.7A). Similarly, if the gene for **green fluorescent protein (GFP, a reporter protein that is usually made** only in jellyfish) is placed on the enhancer of genes encoding the crystallin proteins of the eye lens, GFP expression is seen solely in the lens (Figure 5.7B).



Enhancers are critical in the regulation of normal development. Over the past decade, six generalizations that emphasize their importance for differential gene expression have been made:

1. **Most genes require enhancers for their transcription.**

**2. Enhancers are the major determinant of differential transcription in space (cell type) and time.**

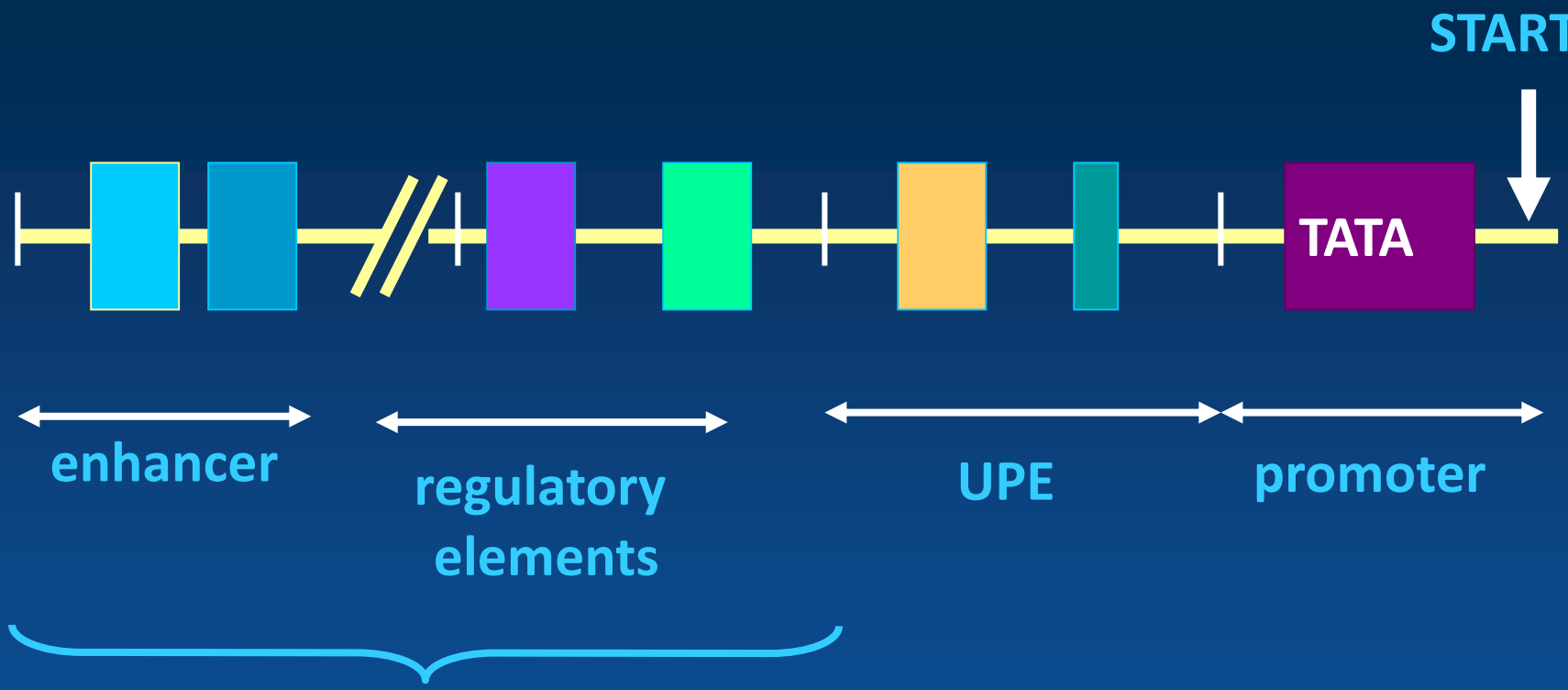
**3. The ability of an enhancer to function while far from the promoter means that there can be** multiple signals to determine whether a given gene is transcribed. A given gene can have several enhancer sites linked to it, and each enhancer can be bound by more than one transcription factor.

**4. The interaction between the proteins bound to the enhancer sites and the transcription initiation** complex assembled at the promoter is thought to regulate transcription. The mechanism of this association is not fully known, nor do we comprehend how the promoter integrates all these signals.

**5. Enhancers are modular. There are various DNA elements that regulate temporal and spatial** gene expression, and these can be mixed and matched. As we will see, the enhancers for endocrine hormones such as insulin and for lens-specific proteins such as crystallins both have sites that bind Pax6 protein. But Pax6 doesn't tell the lens to make insulin or the pancreas to make crystallins, because there are other transcription factor proteins that also must bind. It is the combination of transcription factors that causes particular genes to be transcribed.

**6. A gene can have several enhancer elements, each turning it on in a different set of cells**

**7. Enhancers can also be used to inhibit transcription. In some cases, the same transcription** factors that activate the transcription of one gene can be used to repress the transcription of other genes. These "negative enhancers " are also called **silencers.**




binding of cell specific transcription factors

Transcription factors are proteins that bind to enhancer or promoter regions and interact to activate or repress the transcription of a particular gene.

# Transcription factors have three major domains:

DNA-binding domain: recognizes a particular DNA sequence

Trans-activating domain: activates or supresses the transcription of the gene whose promoter or enhancer it has bound

Protein-protein interaction domain: allows the transcription factor's activity to be modulated by other regulatory proteins

Gene activity depends on

TF concentation,

TF quality (phosphorylated or not)

and TF combinations.

# Structural motifs present in the DNA binding domains

Leucine zipper

Helix-turn-helix, HTH

Zn-finger

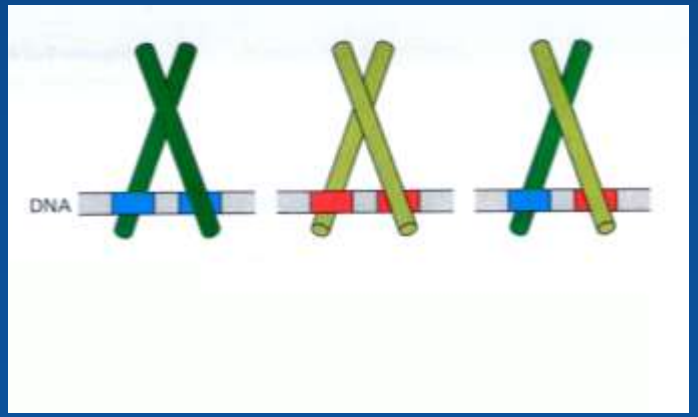
Helix-loop-helix, HLH

# Leucine-zipper

-the dimeric Y shaped protein contains two extended $\alpha$-helices that ”grip” the DNA, much like a pair of scissors at two adjacent major grooves separated by about half a turn of the double helix

-the motif contains the hydrophobic amino acid leucine at every seventh position in the sequence, they are required for dimerisation

-C/EBP, AP1 (liver differentiation, fat cell specification)

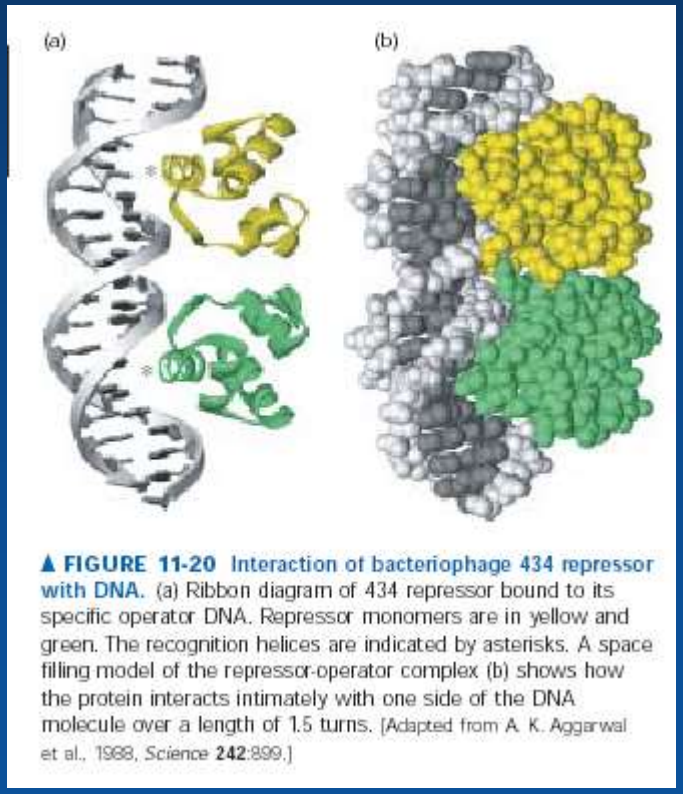


# Helix-turn-helix, HTH

-is present in many bacterial repressor protein

-dimerisation

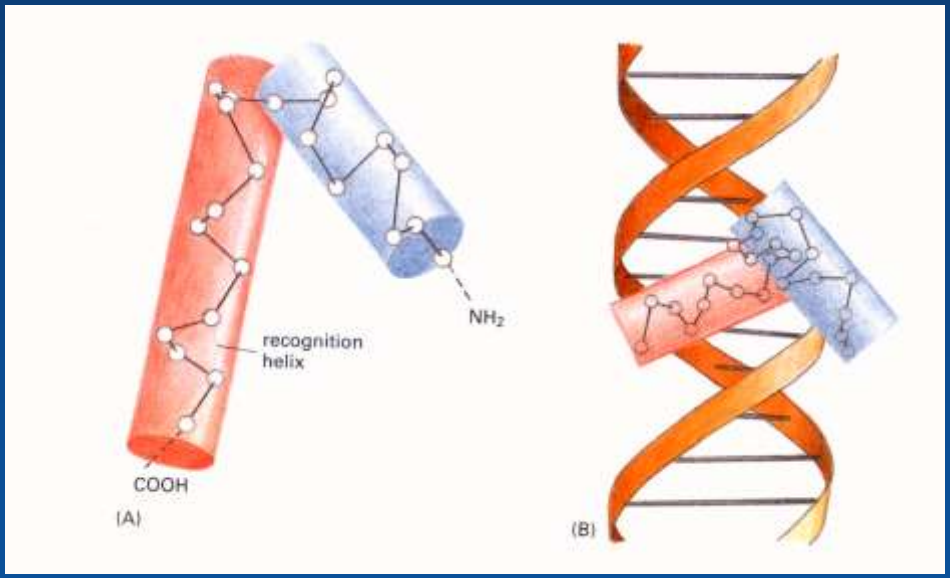
-homeobox, POU (Oct genes, Pit1), paired (Drosoph. paired, Pax genes), winged HTH (c-ets, PU.1)




▲ **FIGURE 11-20** Interaction of bacteriophage 434 repressor with DNA. (a) Ribbon diagram of 434 repressor bound to its specific operator DNA. Repressor monomers are in yellow and green. The recognition helices are indicated by asterisks. A space filling model of the repressor-operator complex (b) shows how the protein interacts intimately with one side of the DNA molecule over a length of 1.5 turns. [Adapted from A. K. Aggarwal et al., 1988, *Science* **242**:899.]

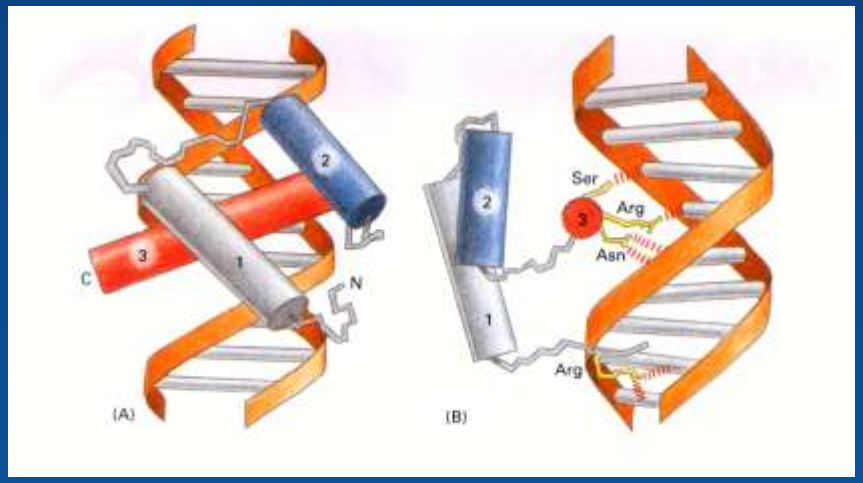
# Helix-turn-helix structure

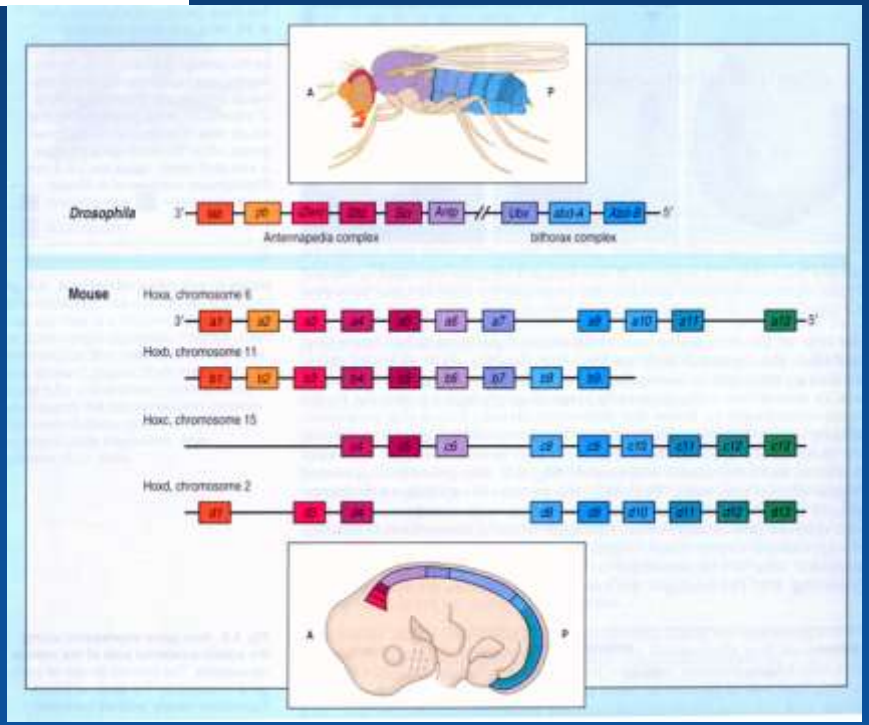
-two $\alpha$-helices, one of them is called recognition helix



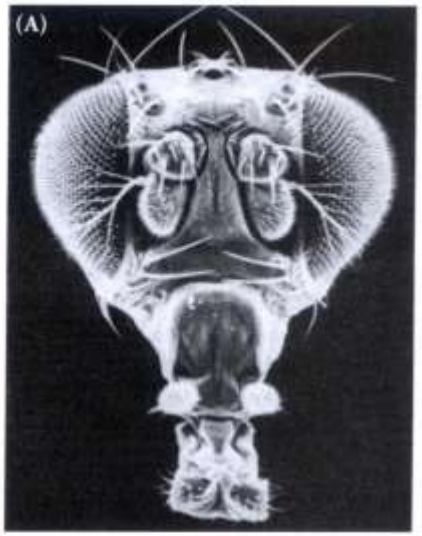
# Homeobox genes

Many eukaryotic transcription factors that function during development contain a conserved 60-residue DNA-binding motif that is similar to the helix-turn-helix motif of bacterial repressors. Called **homeodomain** proteins, these transcription factors were first identified in *Drosophila* mutants in which one body part was transformed into another during development (Chapter 15). The conserved homeodomain sequence has also been found in vertebrate transcription factors, including those that have similar master control functions in human development.

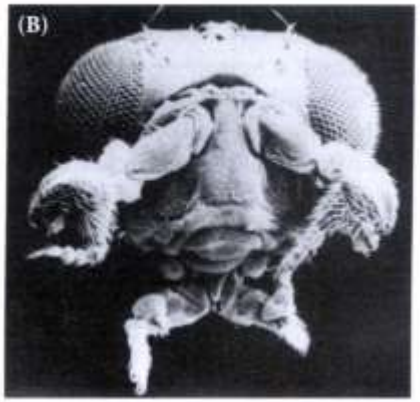


# Homeotic transformation in Drosophila



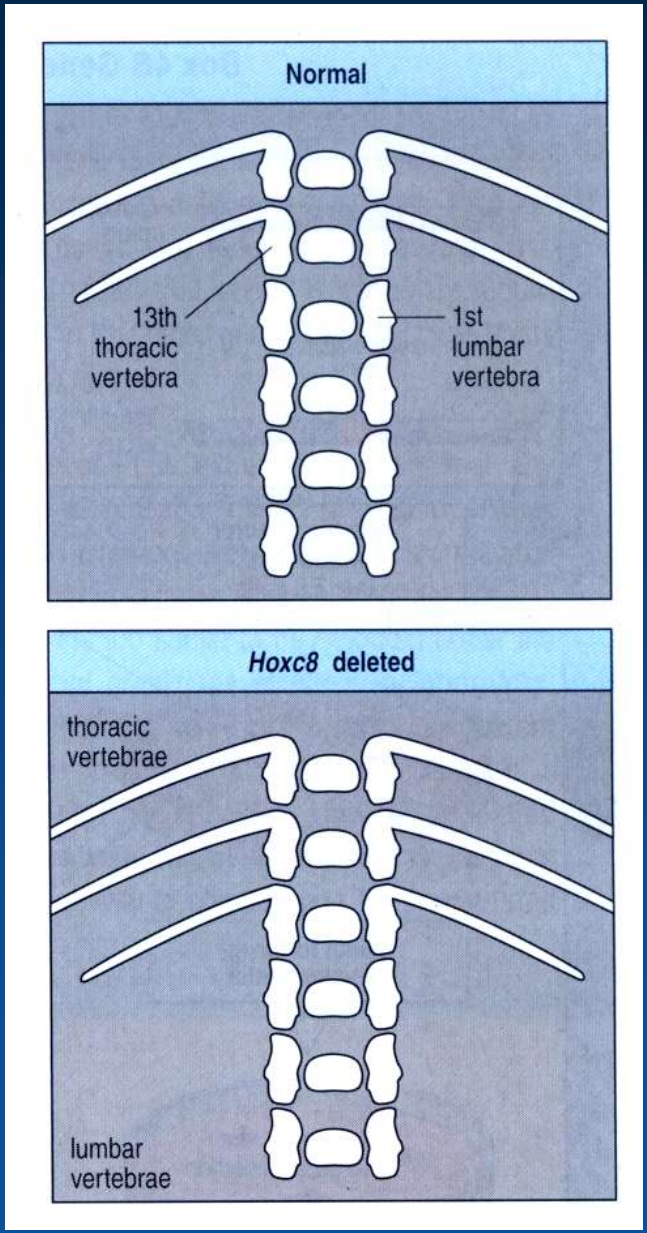

Wild type




Antennapedia mutant, flies with this mutation have the antennae converted into legs

Bithorax/postbithorax mutant
Halteres are converted into wings

# Homeotic transformation in the mouse



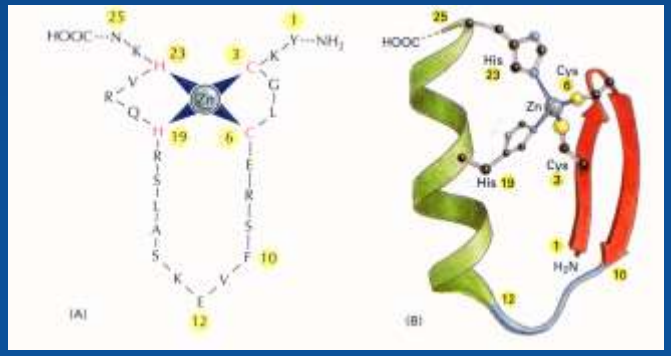

In loss-of-function homozygous mutants of Hoxc8, the first lumbar vertebra is transformed into a rib-bearing thoracic vertebra. The mutation has resulted in the transformation of the lumbar vertebra into a more anterior structure.

# Zinc finger

-number of different eukaryotic protein have regions that fold around a central $Zn^{2+}$ ion, producing a compact domain from a relatively length of the polypeptide chain

-$\alpha$-helix recognise the DNA

-WT1, Krüppel, Engrailed (kidney, gonad and macrophage development, Drosophila segmentation)
steroid receptors

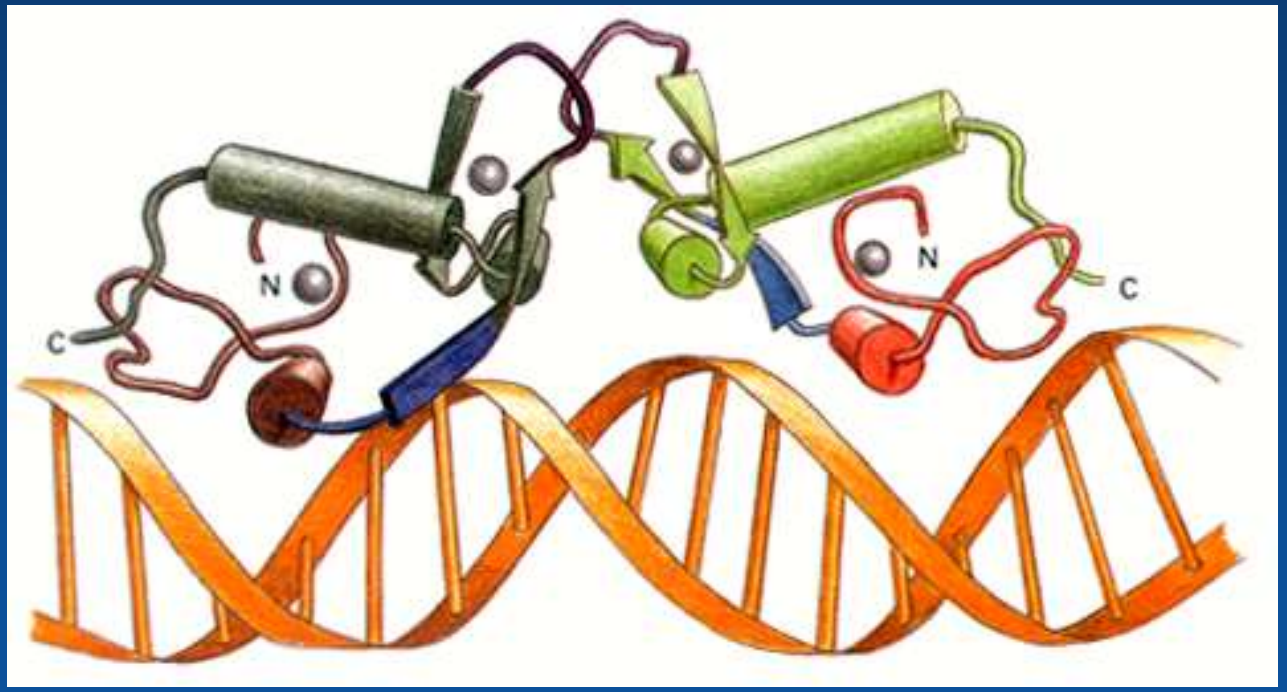


# Steroid receptor superfamily

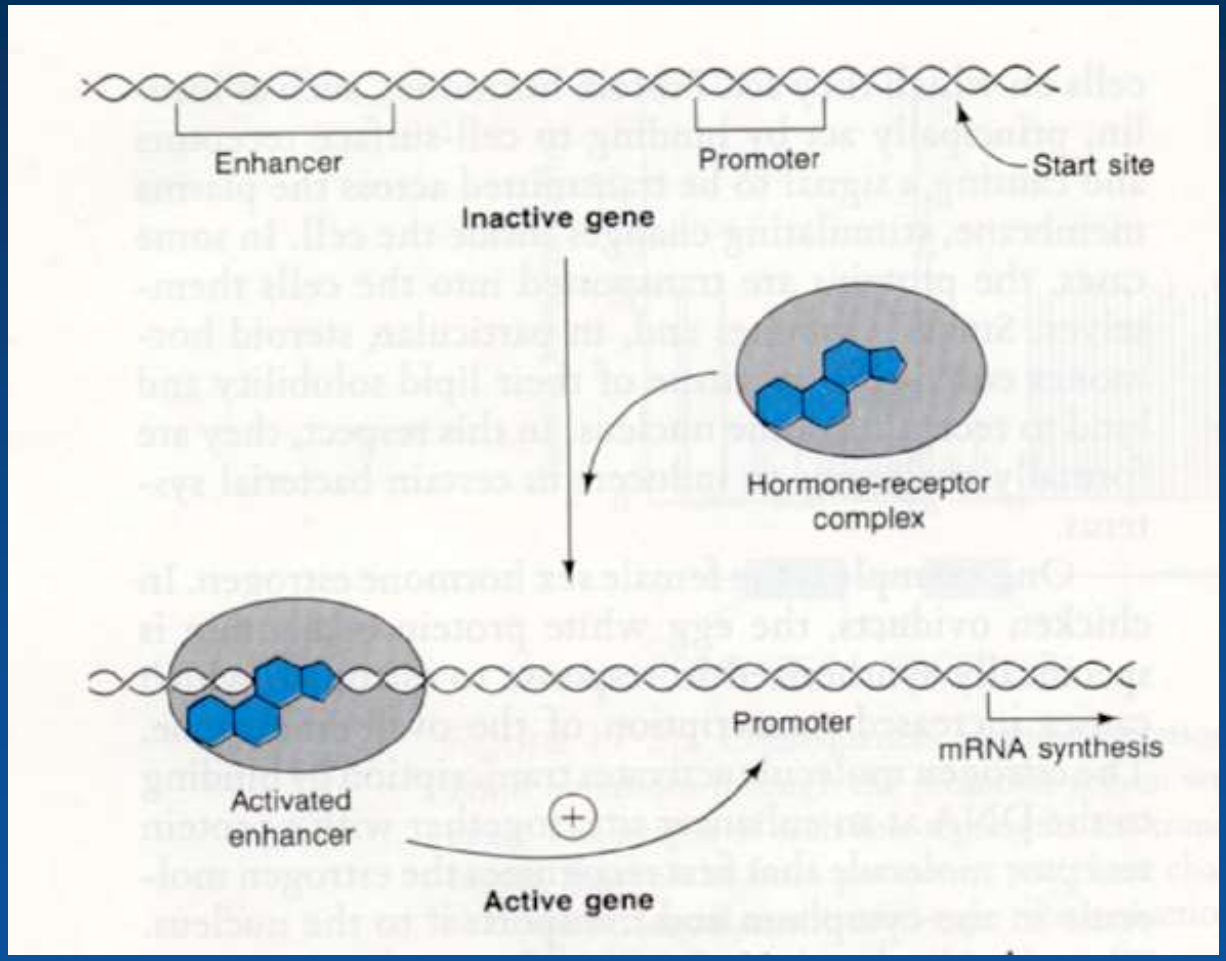
-intracellular (nuclear) receptors

-C4 Zn-finger, presence of two groups of four critical cysteine bounded by the $Zn^{2+}$

-homo or hetero dimerisation



# Steroid hormon/receptor complexes as special cases of trans factors



# GATA factors

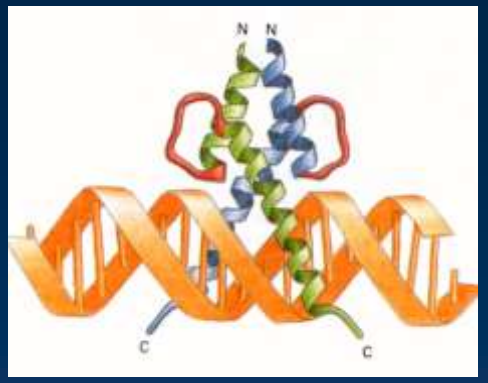
-GATA-1, 2, 3 (embryonic hematopoiesis)

-GATA-4, 5, 6 (heart development)

-two C2H2 Zn-finger, which recognize the WGATAR (W=A or T and R=A or G), or the CGATGG and AGATTA sequences

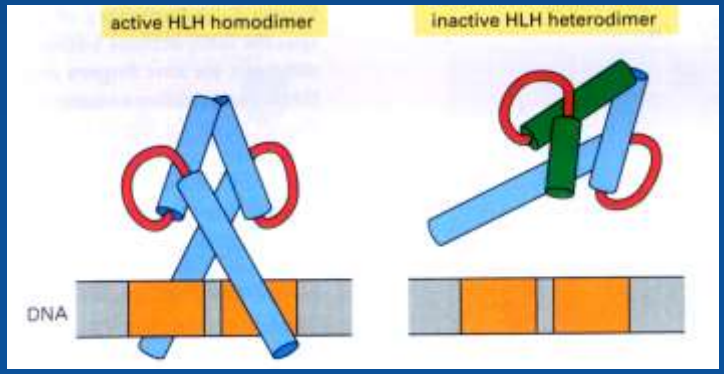
-only the Zn-finger is conserved between family members

# Helix-loop-helix motif

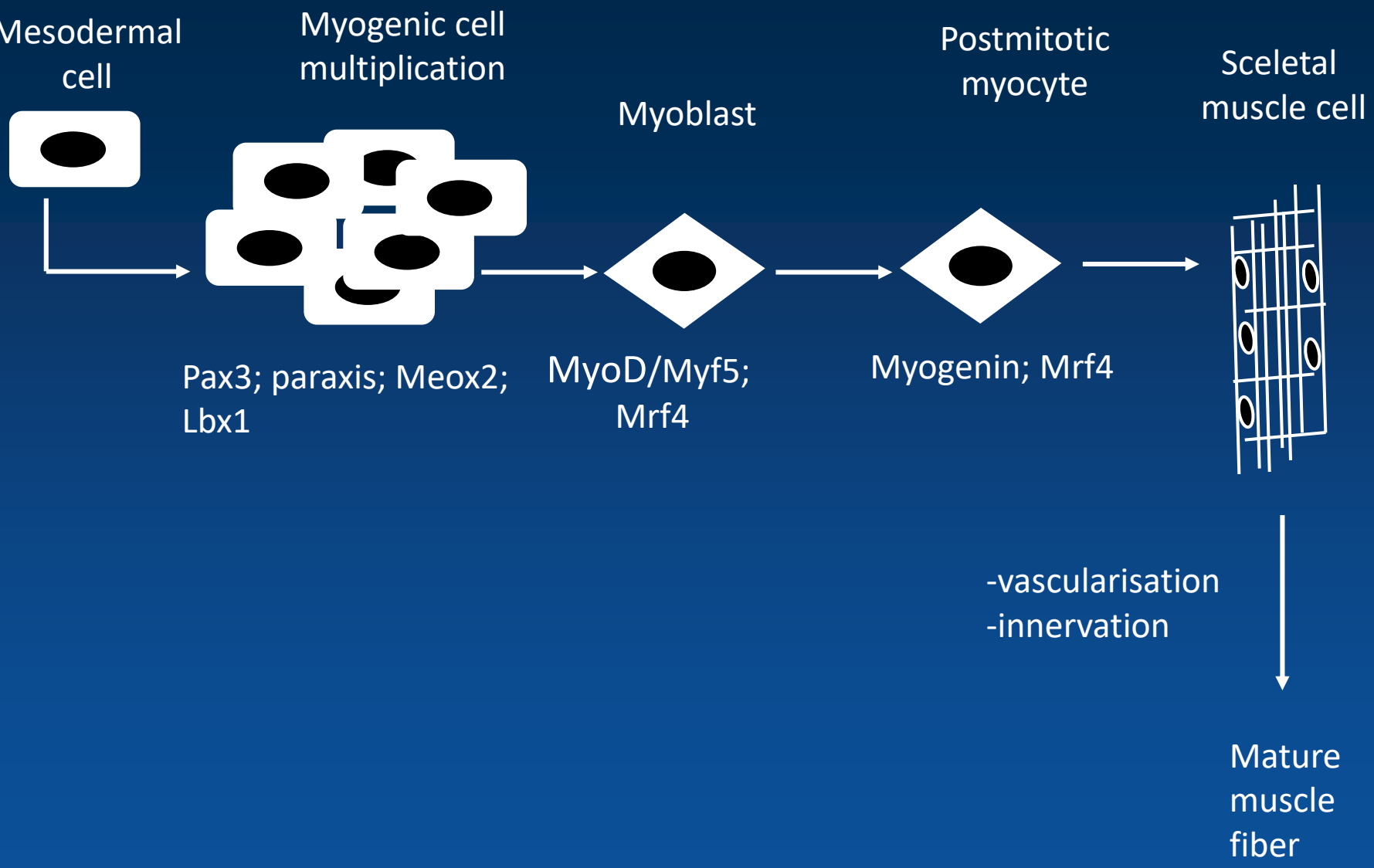



Basic Helix-Loop-Helix (bHLH) Proteins The DNA-binding domain of another class of dimeric transcription factors contains a structural motif very similar to the basic-zipper motif except that a nonhelical loop of the polypeptide chain separates two α-helical regions in each monomer (Figure 11-22b). Termed a basic helix-loop-helix (bHLH), this motif was predicted from the amino acid sequences of these proteins, which contain an N-terminal α helix with basic residues that interact with DNA, a middle loop region, and a C-terminal region with hydrophobic amino acids spaced at intervals characteristic of an amphipathic α helix. As with basic-zipper proteins, different bHLH proteins can form heterodimers.

lehetővé tesz



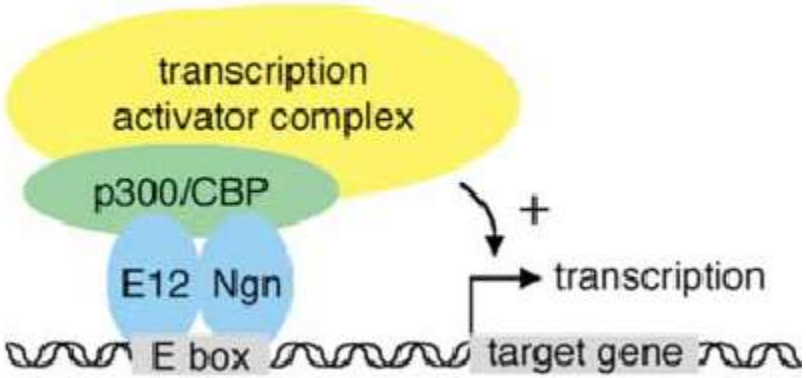
# Orderly appearence of TFs during sceletal muscle development

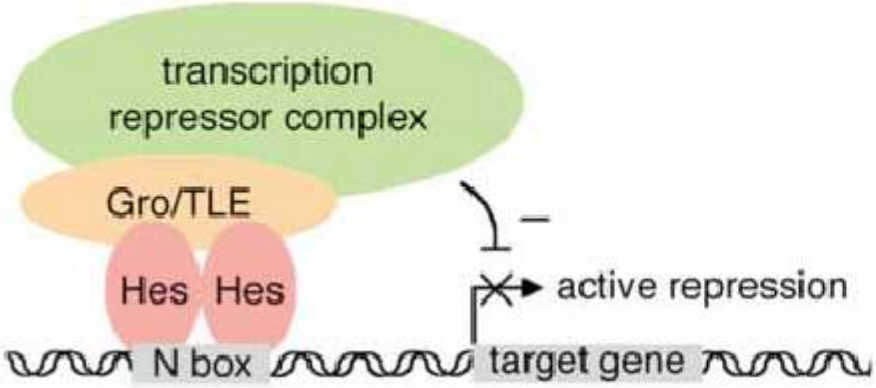


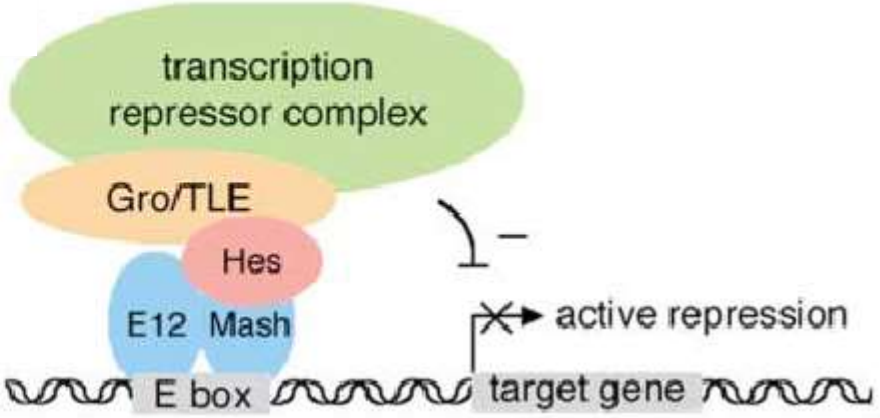

**After Shahragim Tajbakhsh Current Opinion in Genetics & Development 2003, 13:413–422 and Charge´Sophie B. P. and Michael A. Rudnicki. *Physiol Rev* 84: 209–238, 2004**

# bHLH transcription family in neurogenesis

| bHLH Factor | Related Factor in Drosophila | Function | DNA Element |
|---|---|---|---|
| NeuroD | Atonal* | activate transcription | E box |
| Ngn | Tap/Biparous | | |
| Olig | Oli | repressor? | |
| Mash | Achaete-Scute | activate transcription | |
| E protein | Daughterless | | |
| Id | EMC | sequester E protein | no DNA binding |
| Hes | Hairy/(E/Spl) | repress transcription | N box |



# TF hierarchy during hematopoiesis

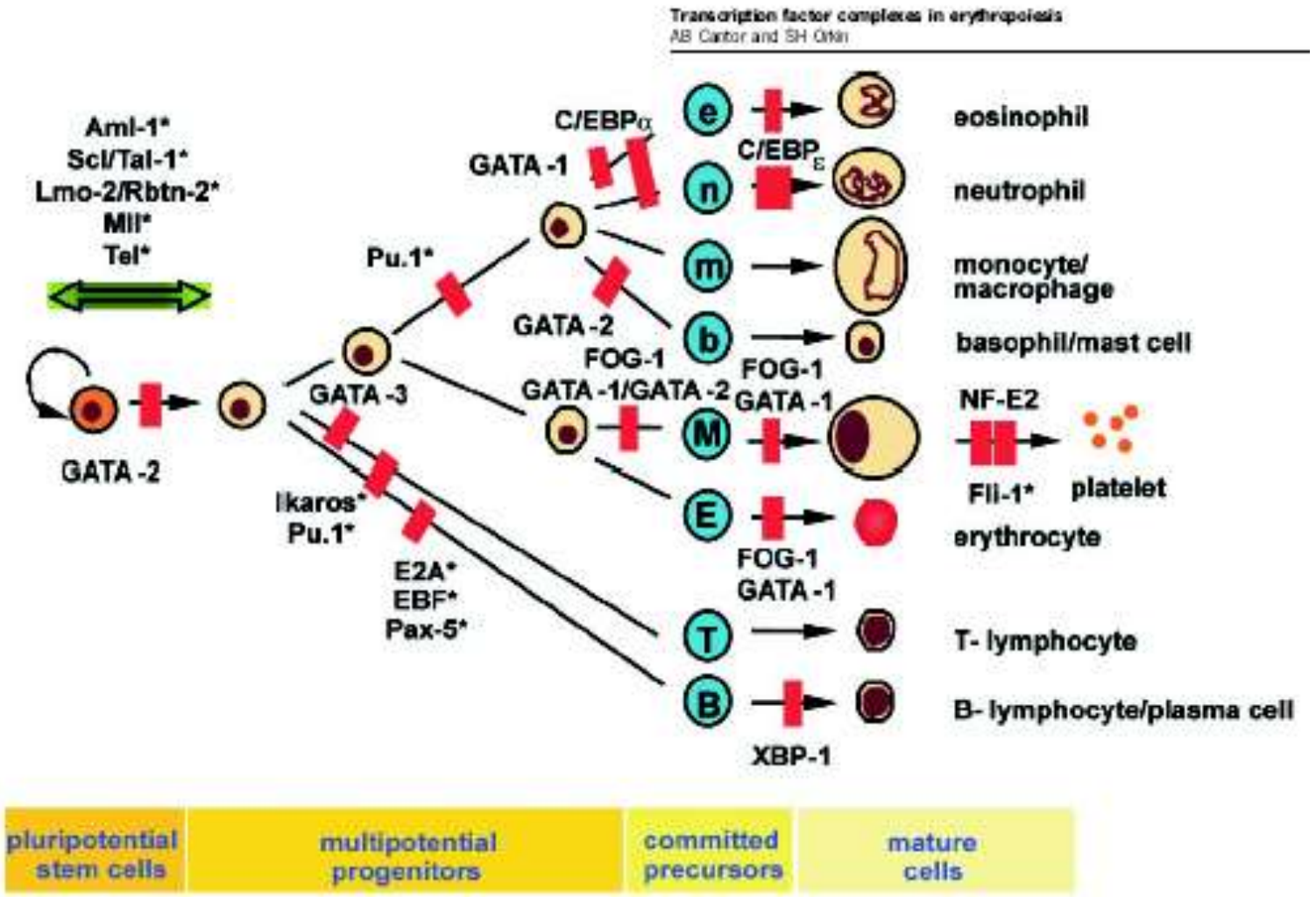
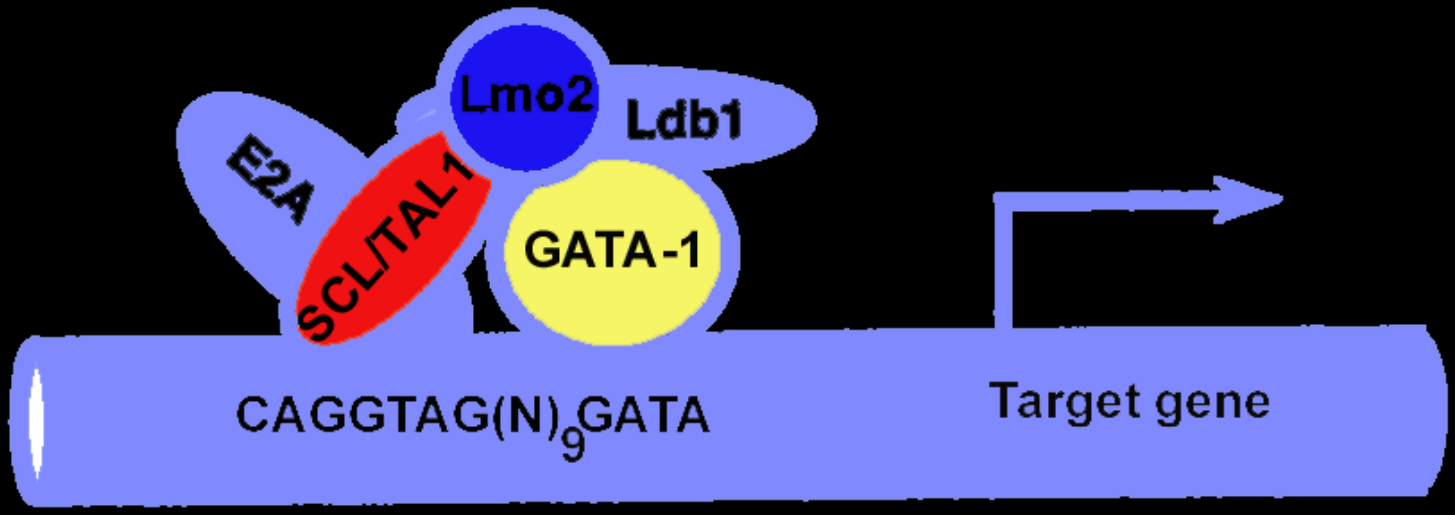



**Figure 1** Transcription factor requirements in hematopoiesis. Schematic representation of hematopoietic lineage pathways from pluripotential stem cells to mature blood elements. Red bars represent the location of maturation arrest observed in the absence of the corresponding transcription factor. Transcription factors associated with chromosomal translocations or viral insertions in human and murine leukemias are denoted by an asterisk. e, eosinophil; n, neutrophil; m, monocyte/macrophage; b, basophil/mast cell; M, megakaryocyte; E, erythrocyte; T, T-lymphocyte; B, B-lymphocyte

# Combinatorial action of transcription factors

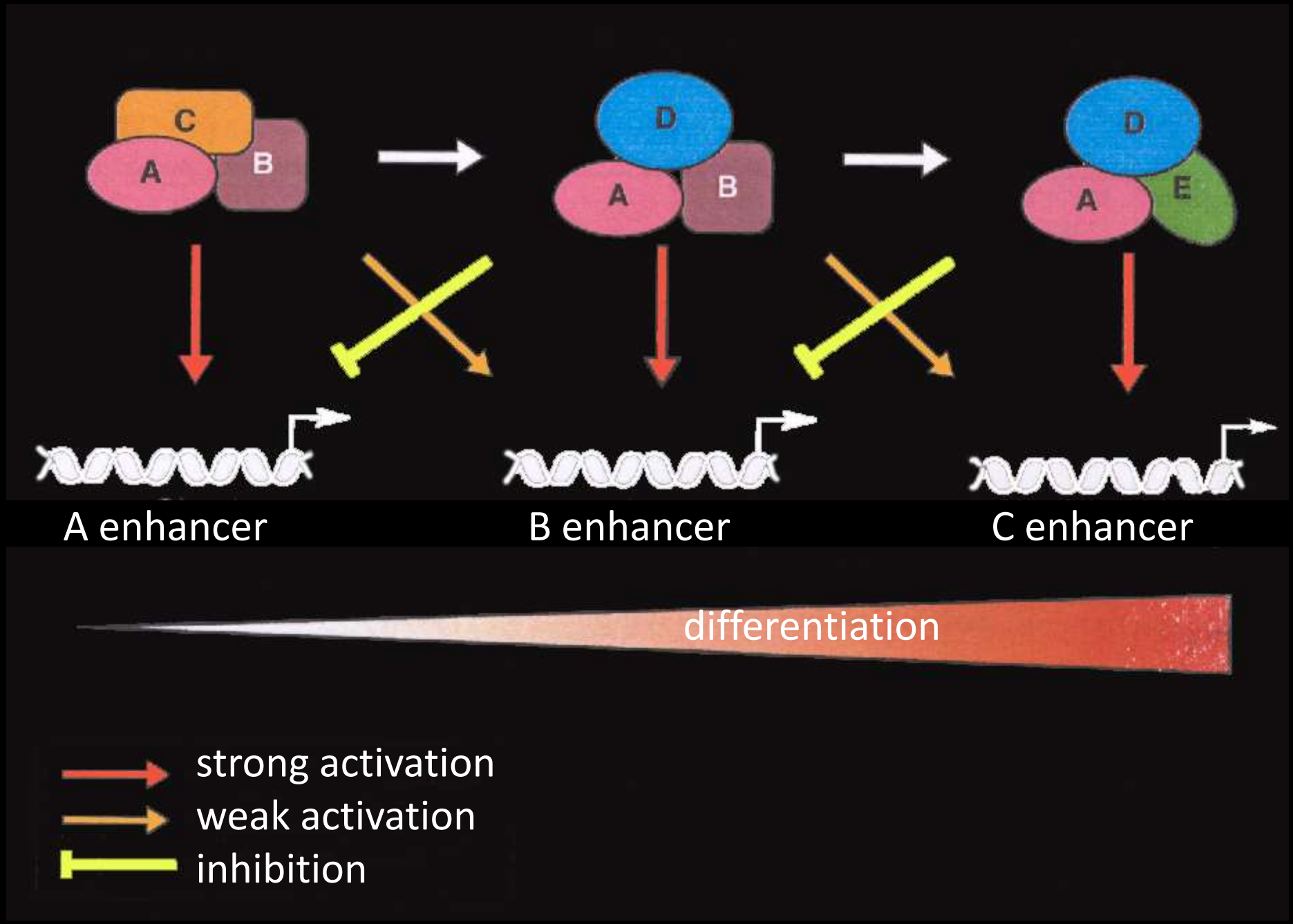
**Fig. 8.** Model of the Lmo2-containing oligomeric DNA-binding complex. The oligomeric complex binds to the E-box–GATA motif with a restricted spacing of 9 bp (at least as determined by CASTing with the R76 oligonucleotide). The stoichiometry of the complex is unknown, but DNA-binding modules within the oligomeric complex are provided by E47–TAL1 heterodimer (E-box binding) and GATA-1 (GATA binding). The Lmo2 LIM-only protein, together with Ldb1, links the two DNA-binding moieties. Target genes with E-box–GATA motif(s) provide recognition sites for the oligomeric complex. Transcriptional transactivation of these genes is a possible consequence of binding to target genes, but repressive functions may also occur, perhaps depending on the stage of haematopoiesis at which binding occurs.




Wadman et al. 1997 The EMBO Journal 16: 3145–3157

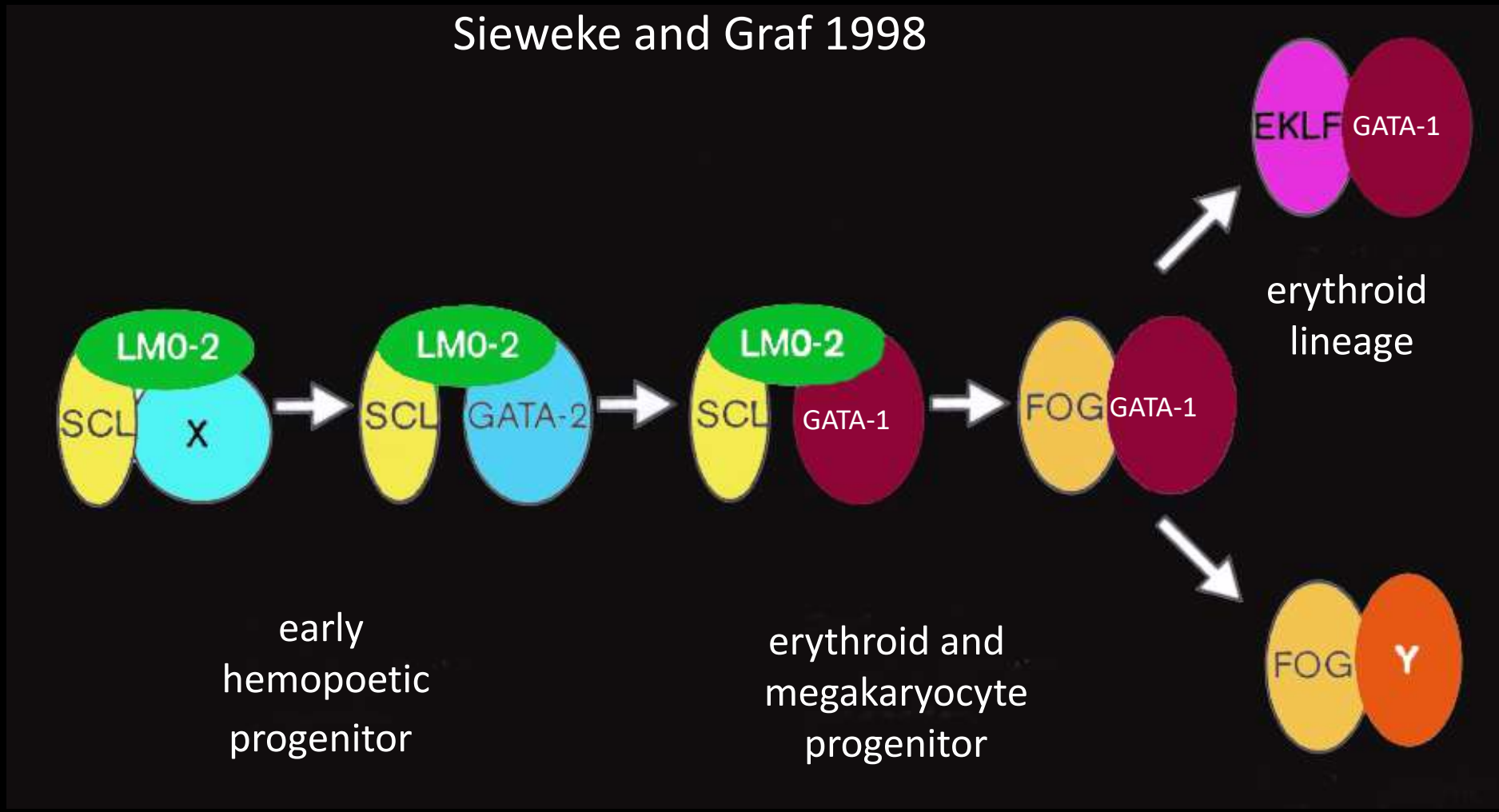
# The „coctail party” model

In the commitment of cell lineages, lineage-specific transcription factors (TF), working with general TFs play important roles. In addition, lineage-specific gene expression appears to be regulated not by single master regulators but by the combinations of transcription factors (160). Expression of a given transcription factor may have different consequences in different cell types, in part because of the different transcriptional environment. The combinatorial action of transcription factors is more important in the control of gene expression, which occurs by many physically interacting factors, forming large, multiprotein complexes. During differentiation, lineage specific complexes are selected, and define the direction of commitment. Sieweke and Graf suggested in their model that these complexes obtain new functions during haematopoietic differentiation through successive changes in composition, such as the topic of discussion changes and take other directions as new people join and others leave a cocktail party. The changes in the composition drive to sequential gene expression activations and inhibitions responsible for specific determination and differentiation (161).


C
A
B
D
A
B
D
A
E
A enhancer
B enhancer
C enhancer
differentiation
strong activation
weak activation
inhibition

# The „coctail party” model

Sieweke and Graf 1998



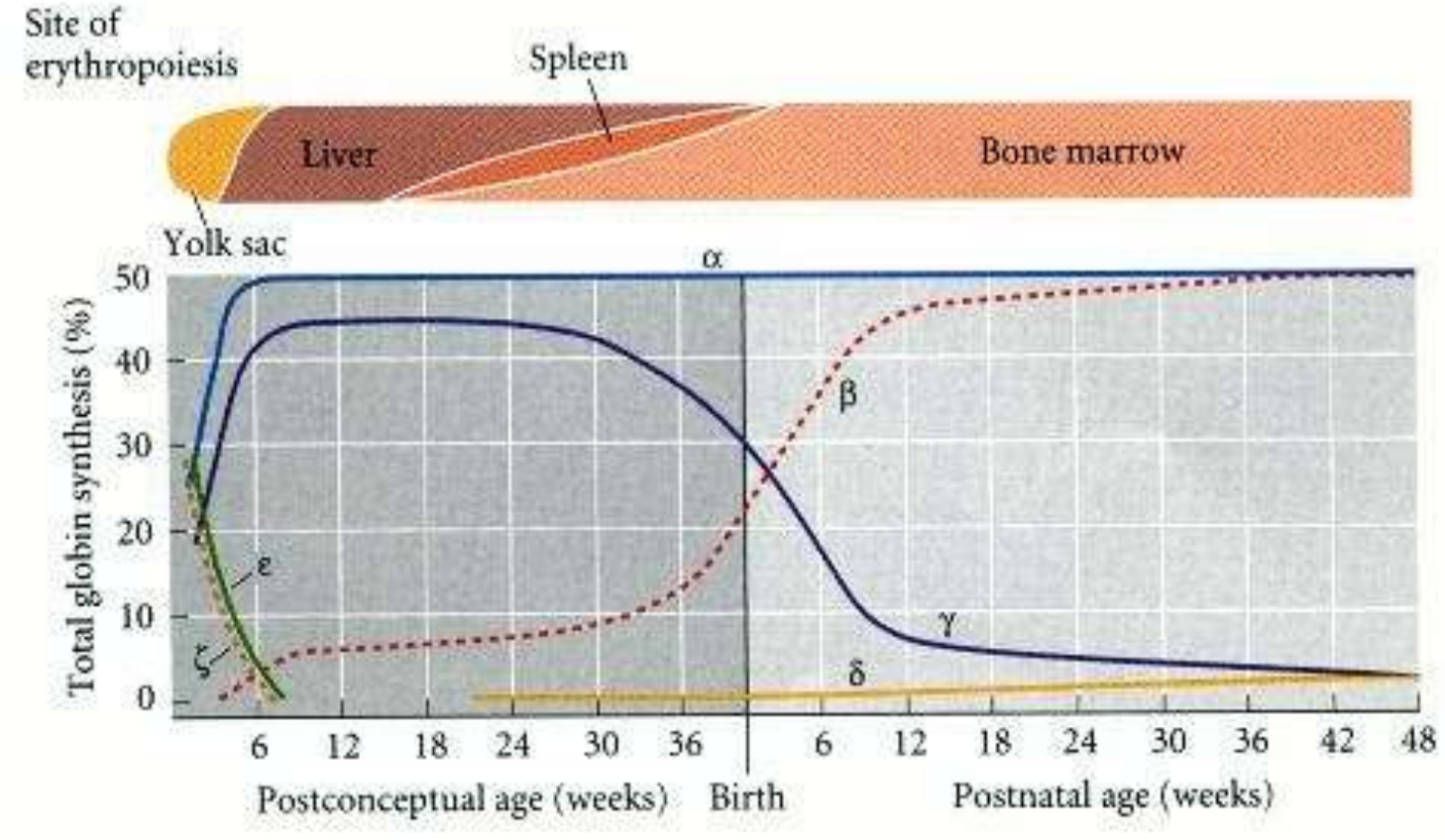
# Return to the cis elements

# „Super-enhancers”: Locus control regions (LCRs)

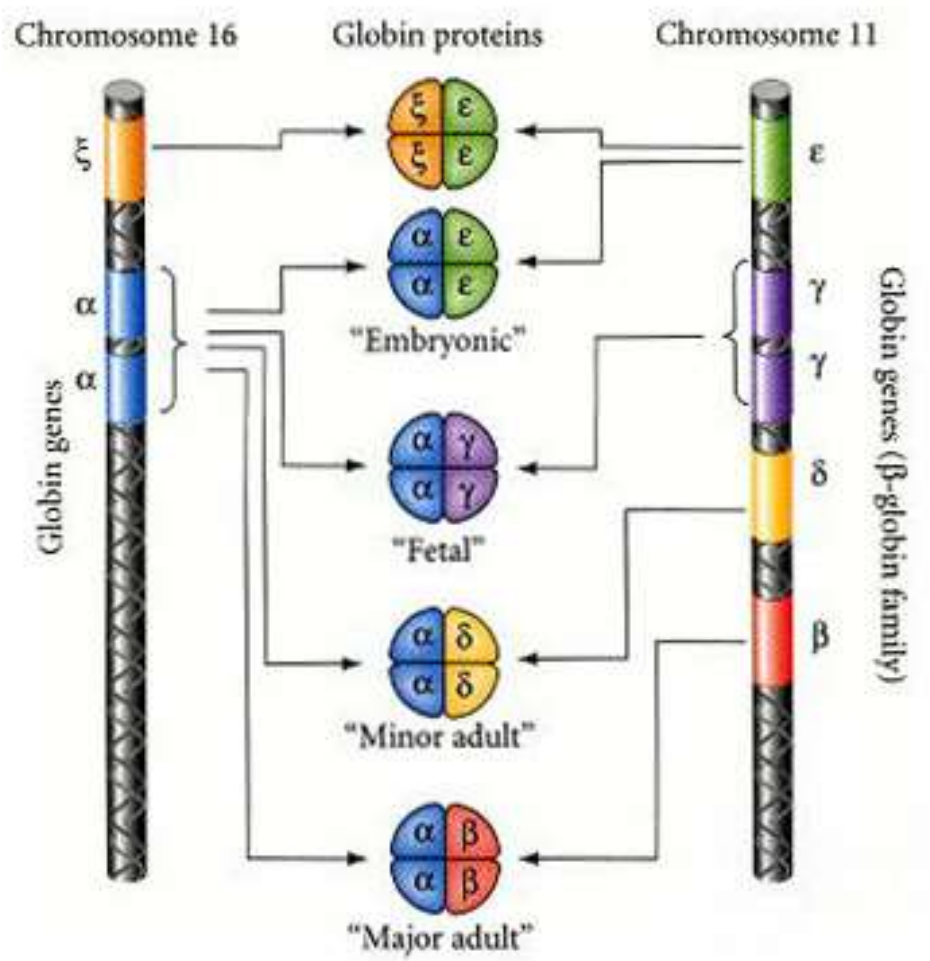
There are some regions of DNA called **locus control regions (LCRs), which function as "super-enhancers." These LCRs establish an "open„** chromatin configuration, inhibiting the normal repression of transcription over an area spanning several genes. The mechanism by which the LCR opens up the chromatin is not yet known.

# Regulation of the β-globin genes as an example

In many species, including chicks and humans, the embryonic or fetal hemoglobin differs from that found in adult red blood cells.



# Human globin genes




The human β-globin gene cluster contains five genes- ε, γG, γG, δ, and β. These are expressed at different times during development. The protein products of all these genes combine with globins encoded by the α-globin complex to form physiologically different hemoglobins at each of these three stages of development.

The discovery of the human globin LCR came from studies of the genetic disease β-**thalassemia.**
**This anemic condition results** from a lack of β-globin and can be caused in several ways.
The usual causes of β-thalassemia involve deletions or mutations in either the coding region of the β-globin gene or its promoter. However**, in some patients, there is a deletion in a region *upstream from the β-globin* gene family, while the genes themselves are normal.** Moreover, without this upstream region, the β-globin family DNA was found to be DNase I-insensitive (van der Ploeg et al. 1980; Kioussis et al. 1983). **DNase I treatment** is used to see whether the DNA in chromatin is accessible to transcription factors. If the DNA in chromatin is not digested by DNase I, it means that DNase I cannot reach it, and therefore, transcription factors couldn't reach it either. Promoters are usually DNase I-sensitive in the cells where they function; and they are usually DNase I-insensitive in those cells where they are not active (Weintraub and Groudine 1976; Stalder et al. 1980).
So it appeared that **there was a region of DNA upstream from the β-globin gene cluster that was responsible for "opening up" the chromatin of the genes, making them accessible to transcription factors. This region of DNA was termed the β-globin locus control region.**

The control region that regulate expression of the β-globin gene cluster are complex and extensive. Each gene has a promoter and control sites immediately upstream (to the 5' side) of the transcription starting point, and there is also an enhancer downstream (to the 3' side) of the β-globin gene, which is the last gene in the cluster. But these local control sequences, which contain binding sites for transcription factors specific for erythroid cells, as well as for other more widespread transcriptional activators, are not sufficient to provide properly regulated expression of the β-globin genes.

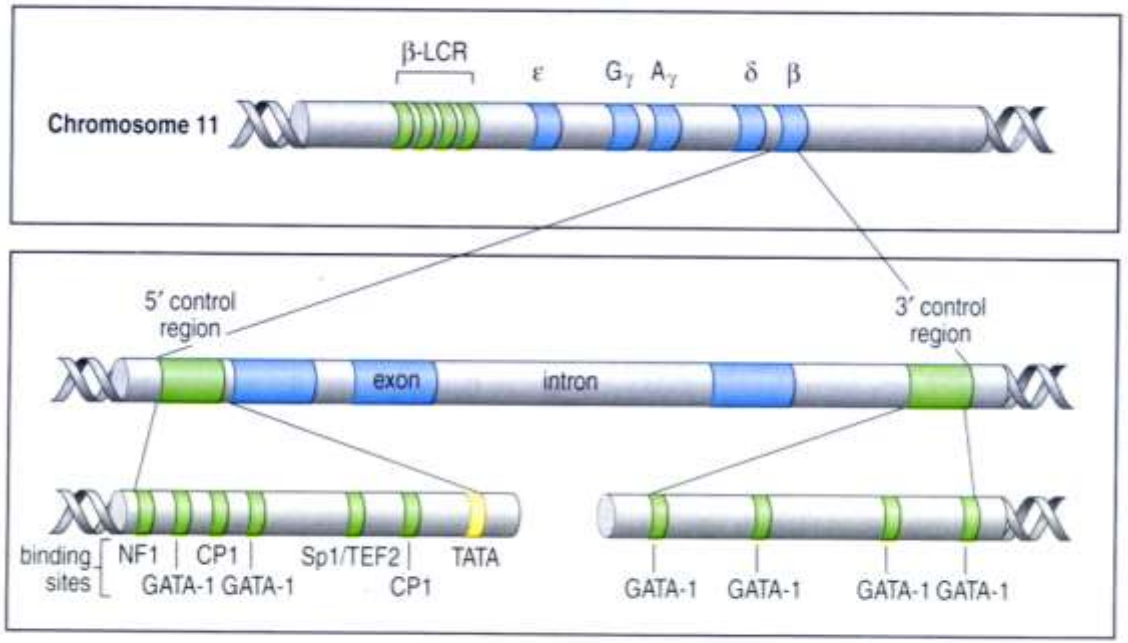
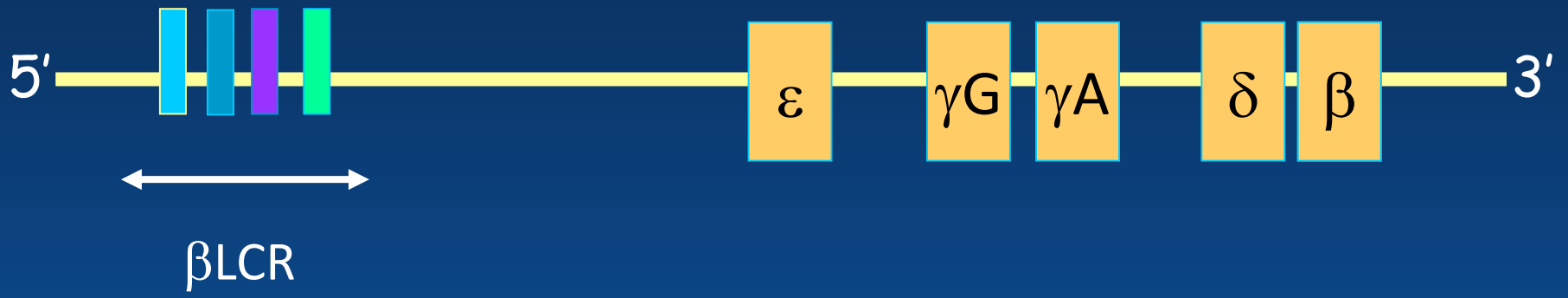



**Fig. 9.25 The control regions of the β-globin gene**. The β-globin gene is part of a complex of other β-type globin genes and is only expressed in the adult. Binding sites for the relatively erythroid-specific transcription factor GATA-1, as well as for other non tissue-specific transcription factors, such as NF1 and CP1, are located in the control regions. Upstream of the whole β-globin cluster, in the locus control region (LCR), are additional control regions required for high-level expression and full developmental regulation of the β-globin genes.

# The role of the LCR (locus control region) in the transcription of the β-globin genes




The locus control region for the β-globin gene complex is located far upstream from its most 5´ member. This LCR contains four sites that are DNase I-hypersensitive only in erythroid precursor cells. Sites are said to be **DNase-hypersensitive** when the DNA in the chromatin can be digested there with only small amounts of DNase I. In most instances, a site is thought to be DNase I hypersensitive when it lacks nucleosomes (Elgin 1988). These sites in the LCR are therefore within nucleosome coils in most cells' nuclei, but in the precursors of the red blood cells, this DNA is exposed. The entire LCR is necessary for activating high levels of erythroid cell-specific transcription of the entire β-globin gene family on human chromosome 11 (Grosveld et al. 1987). Deletion or mutation of the LCR causes the silencing of all these genes.

# How does it work?

Model for the control of globin-gene switching envisages an intercation of LCR-bound proteins with proteins bound to the promoters of succesive globin genes.
The DNA between the LCR region and the globin genes is thought to loop in such a way that proteins binding to the LCR can physically interact with proteins bound to the globin gene promoters.

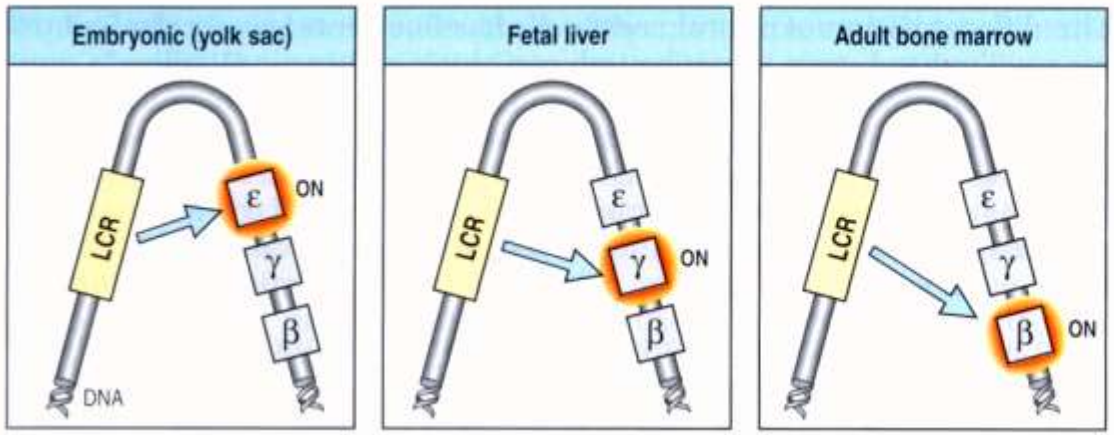



**Fig. 9.26 A possible mechanism for the successive activation of β-family globin genes by the LCR during development.** The LCR (locus control region) is thought to make contact with the promoter of each gene in succession, at different stages of development, thus controlling their temporal expression. After Crossley, M., *et al.*: 1993.

Mol. Cells 34, 1-5, July 31, 2012
DOI/10.1007/s10059-012-0048-8
pISSN: 1016-8478 eISSN: 0219-1032

Molecules and Cells
http://molcells.org

Minireview

# Chromatin Loop Formation in the β-Globin Locus and Its Role in Globin Gene Transcription

AeRi Kim[1], and Ann Dean*

**Although linearly distant along mouse chromosome 7 and human chromosome 11, the mammalian β-globin gene is located in close proximity to the upstream locus control region enhancer when it is actively transcribed in the nuclear chromatin environment of erythroid cells. This organization is thought to generate a chromatin loop between the LCR, a powerful enhancer, and active globin genes by extruding intervening regions containing inactive genes. Loop formation in the β-globin locus requires erythroid specific transcriptional activators, co-factors and insulator-related factors. Chromatin structural features such as histone modifications and DNase I hypersensitive site formation as well as nuclear localization are all involved in loop formation in the locus through diverse mechanisms. Current models envision the formation of the loop as a necessary step in globin gene transcription activation, but this has not been definitively established and many questions remain about what is necessary to achieve globin gene transcription activation.**

Recent data about the LCR

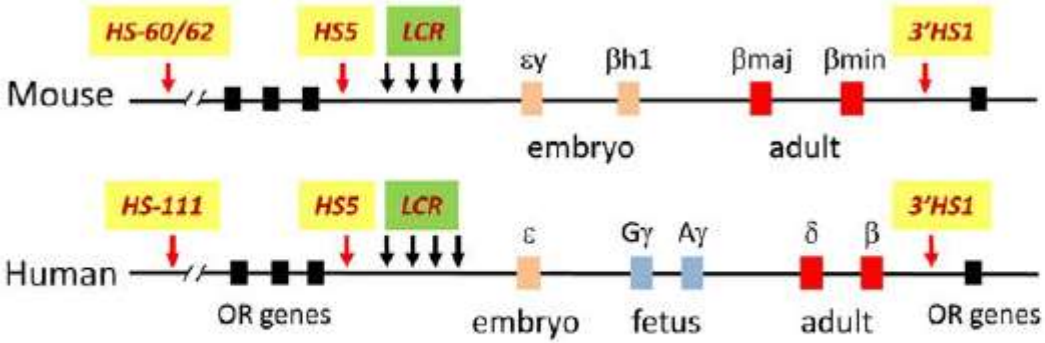



**Fig. 1.** The β-globin loci in mouse and human. The β-globin loci consist of the LCR, the globin genes and insulator elements HS5 and 3′HS1. In the mouse locus, the εy and βH1 are transcribed the in embryo, while the βmaj and βmin are transcribed in the adult. In the human locus, the globin genes are transcribed in three developmental stages; ε in the embryo, Gγ and Aγ in the fetus, and δ and β in adult.

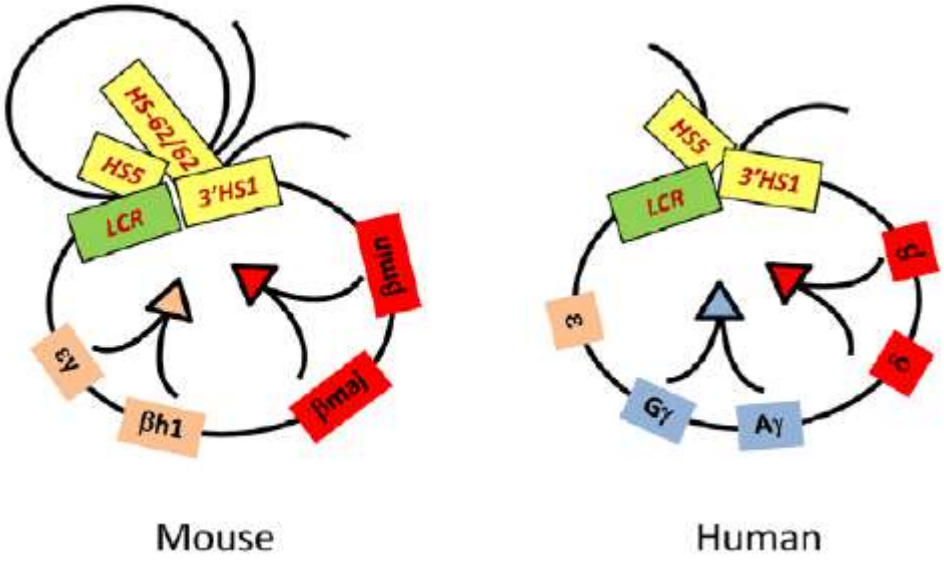



**Fig. 2.** Chromatin loop formation it the β-globin loci. The globin genes are closely positioned with HS-60/62, HS5, LCR and 3′HS1 in the mouse β-globin locus when they are actively transcribed. In the human locus, the fetal genes and adult genes are in close proximity with HS5, LCR, and 3′HS1 when they are transcribed. The close positioning of active genes with HSs makes chromatin loops by extruding regions between them.

Published in final edited form as:
*J Cell Biochem*. 2013 September ; 114(9): 1997–2006. doi:10.1002/jcb.24542.

# Cell-cycle specific association of transcription factors and RNA polymerase II with the human β-globin gene locus†

**Michael Rosenberg**, **Alex Xiucheng Fan**, **I-Ju Lin**, **Shermi Y Liang**, and **Jörg Bungert**
Department of Biochemistry and Molecular Biology, Center for Epigenetics, Genetics Institute, Powell Gene Therapy Center, University of Florida, Gainesville, Florida, 32610

## Abstract

The human β-globin genes are regulated by a locus control region (LCR) and are expressed at extremely high levels in erythroid cells. How transcriptional fidelity of highly expressed genes is regulated and maintained during the cell cycle is not completely understood. Here, we analyzed the association of transcription factor USF, the co-activator CBP, topoisomerase I (Topo I), basal transcription factor TFIIB, and RNA polymerase II (Pol II) with the β-globin gene locus at specific cell-cycle stages. The data demonstrate that while association of Pol II with globin locus associated chromatin decreased in mitotically arrested cells, it remained bound at lower levels at the γ-globin gene promoter. During early S-phase, association of CBP, USF and Pol II with the globin gene locus decreased. The reassociation of CBP and USF2 with the LCR preceded reassociation of Pol II, suggesting that these proteins together mediate recruitment of Pol II to the β-globin gene locus during S-phase. Finally, we analyzed the association of Topo I with the globin gene locus during late S-phase. In general, Topo I association correlated with the binding of Pol II. Inhibition of Topo I activity reduced Pol II binding at the LCR and intergenic regions but not at the γ-globin gene promoter. The data demonstrate dynamic associations of transcription factors with the globin gene locus during the cell cycle and support previous results showing that specific components of transcription complexes remain associated with highly transcribed genes during mitosis.

Biochimica et Biophysica Acta 1829 (2013) 963–969

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

# Histone acetylation contributes to chromatin looping between the locus control region and globin gene by influencing hypersensitive site formation

Yea Woon Kim, AeRi Kim *

*Department of Molecular Biology, College of Natural Sciences, Pusan National University, Busan 609-735, South Korea*

ARTICLE INFO

*Article history:*
Received 22 February 2013
Received in revised form 27 March 2013
Accepted 11 April 2013
Available online 20 April 2013

*Keywords:*
Histone acetylation
Chromatin loop
Hypersensitive site
β-Globin locus
Transcription

ABSTRACT

Chromatin loops are formed between enhancers and promoters and between insulators to regulate gene transcription in the eukaryotic genome. These transcription regulatory elements forming loops have highly acetylated histones. To understand the correlation between histone acetylation and chromatin loop formation, we inhibited the expression of histone acetyltransferase CBP and p300 in erythroid K562 cells and analyzed the chromatin structure of the β-globin locus. The proximity between the locus control region (LCR) and the active $^{G}$γ-globin gene was decreased in the β-globin locus when histones were hypoacetylated by the double knockdown of CBP and p300. Sensitivity to DNase I and binding of erythroid specific activators were reduced in the hypoacetylated LCR hypersensitive sites (HSs) and gene promoter. Interestingly, the chromatin loop between HS5 and 3′HS1 was formed regardless of the hypoacetylation of the β-globin locus. CTCF binding was maintained at HS5 and 3′HS1 in the hypoacetylated locus. Thus, these results indicate that histone acetylation contributes to chromatin looping through the formation of HSs in the LCR and gene promoter. However, looping between insulators appears to be independent from histone acetylation.



INSULATOR? →

[1] M. Bartkuhn, R. Renkawitz, Long range chromatin interactions involved in gene regulation, Biochim. Biophys. Acta 1783 (2008) 2161–2166.
[2] A. Dean, In the loop: long range chromatin interactions and gene regulation, Brief. Funct. Genomics 10 (2011) 3–10.
[3] S. Kadauke, G.A. Blobel, Chromatin loops in gene regulation, Biochim. Biophys. Acta 1789 (2009) 17–25.

Thank you for your attention!

References:

S. F. Gilbert: Developmental Biology, Sinauer associates, Inc. Publishers

B. M. Carlson: Human Embryology and Developmental Biology, 4E, Elsevier

L. Wolpert: Principles of development, Oxford Univ. Press

Molecular biology of the cell, 5th ed., edited by Alberts et al., Garland Science

http://web.wi.mit.edu/young/pub/RNAPolymerase.html

Lectures of ELTE TTK Dept. of Genetics