

MUNI
SCI

Bi4025en

Molecular Biology

Mgr. Jiří Kohoutek, Ph.D.

Lecture 7

- Regulation of gene expression in prokaryotes and eukaryotes.

Why to regulate gene expression?

- The products of all genome genes are not necessary at every moment of a cell's life - mechanisms to ensure gene expression at the right time and in the right place (spatiotemporal regulation).
- Variability of the external environment. Responses to signals from the environment (e.g. temperature, osmotic pressure, nutrient availability,..) and signals from other cells, tissues and organs (e.g. developmental processes, injuries,..).
- Variability of the internal environment of the cell during the cell cycle, cell commitment, differentiation and etc.
- Gene expression is highly energetic process.

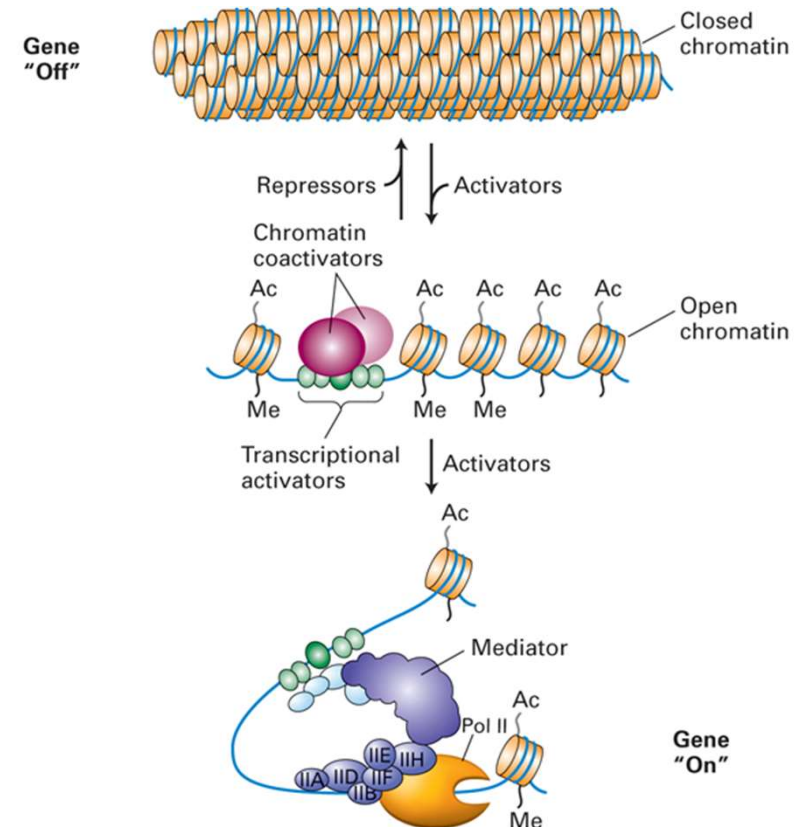
- Regulation of gene expression in the eukaryotes.

Eukaryotic gene expression

- Eukaryotic gene expression (like us!) can be controlled at various stages, from the availability of DNA to the production of mRNAs to the translation and processing of proteins.
- Different genes are regulated at different points, and it's not uncommon for a gene (particularly an important or powerful one) to be regulated at multiple steps.
- Chromatin accessibility.
- Transcription.
- RNA processing.
- RNA stability.
- Translation.
- Protein activity.

Components of eukaryotic gene expression

- **Enhancer/Silencer** – DNA control element far from or close a gene or intron.
- **Activator** – bind to enhancers to turn on transcription of a gene.
 - Transcription factors.
 - Needed for transcription to begin.
- **Repressors** – bind to silencers.
 - Turn off transcription.
 - Block activators from binding to enhancers.

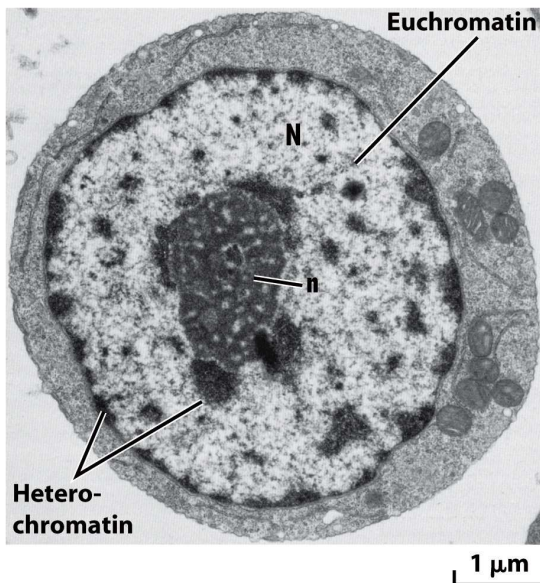
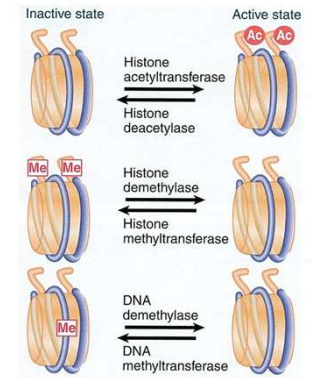


Components of eukaryotic gene expression

- Co-activator and Co-repressor - the distinction between an activator/co-activator and repressor/co-repressor is based on **whether or not the protein binds specifically to DNA**.
- Namely, **activators/repressors have DNA binding domains** that allow them to bind to DNA.
- **Co-activators/co-repressors** typically don't bind to specific sequences in DNA.
- They typically **exert their effects** on transcription initiation **via protein-protein interactions** within transcription initiation complexes at promoters, or by modifying histone tails.

Chromatin accessibility

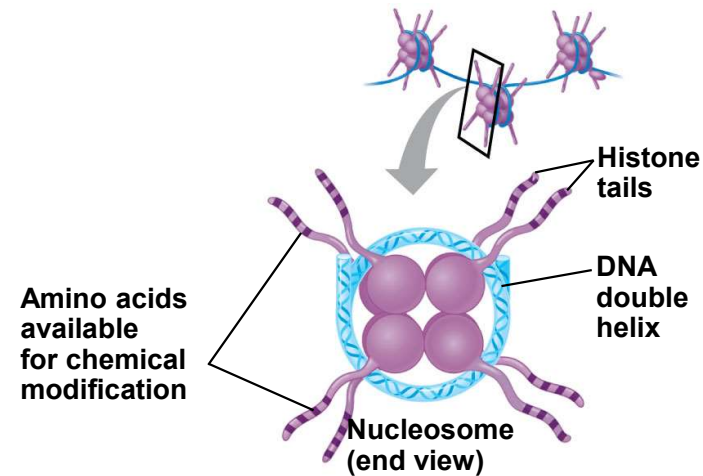
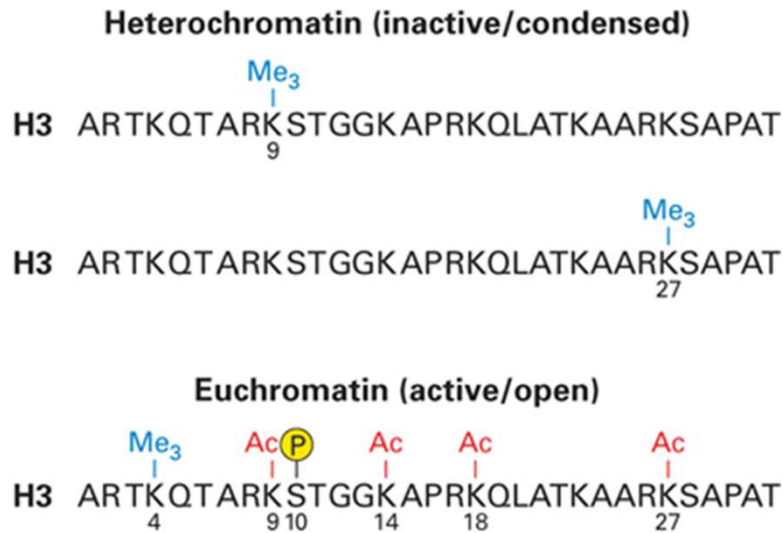
- Chromatin accessibility - the structure of chromatin (DNA and its organizing proteins) can be regulated. More open or “relaxed” chromatin makes a gene more available for transcription.
- Interphase chromatin exists in **two different condensation states**.



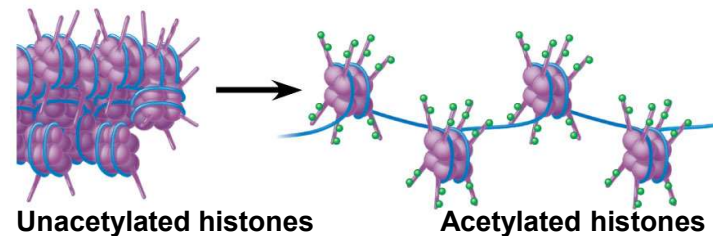
- **Heterochromatin** is a condensed form that has a condensation state similar to chromatin found in metaphase chromosomes. Heterochromatin typically is found at centromere and telomere regions, which remain relatively condensed during interphase.
- **Euchromatin** is considerably less condensed. Most transcribed genes are located in regions of euchromatin.

Chromatin accessibility

- Transition is between state of hetero- and euchromatin is mediated by various modifications of histones.



(a) Histone tails protrude outward from a nucleosome

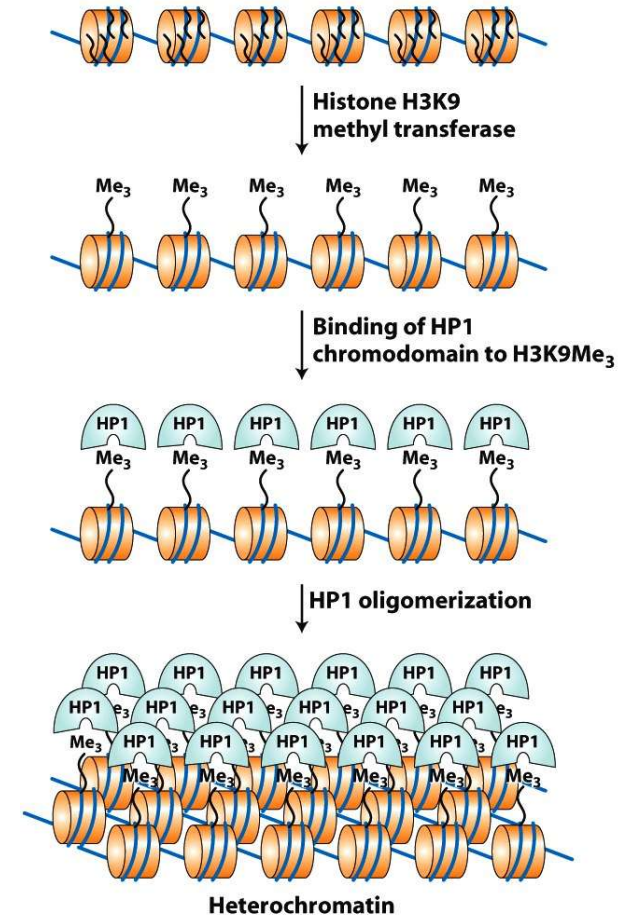


(b) Acetylation of histone tails promotes loose chromatin structure that permits transcription

© 2011 Pearson Education, Inc.

Euchromatin to heterochromatin transition

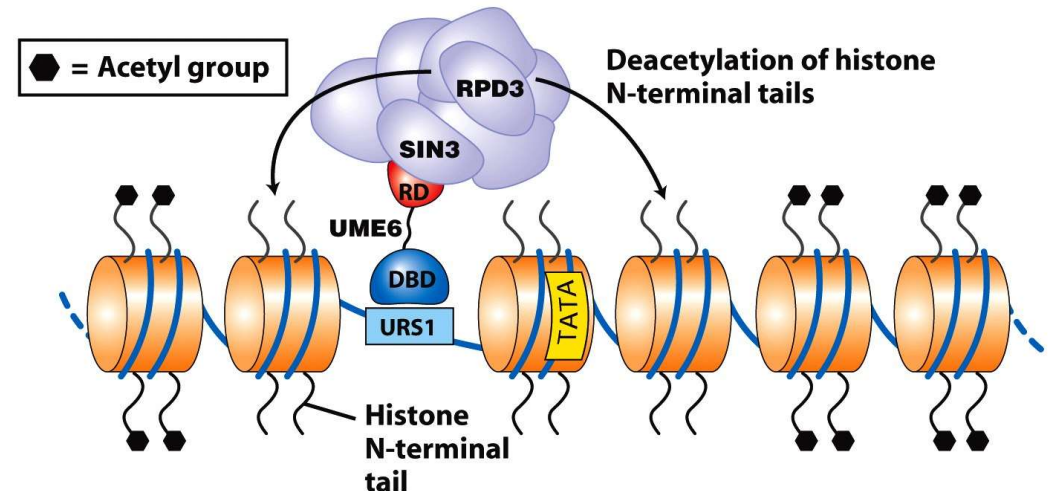
- The trimethylation of histone H3 at lysine 9 (H3K9Me₃) plays an important role in promoting chromatin condensation to heterochromatin.
- Trimethylated sites are bound by heterochromatin protein 1 (HP1) which self-associates and oligomerizes resulting in heterochromatin. Heterochromatin condensation is thought to spread laterally between “boundary elements” that mark the ends of transcriptionally active euchromatin.
- Recruitment of the H3K9 histone methyl transferase (HMT) to HP1 sites promotes heterochromatin spreading by catalyzing H3 methylation.



Euchromatin to heterochromatin transition

- Another mechanism involved in heterochromatin formation from euchromatin is direct **histone deacetylation**.
- The UME6 repressor binds to URS1 control elements and recruits a co-repressor complex containing SIN3 and RPD3 to these sites.
- RPD3 is a histone deacetylase, an enzyme removing acetyl groups from histones in the vicinity of the URS1 sequence. The nucleosomes bound to DNA in this region (which contains a TATA box promoter) subsequently condense, and expression of the gene is repressed.

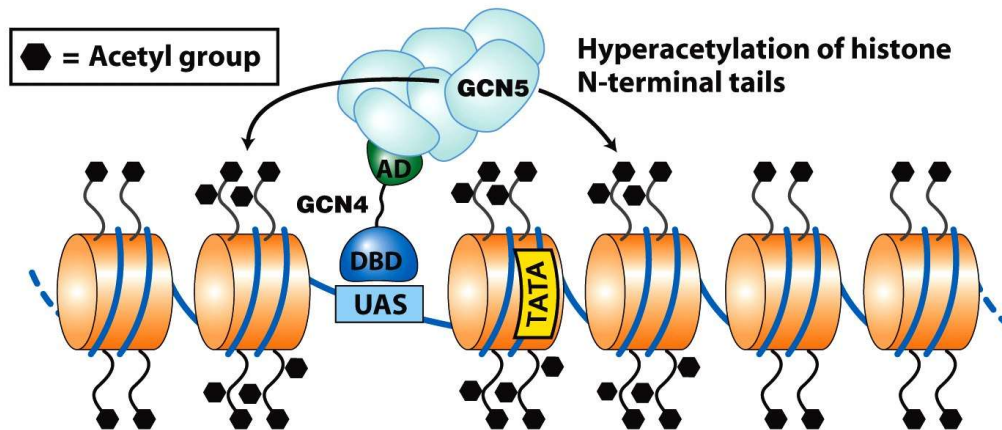
Repressor-directed histone deacetylation



Histone acetylation

- In general, the genes can be turned on by **histone acetylation** and chromatin **decondensation**.
- For instance, GCN4 activator first binds to its UAS upstream of the TATA box of a regulated gene and recruits a co-activator complex containing the GCN5.

Activator-directed histone hyperacetylation



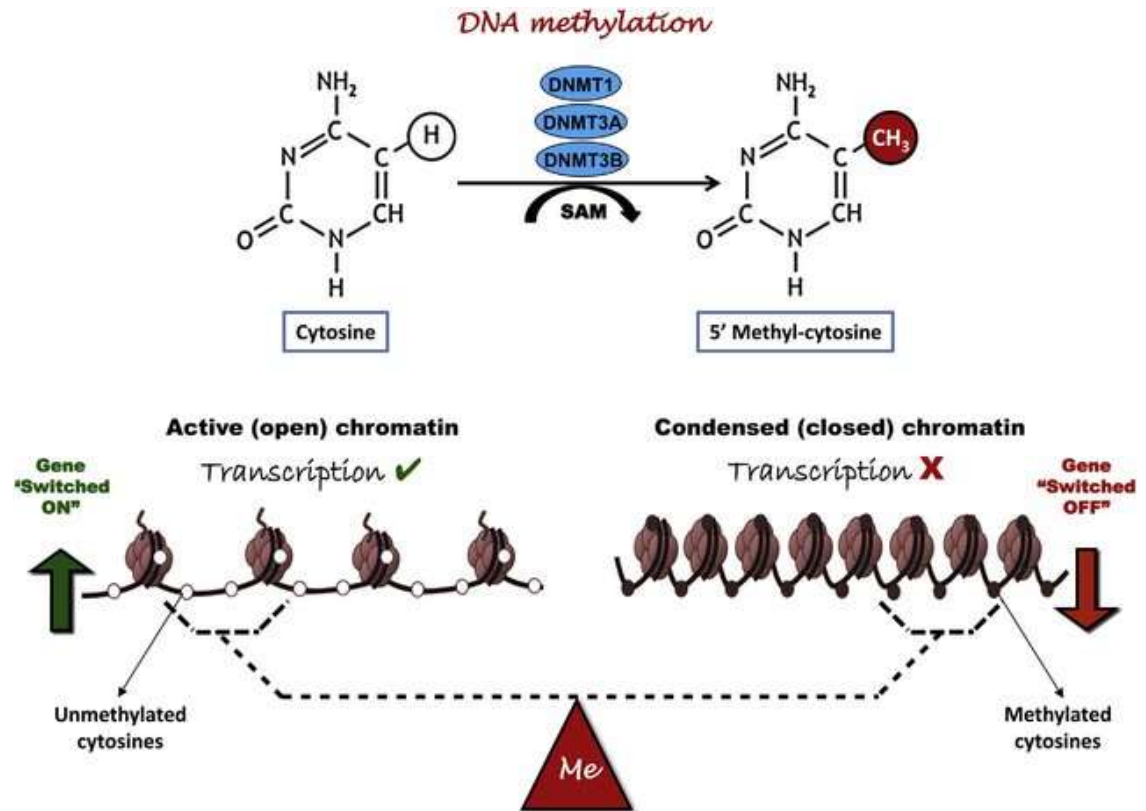
- The GCN5 acetylates histone N-terminal tails.
- **Hyperacetylation of histones** leads to chromatin decondensation.
- General TFs and RNA Pol II are then able to interact with the promoter, and the gene is transcribed.

DNA Methylation

- Cytosine methylation in higher eukaryotes: 10-30%.
- Mediated by DNA methyltransferases (DNMTs).
- Targeted sequence is short: GC in animals and GNC in plants.
- DNA methylation typically weakens gene expression.
- Genes with continuous transcription mostly do not have GC methylated islands.
- Methyl groups protrude into a large DNA groove „ thus preventing proper binding of transcription factors“.

DNA Methylation

- The process of DNA methylation involves the transfer of **methyl group** from S-adenosylmethionine (SAM) to the **C-5 position of cytosine**, catalyzed by DNA methyltransferases.
- DNA methylation is an "epigenetic switch" that regulates the balance between "open" and "closed" form of chromatin by changing the interactions between DNA and protein.



Transcription - activity of eukaryotic cell

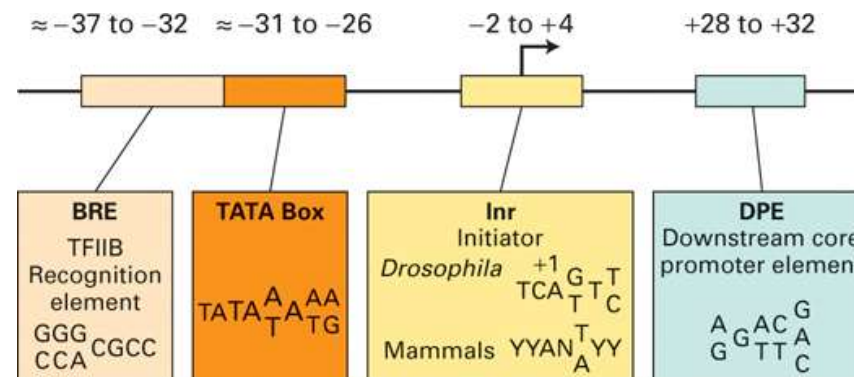
- Transcription is a key regulatory point for many genes. **Sets of transcription factor proteins** bind to specific DNA sequences in or near a gene and **promote or repress** its transcription into an RNA.
- Unlike prokaryotes, eukaryotic genes are not completely turned on or off, but there is modulation of transcription.
- **Basal transcription** - with the participation of basal TF, transcriptional levels, minimum level of transcription.
- **Constitutive transcription** - with the participation of basal and constitutive TFs, allowing different transcription rates of different gene.
 - **General TF = basal + constitutive, activate operational genes**
- **Induced transcription** - transcription regulated by inducible specific TFs which activity is influenced by stimuli from the external environment.
 - **Specific TF= cellular and time-specific regulatory proteins.**

Requirements for transcription initiation

- Putting RNA-polymerase in the active state.
- Binding of TF to the promoter (with the participation of activators and coactivators, necessary to create a pre-initiation complex.
- Binding of specific (inducible) TFs to transcription enhancers with unique response sequences (RE).
- TF interaction allowing the transcription promoter and enhancer to interact.
- Active RNA polymerase state.

Promoter structure

- The **promoter** of a eukaryotic gene can be defined as a sequence that **sets the transcription start site** for RNA polymerase.
- Strong RNA Pol II promoters contain an A/T rich sequence known as the **TATA box** located 26-31 bp upstream of the start site.
- Other genes have alternative sequence elements known as **initiators** (Inr) which also serve as promoters that set the RNA Pol II start site.
- Finally, **CG-rich repeat sequences** (CpG islands) are used by RNA Pol II as promoters in 60-70% of genes. Most of these genes are weakly expressed.



RNA polymerase II require five types of proteins

- Basal (general) transcription factors.
- Architectural regulators to facilitate DNA looping.
- Transcription factors, transcription activators.
 - Proteins that bind to upstream or downstream activator sequences (UASs).
- Chromatin modification/remodeling proteins.
- Coactivators
 - Act indirectly (with other proteins, not with DNA).

Transcription initiation by RNA Pol II

- RNA Polymerase II (RNA Pol II) requires general TFs in addition to tissue-specific transcription factors for transcription of most genes *in vivo*.
- **General transcription factors**, TFs, position RNA Pol II at start sites and **assist the enzyme in melting promoter DNA**. General TFs are highly conserved across species. The general TFs used at TATA box.
- Architectural regulators facilitates **DNA looping**. Such as, TFIID consists of TBP (TATA box binding protein) and 13 TBP-associated factors (TAFs).

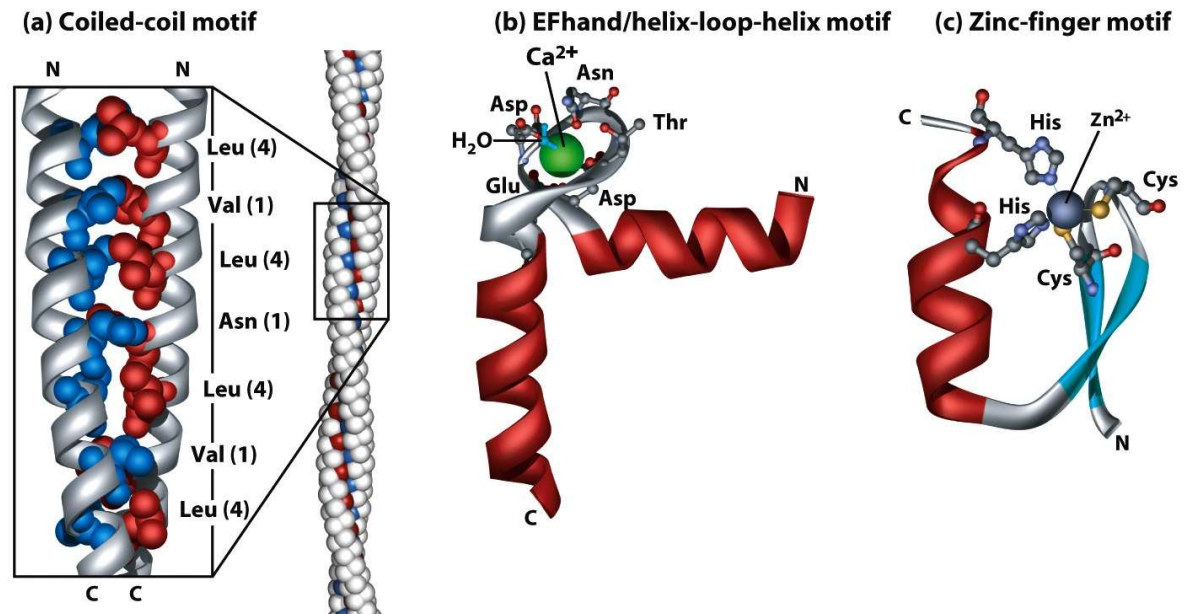
Characteristics of transcription factors (TF)

- Provide a response to various extracellular or intracellular stimuli indicating the need to turn on one or more genes.
- Unlike most proteins, they are able to enter the nucleus.
- Recognize and bind to specific DNA sequences.
- Make contact with the transcription apparatus, either directly or indirectly.

Characteristics of transcription factors (TF)

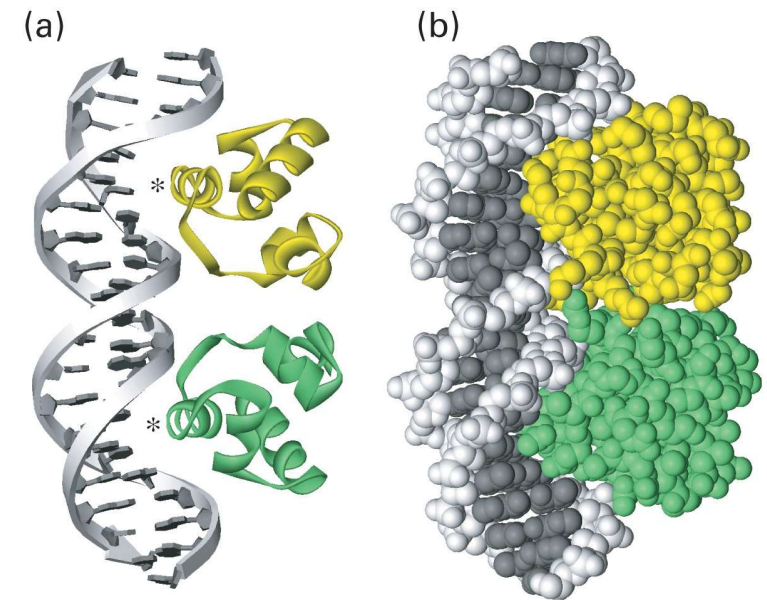
- DNA-binding domains of several transcription factors employ similar tertiary structure – motif.
- DNA-binding motifs are evolutionarily conserved AA sequence which have a defined conformation since AA sequence ultimately determines structure.

- A given DNA-binding motif can occur in a number of proteins where it carries out the same or similar functions.
- For examples of the **coiled-coil**, **EF hand/helix-loop-helix**, and **zinc-finger motifs**.



DNA-binding motifs in TFs

- DNA-binding motifs bind specifically to DNA via non-covalent interactions.
- **Helix – Turn – Helix**. The second helix in this motif (the DNA recognition helix) typically binds to a specific sequence of bases in DNA.
- Alpha-helices are one of the most common types of DNA-binding sequences. The side-chains of residues within the α -helix often bind to the surfaces of bases exposed in the major groove of double-helical DNA. Binding to phosphates and bases in the minor groove typically is less important.
- Helix-turn-helix TFs are common in bacteria.



DNA-binding motifs in TFs

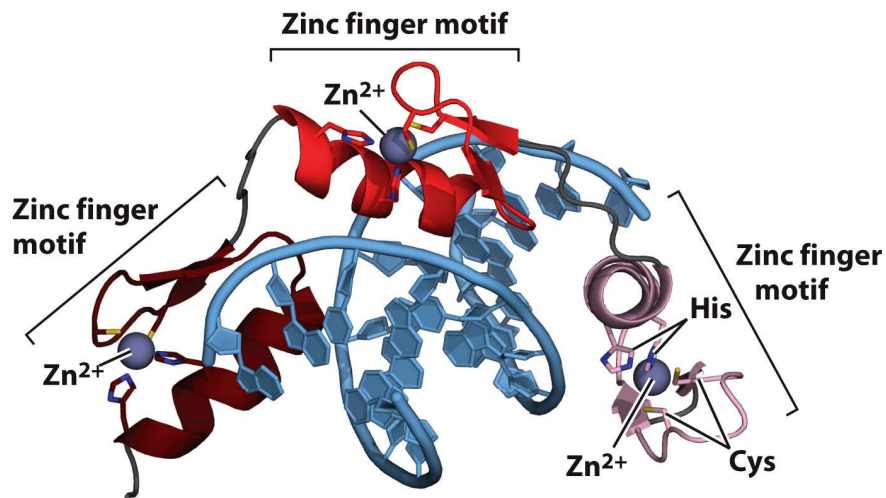


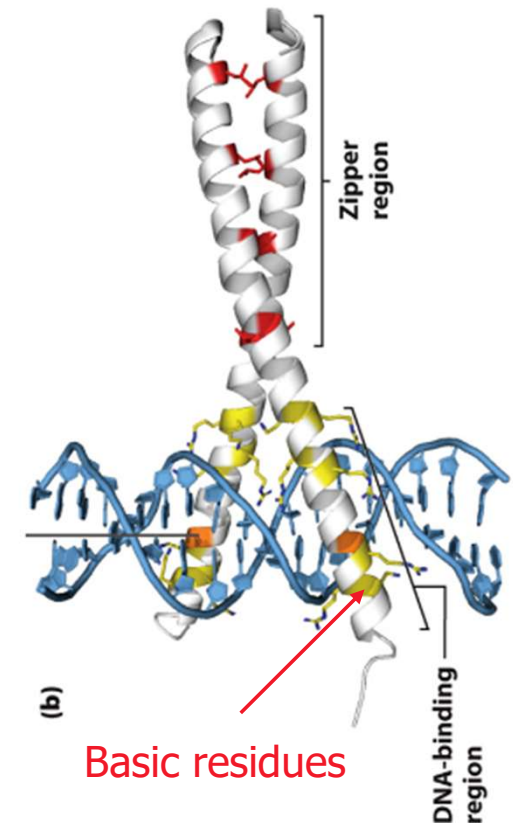
Figure 28-12
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

- Zinc finger - most common DNA-binding motif in human and multicellular animal.
- Two types of zinc finger TFs:
- C2H2 zinc finger TFs - its 2 cysteine and 2 histidine residues bind to zinc ions (Zn²⁺) and the α -helix containing the 2 histidines binds to bases in the major groove.
 - C4 zinc finger TFs is much less common.

- Most TFs containing this motif are dimeric. Nuclear receptors, which bind steroid hormones and other compounds, contain this motif. Zinc ions are bound to the DNA recognition helix of this motif, which contacts bases in the major groove.

DNA-binding motifs in TFs

- Leucine-zipper TFs
- Contain extended α -helices wherein every 7th amino acid is leucine. This periodicity creates a nonpolar face on one side of the helix that is ideal for dimerization with another such protein via a **coiled-coil motif**.
- So-called basic zipper (bZip) TFs have a similar structure except that some leucines are replaced by other nonpolar amino acids.
- The N-terminal ends of both leucine-zipper and bZip proteins contain **basic amino acids that interact with bases in the major groove**.



DNA-binding motifs in TFs

- Helix – Loop – Helix TF
- Another class of TF, the basic helix-loop-helix (bHLH) proteins are similar to bZip proteins, but contain a loop between the DNA recognition helix and the coiled-coil region. bZip and bHLH proteins commonly form heterodimeric TFs.

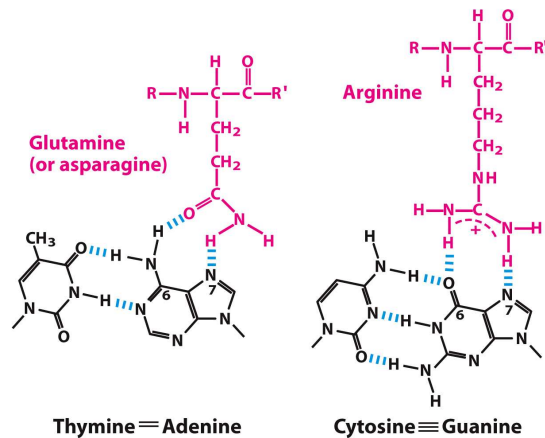
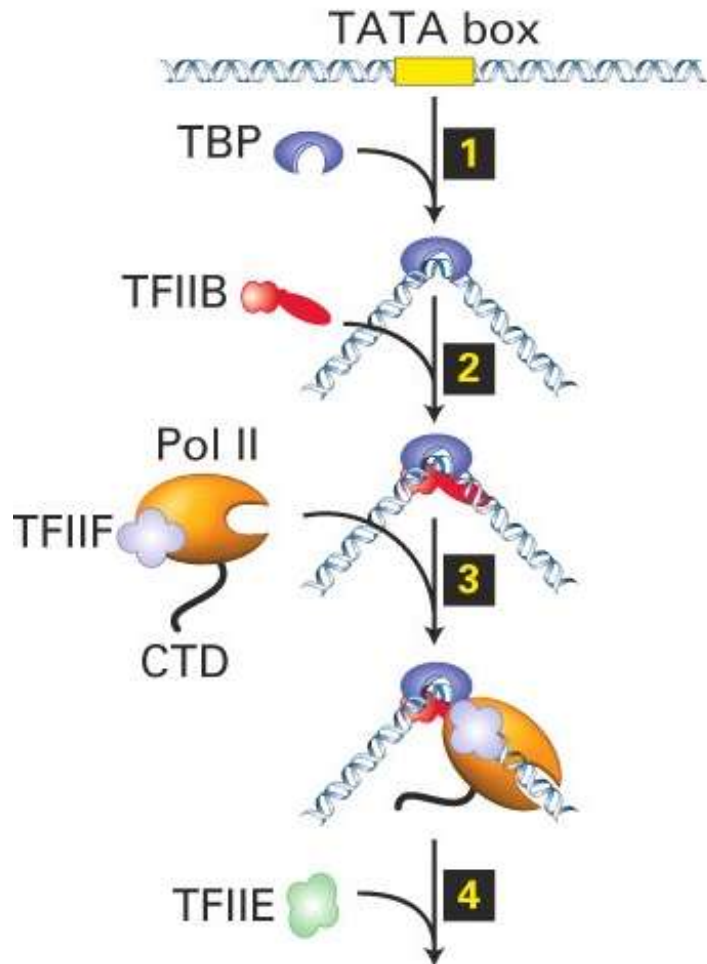


Figure 28-10
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company



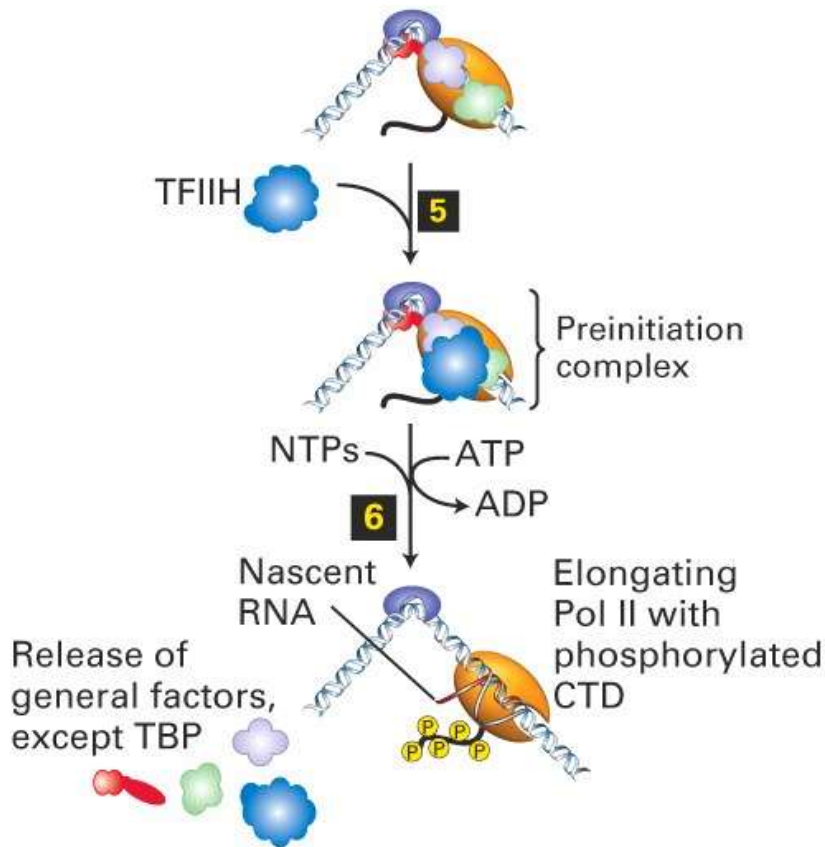
Basic residues

RNA Pol II preinitiation complex formation



- The sequential steps leading to the assembly of the RNA Pol II pre-initiation complex.
- 1. **TBP** binds to the **TATA box** and bends (DNA looping) DNA near the promoter.
- 2. **TFIIB** binds, and then a complex between Pol II and TFIIF loads onto the promoter.
- 3. **TFIIF** positions the **Pol II active site** at the mRNA start site and helps maintain chromatin at the promoter in an uncondensed state.
- 4. **TFIIE** then binds creating a TFIIF docking site.

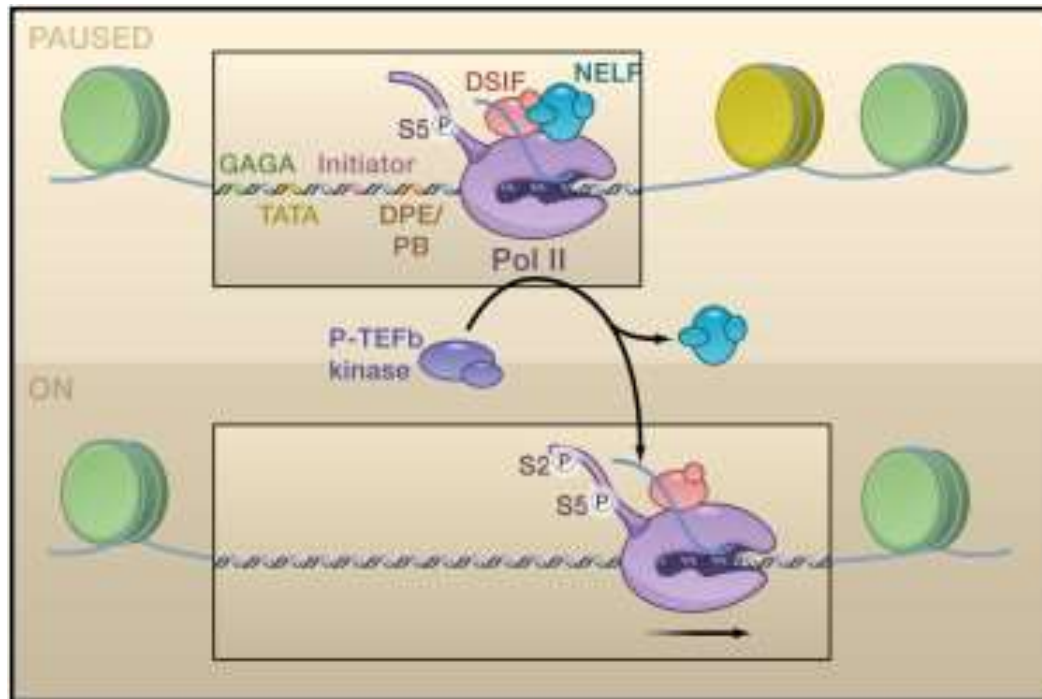
RNA Pol II preinitiation complex formation



- 5. With the addition of TFIID, the assembly of the **pre-initiation complex** is complete.
- 6. Subsequently, one subunit of **TFIID melts DNA at the promoter**, obtaining energy by ATP hydrolysis. RNA Pol II then begins transcribing the mRNA.
- Another subunit of TFIID **phosphorylates the RNA Pol II CTD**, making RNA Pol II highly processive.
- **Tissue-specific TFs** bound to enhancers and promoter-proximal elements also play important roles in transcription initiation *in vivo*.

Transcription – elongation phase

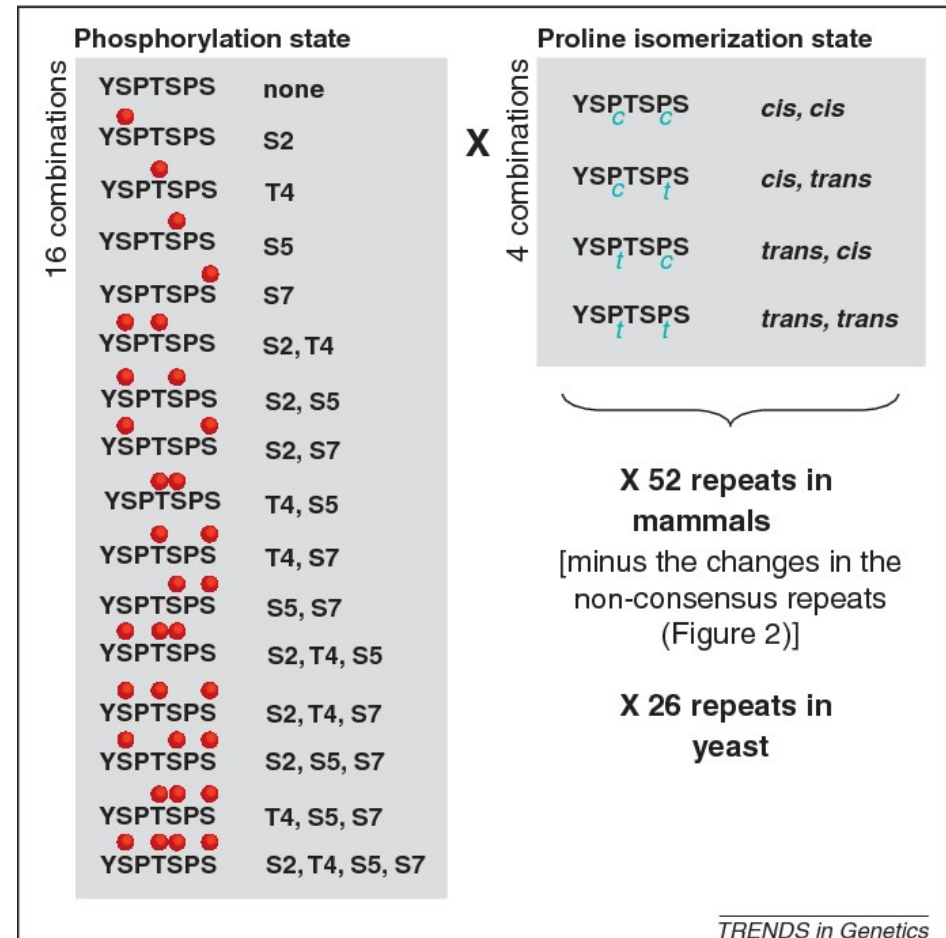
- Right after transcription initiation the RNA polymerase II is paused typically 30–50 nucleotides after transcription initiation site.
- Negative transcription elongation factors DSIF, NELF.



- RNA Pol II undergone promoter escape and contains phosphorylation of serine 5 (Ser5) in the C-terminal domain (CTD).
- Recruitment of P-TEFb (positive transcription elongation factor b) causes phosphorylation of Ser2 in the CTD, resulting in RNA Pol II transcription elongation.

Transcription – elongation phase

- Posttranslational modification, phosphorylation in particular, of CTD of RNA Pol II corresponds with specific phase of transcription.
- Other modification of CTD of RNA Pol II.
 - Acetylation
 - Proline isomerization.
- We talk about **CTD code**.

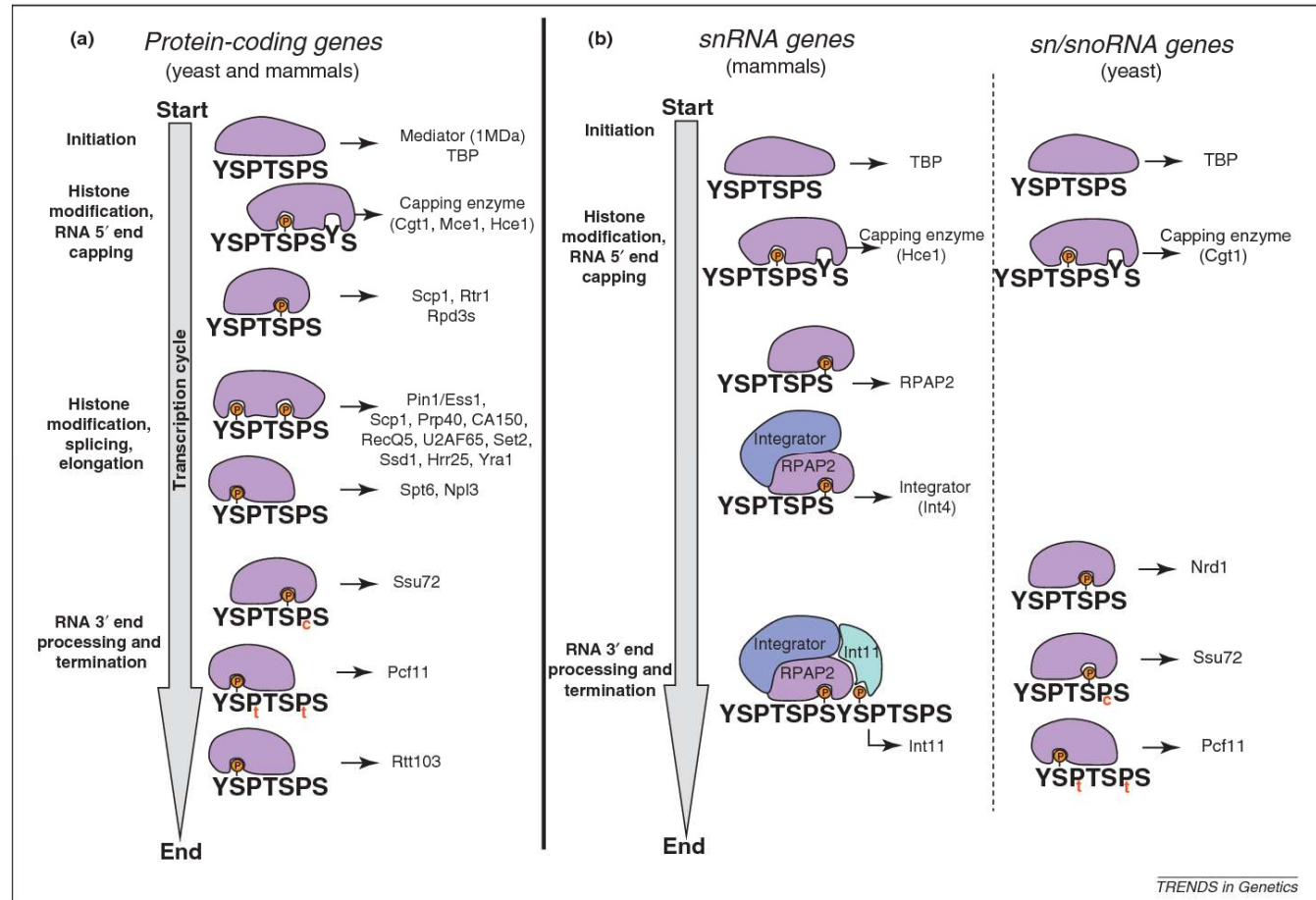


TRENDS in Genetics

MUNI
SCI

Transcription – elongation phase

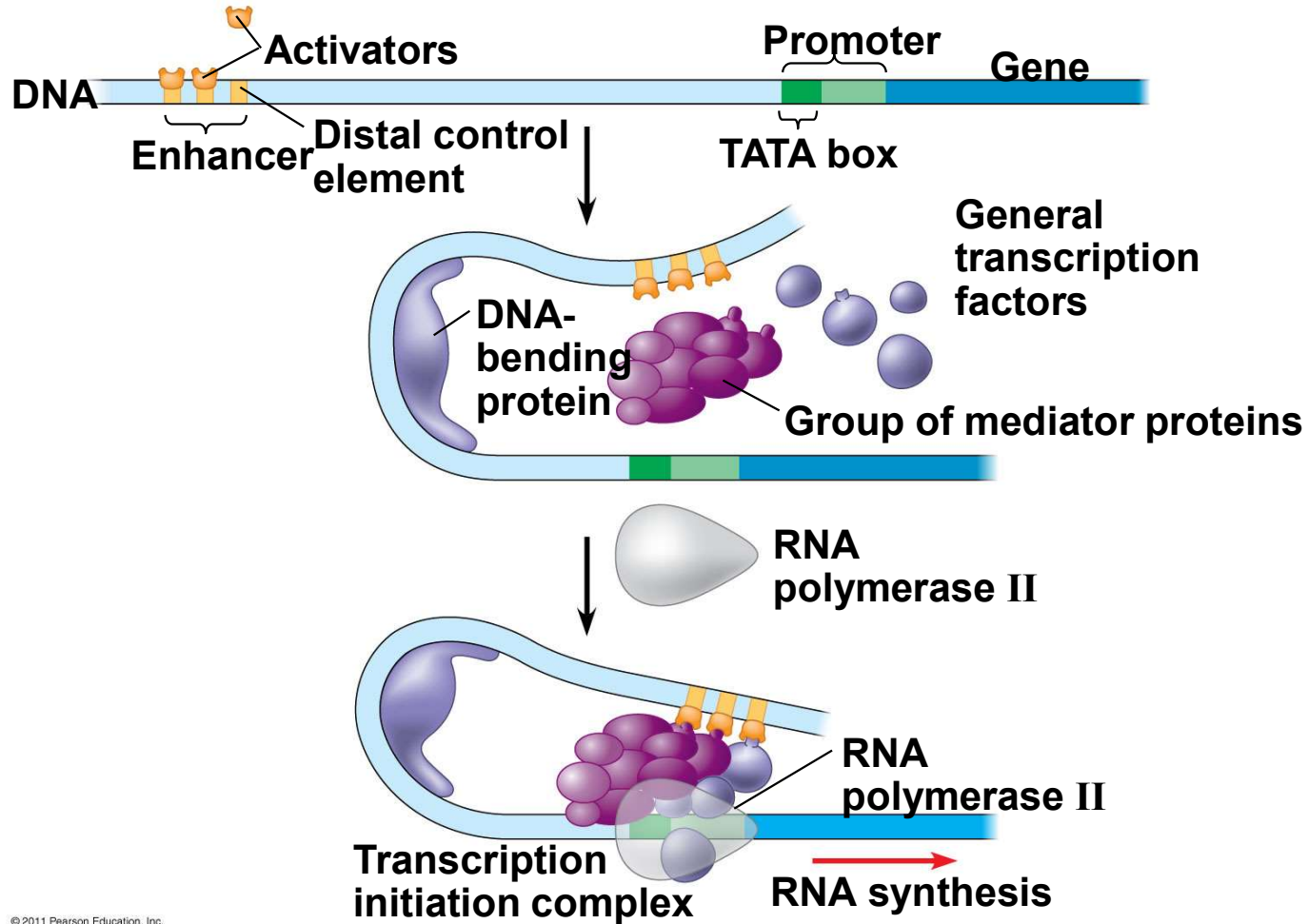
- Modification of CTD of RNA Pol II coincides with various steps in transcription.
 - DNA and chromatin remodeling.
 - Histone modifications.
 - DNA processing.
 - Transcription termination and polyadenylation.



Enhancers and Activators

- Distal control DNA elements, which are called **enhancers**, may be far away from a gene or even located in an intron.
- Enhancers can be **thousands of nucleotides** away from the TATA box of the promoter.
- Can be bound by activators.
- Activators have two domains **DNA-binding**, **protein-binding**, and/or signal molecule-binding domains coupled with **transcription activating domain**.
- Facilitate a sequence of protein-protein interactions that result in transcription of a given gene.
- Some regulate a few genes; some regulate many hundreds of genes.

Enhancers and Activators

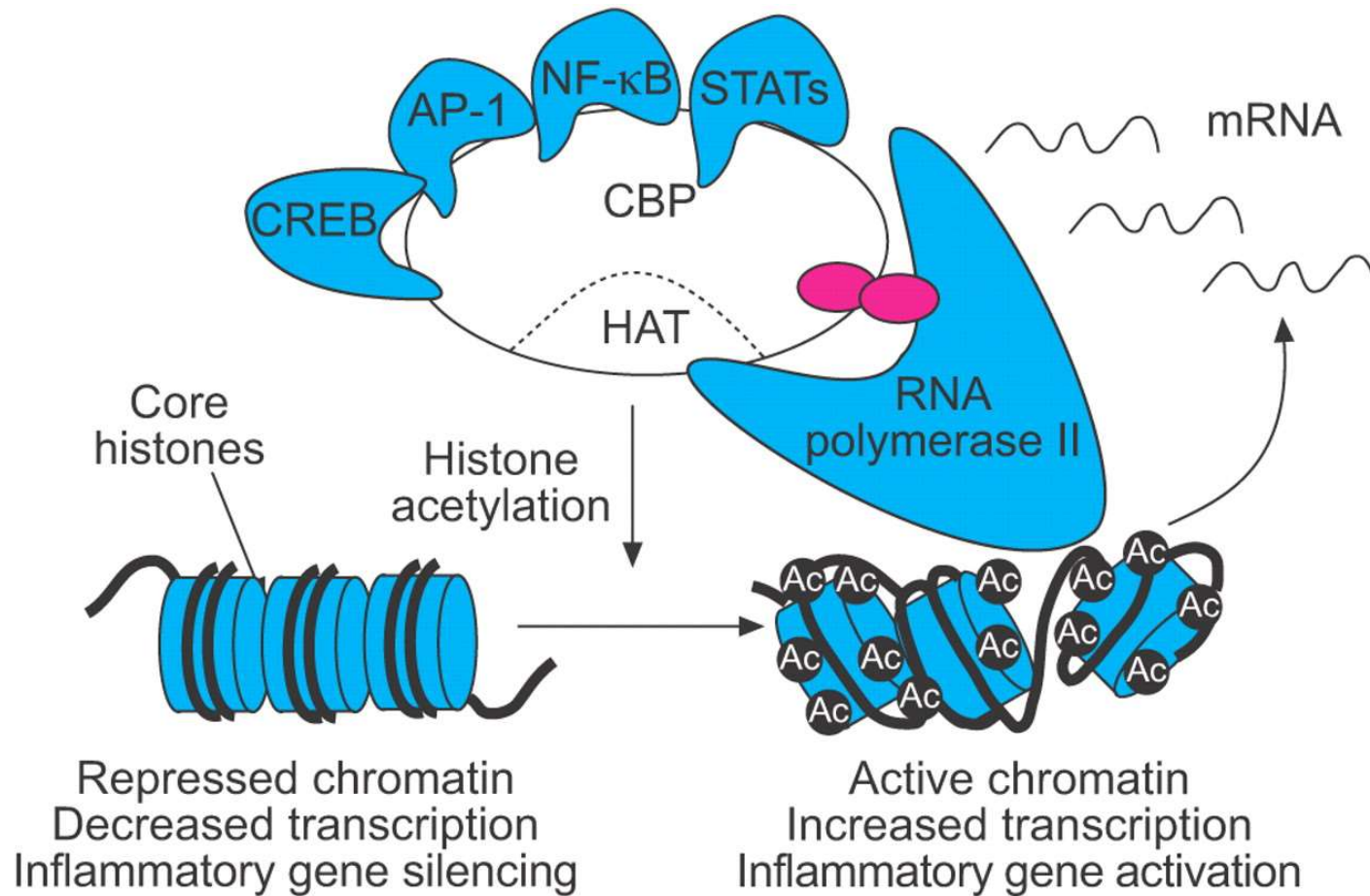


© 2011 Pearson Education, Inc.

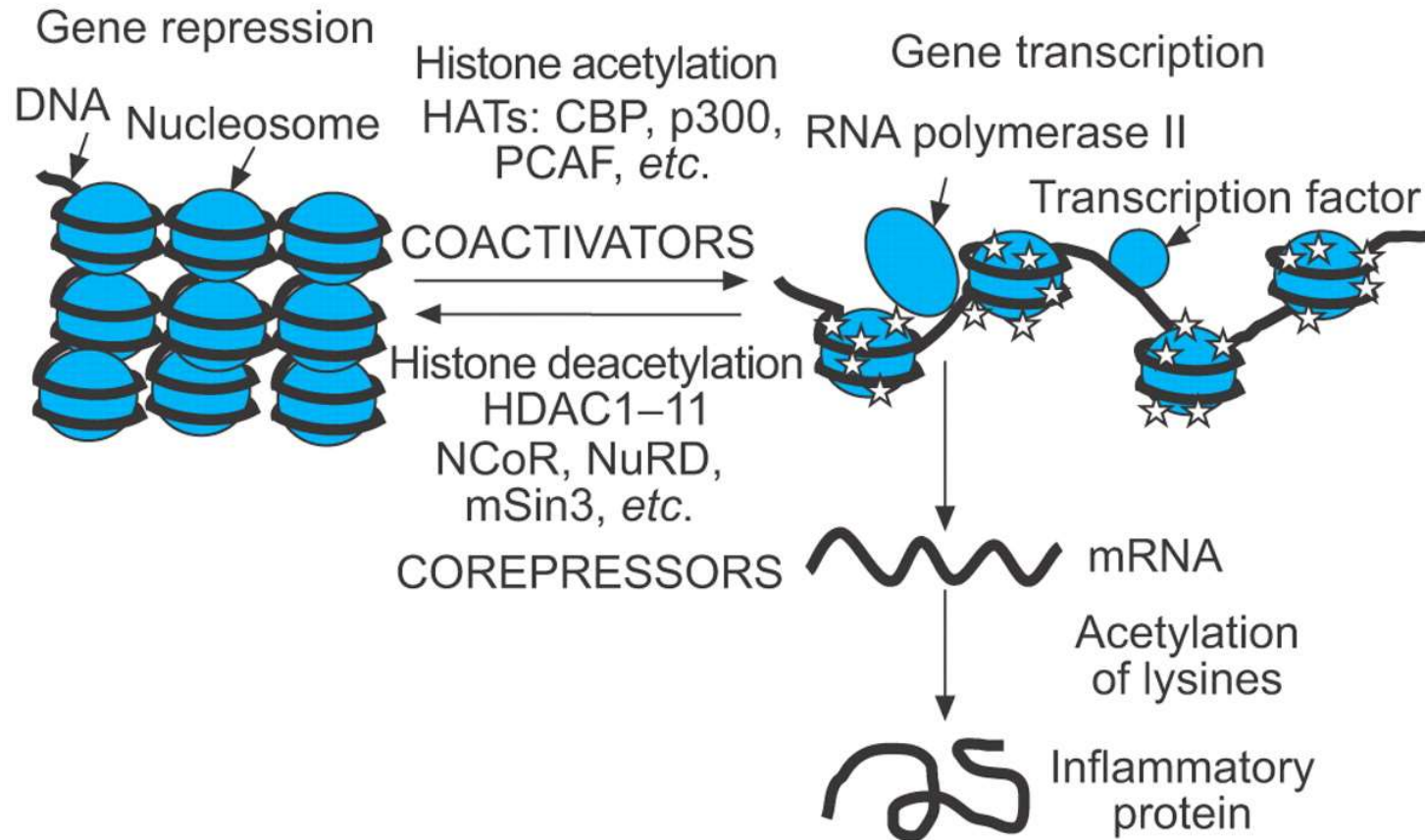
Chromatin remodeling complexes

- CRE - cAMP response element.
- CREB - (cAMP response element-binding protein/CRE-binding protein).
 - Specific transcription factor.
 - Transcription of genes regulated by CREB: somatostatin, c-fos, tyrosine hydroxylase, neuropeptides, enkephalin, genes involved in circadian rhythm control, and more.
- CBP–CREB binding protein.
 - Histone acetyltransferase.
 - Coactivator CREB.
 - Many other transcription factors (c-myc, c-fos, p53, E2F, NF-κB,...).

Chromatin remodeling complexes

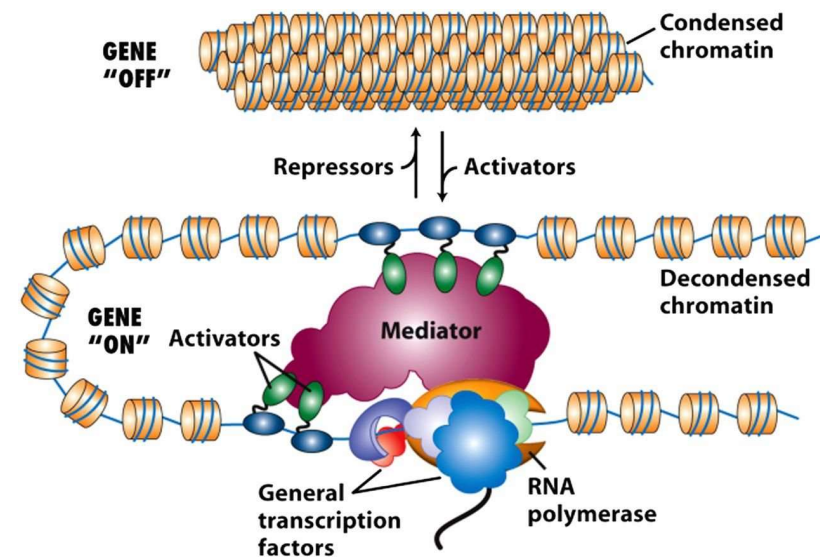
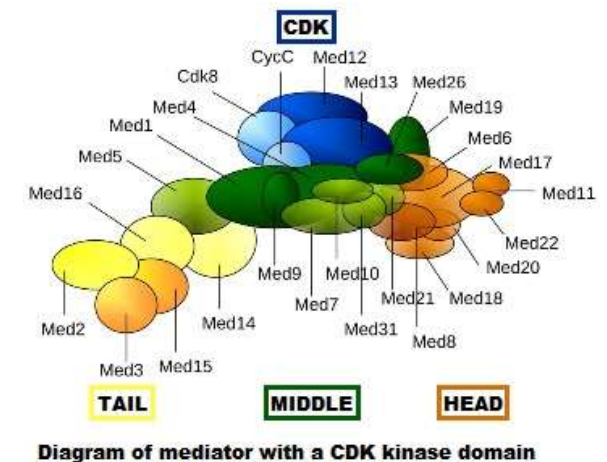


Chromatin remodeling complexes



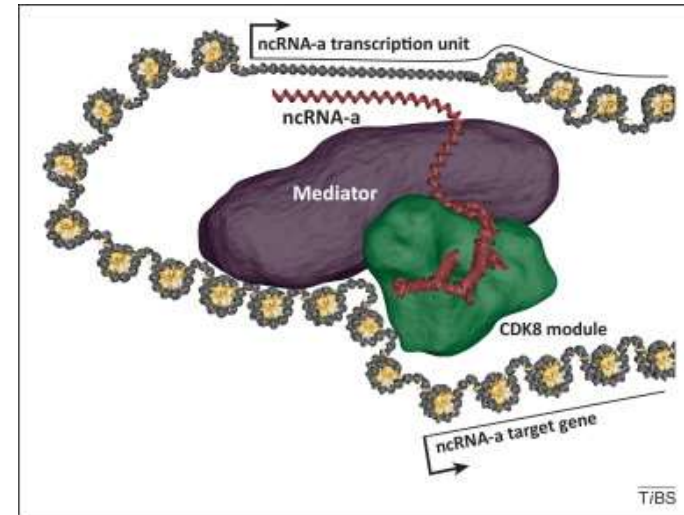
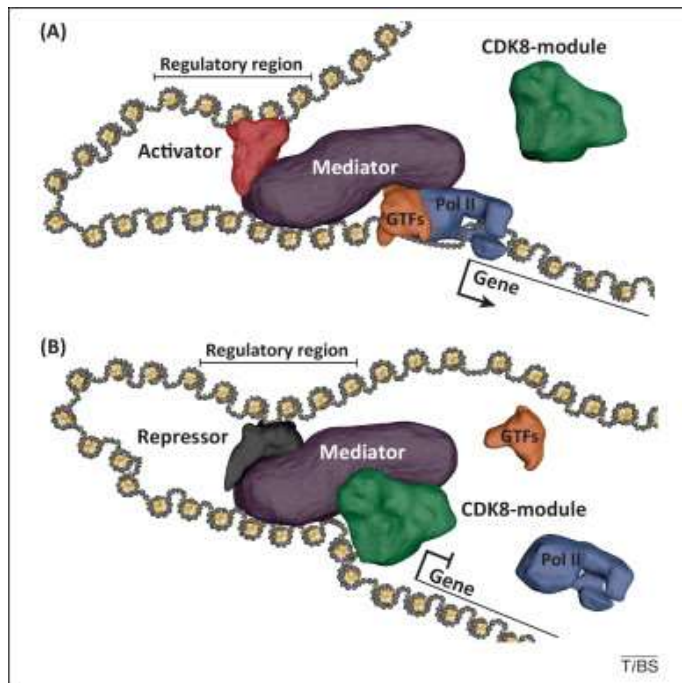
Mediator complex

- Mediator is an ~30-subunit multiprotein complex.
- Mediator functions as a **molecular bridge** between RNA Pol II and transcription factors bound to **enhancers and promoter proximal transcription control sequences**.
- Mediator influences all stages of transcription – from initiation to elongation.
- Non-coding RNA interacts directly with Mediator to influence transcription.



Mediator complex

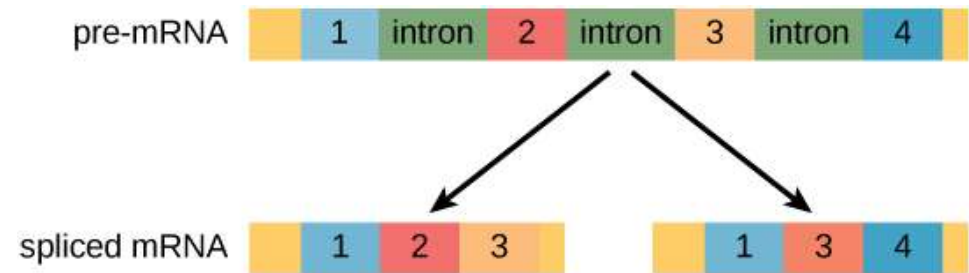
- It can **activate** or **repress** transcription.



- The mammalian Mediator complex interacts directly with a subset of long noncoding RNAs termed **lncRNA-activating (ncRNA-a)**, which activate neighboring genes using a cis-mediated mechanism.

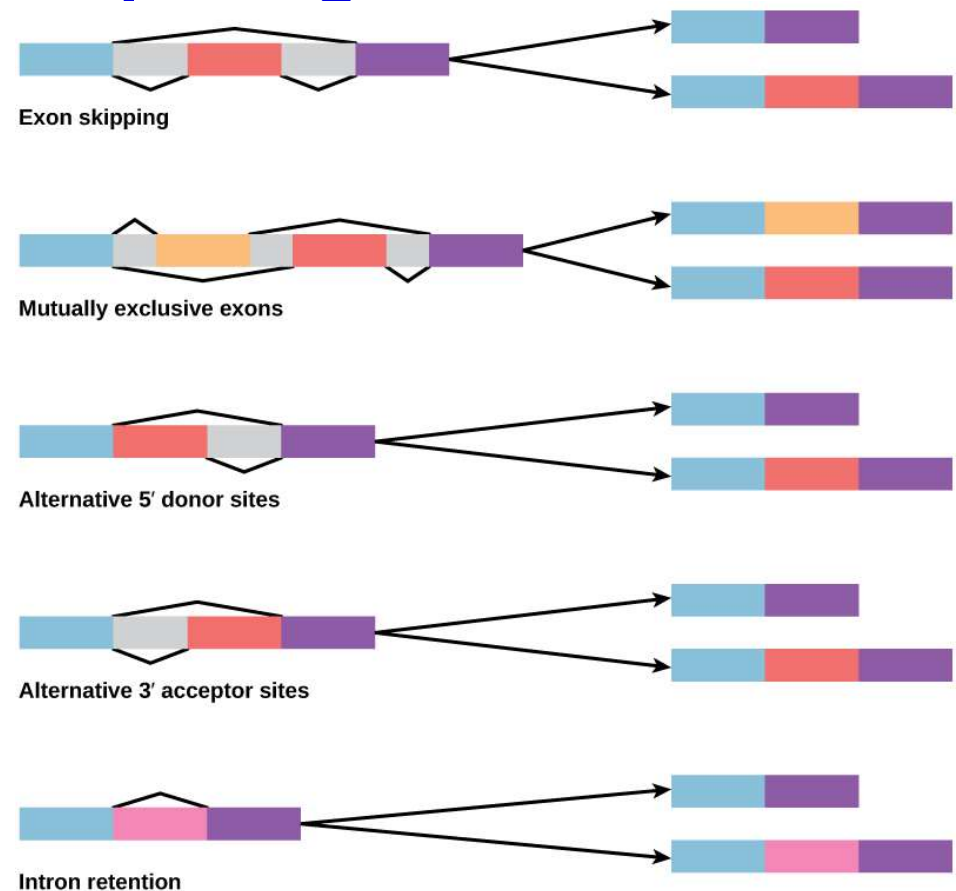
RNA processing

- Splicing, capping, and addition of a poly-A tail to an RNA molecule can be **regulated**.
- **Different mRNAs** may be made from the same pre-mRNA by **alternative splicing**.



Alternative splicing

- **Alternative RNA splicing** is a mechanism that allows **different protein forms** to be produced from one gene when different combinations of introns, and sometimes exons, are removed from the transcript.
- It is a common mechanism of gene regulation in eukaryotes; according to one estimate, 70 percent of genes in humans are expressed as multiple proteins through alternative splicing.
- The cause of many genetic diseases is improper alternative splicing rather than mutations in a sequence.



Splicing Produces Related but Distinct Protein Isoforms

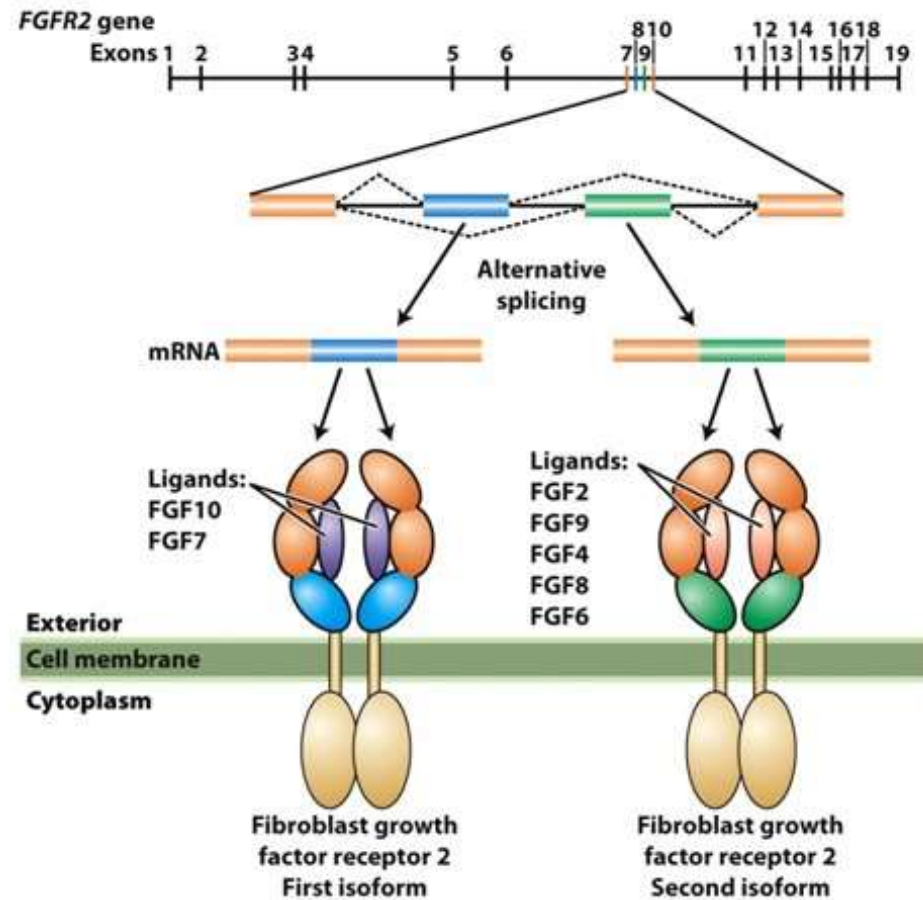
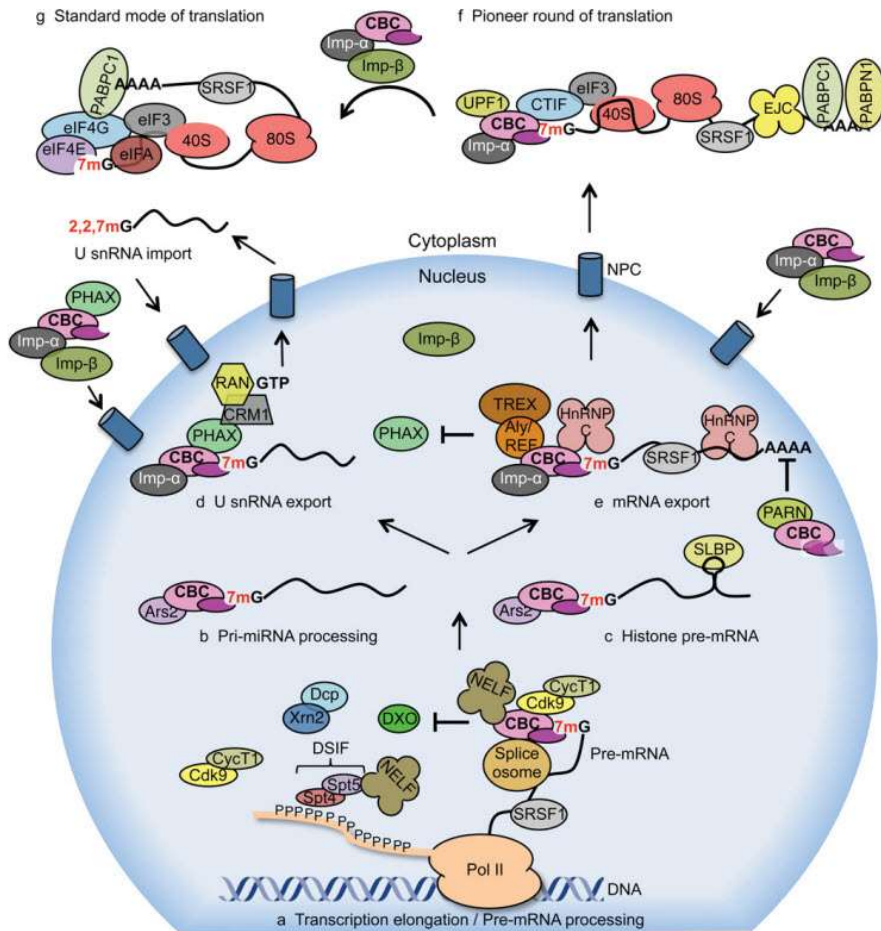


Figure 9-19
Introduction to Genetic Analysis, Tenth Edition
© 2012 W. H. Freeman and Company

RNA stability

- RNA stability. The lifetime of an mRNA molecule in the cytosol affects how many proteins can be made from it.
- Capping.
- Degradation of mRNA.
- Small regulatory RNAs called miRNAs can bind to target mRNAs and cause them to be chopped up.
- Small nuclear RNA – 7SK RNA, U1snRNA.
- Long noncoding RNA.

Regulation of 5' - CAP

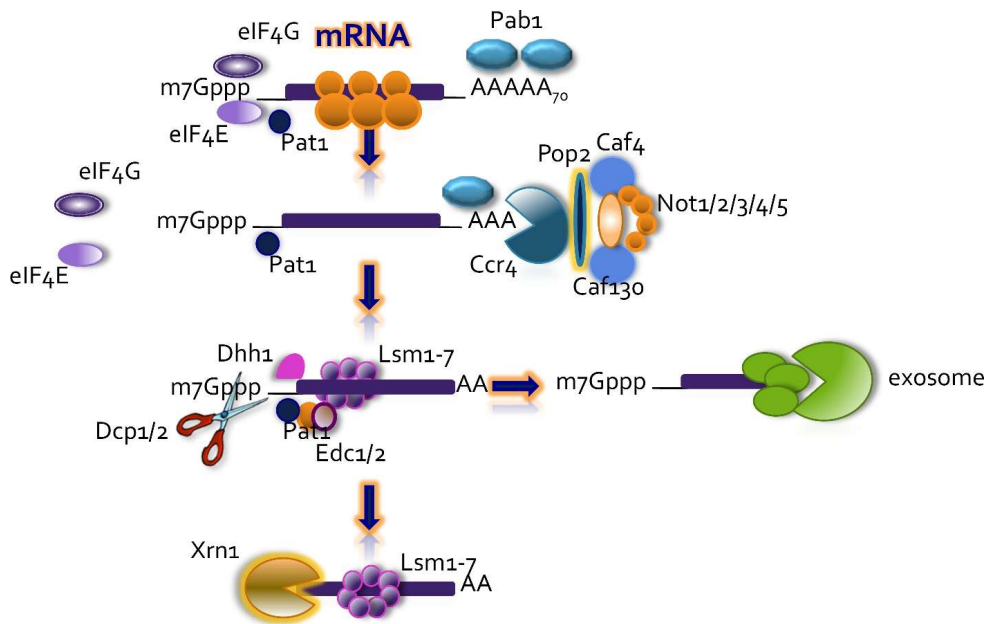


The 5' cap has four main functions:

- Regulation of **nuclear export** of RNA is regulated by the cap binding complex (CBC), which binds to 7-methylguanylate-capped RNA. The CBC is then recognized by the nuclear pore complex and the mRNA exported.
- Prevention of **degradation** by exonucleases.
- Promotion of **translation**.
- Promotion of **5' proximal intron excision**.

Degradation of RNA

- The first step of untranslated mRNA degradation is deadenylation, removal of polyA tail.
- Deadenylated mRNAs are degraded in the 3'-to-5' direction by the exosome complex or, mainly, are transferred to P-bodies for decapping.



- In P-bodies the m7Gppp cap is removed by **Dcp1/2** proteins.
- RNA without cap is degraded in the 5'-to-3' direction by the exoribonuclease **Xrn1**.

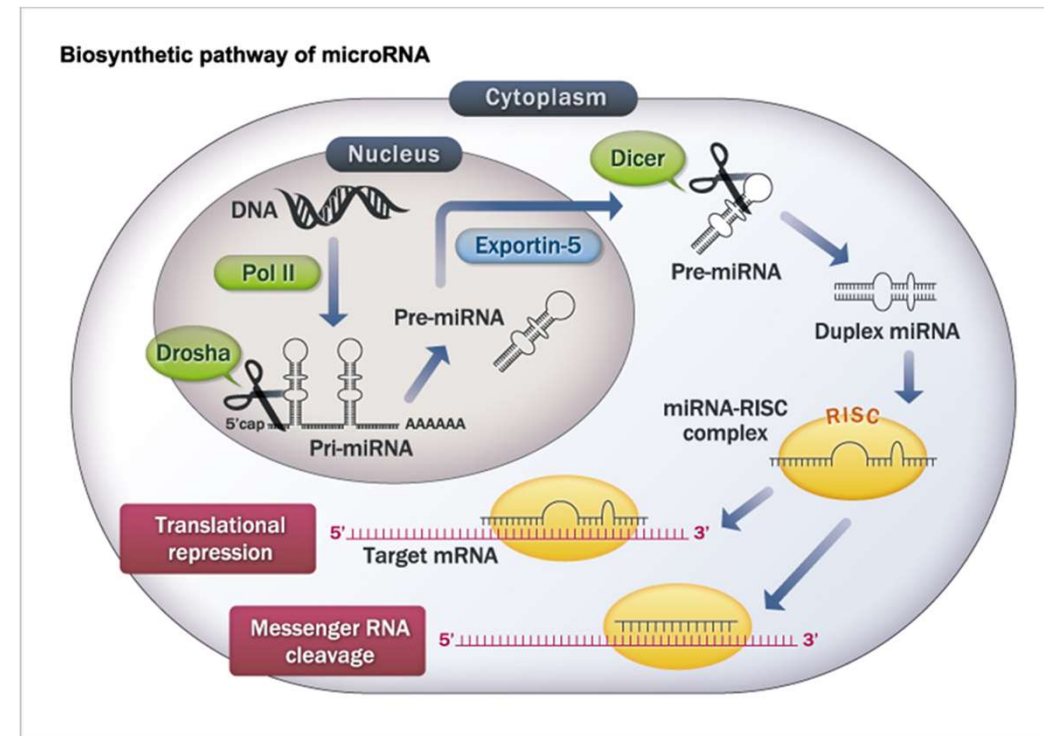
MicroRNA

miRNA

- 18-25 nt, single-strand molecules.
- Post-transcriptionally **regulate gene expression**.
- 2694 human miRNAs (miRBASE in 22, March 2018).
- Evolutionarily conserved and transcribed by RNA Pol II.
- 1-2 % of the genome.
- Genes for miRNAs on all human chromosomes except Y.
- Regulation of the expression of up to 50% of protein-coding genes.
- One miRNA can regulate **tens to hundreds of target mRNAs**.

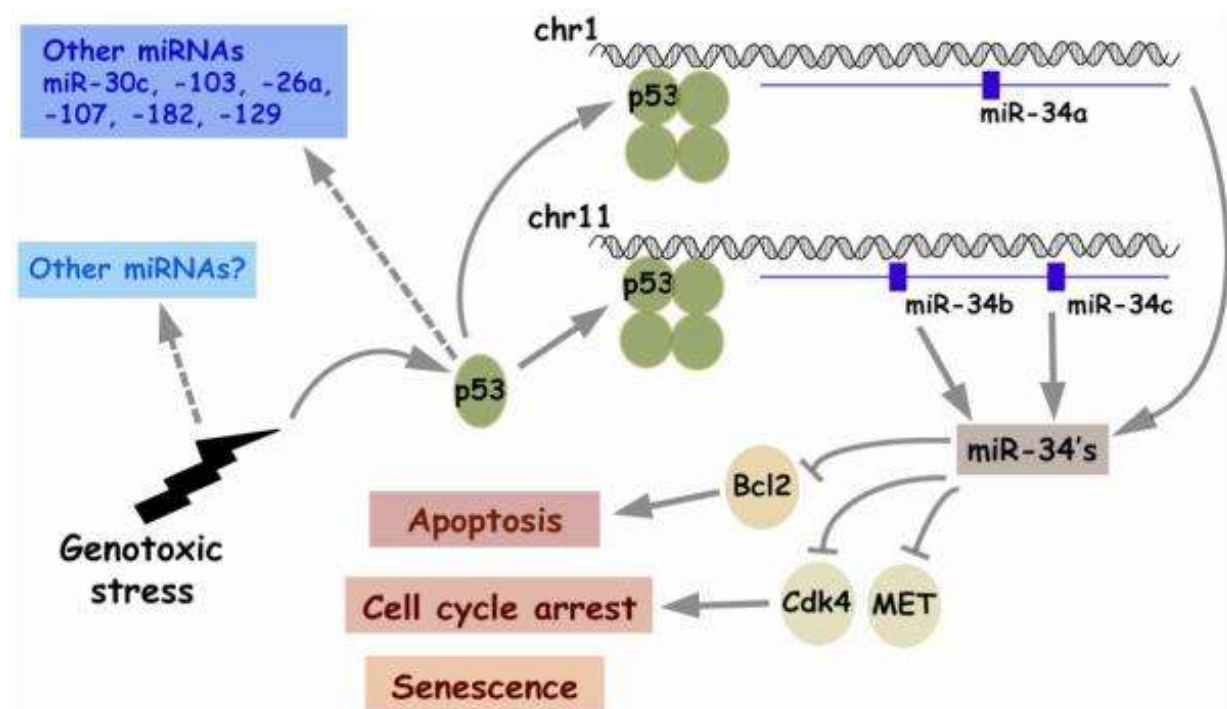
Biogenesis of miRNA

- A miRNA is first transcribed as a long RNA molecule, which forms base pairs with itself and folds over to make a hairpin.
- Next, the hairpin is chopped up by enzymes, releasing a small double-stranded fragment of about 22 nucleotides.
- One of the strands in this fragment is the mature miRNA, which binds to a specific protein to make an RNA-protein complex.
- Targeting of mRNA
 - mRNA cleavage.
 - Translation repression.

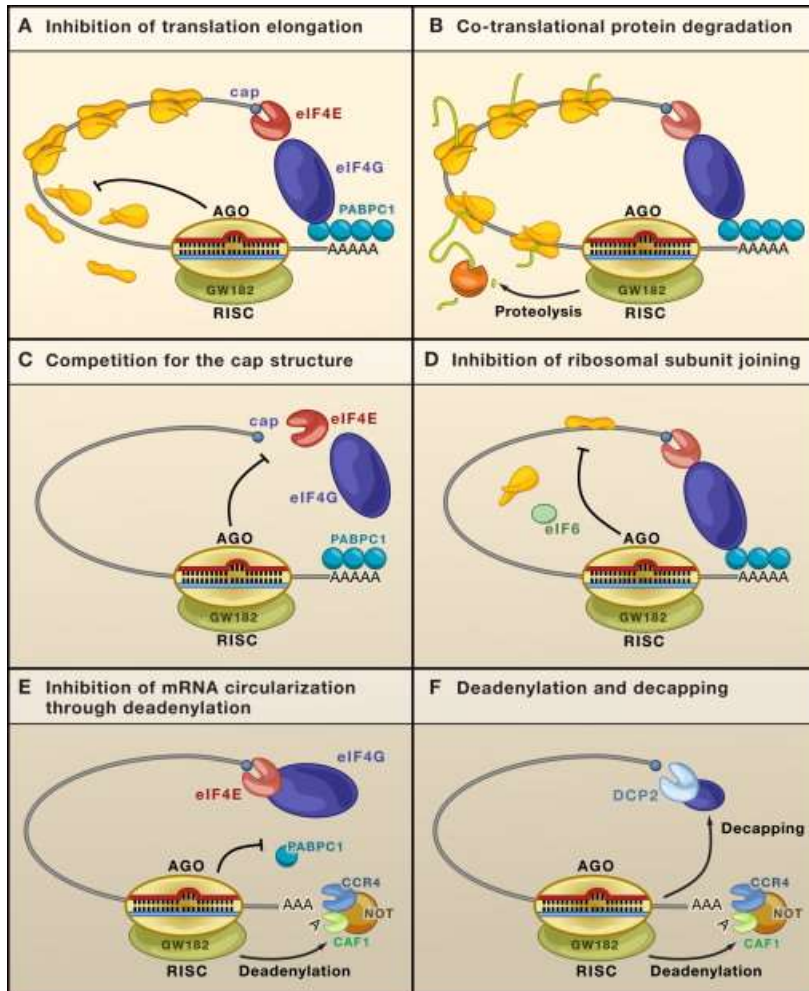


Function of miRNA

- Physiological function
 - Proliferation.
 - Differentiation.
 - Apoptosis and etc.
- Tumorigenesis.



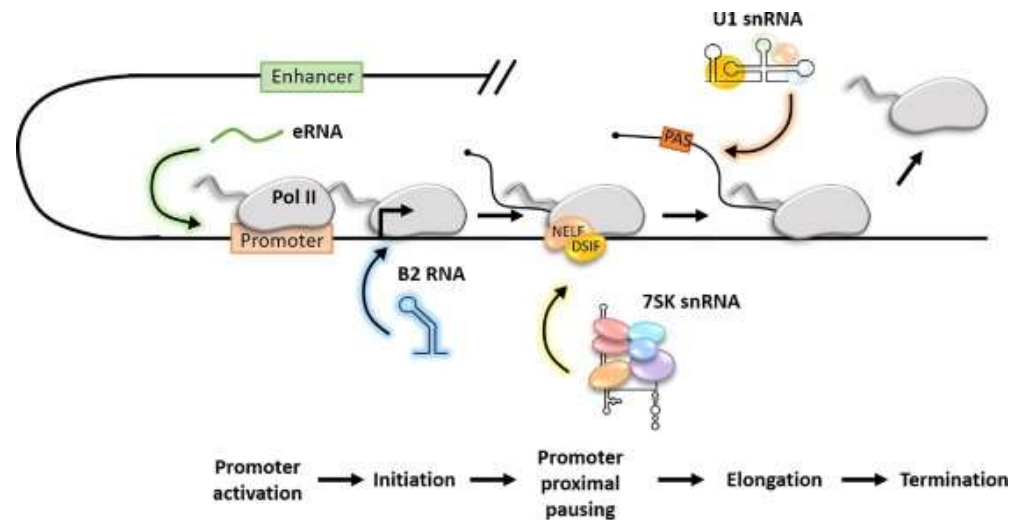
Mechanism of miRNA-mediated gene silencing



- Postinitiation mechanisms (A).
- Cotranslational protein degradation (B).
- Initiation mechanisms. MicroRNAs interfere with a very early step of translation, prior to elongation.
 - Argonaute proteins compete with eIF4E for binding to the cap structure (C).
 - Argonaute proteins recruit eIF6, which prevents the large ribosomal subunit from joining the small subunit (D).
 - Argonaute proteins prevent the formation of the closed loop mRNA configuration by an ill-defined mechanism that includes deadenylation (E).
- MicroRNA-mediated mRNA decay. MicroRNAs trigger deadenylation and subsequent decapping of the mRNA (F).

Small nuclear RNA

- Small nuclear RNA (snRNA) is one of the small RNA with an average size of 150 nt.
- Eukaryotic genomes code for a variety of non-coding RNAs.
- snRNA is a class of highly abundant RNA, localized in the nucleus with important functions in intron splicing and other RNA processing.

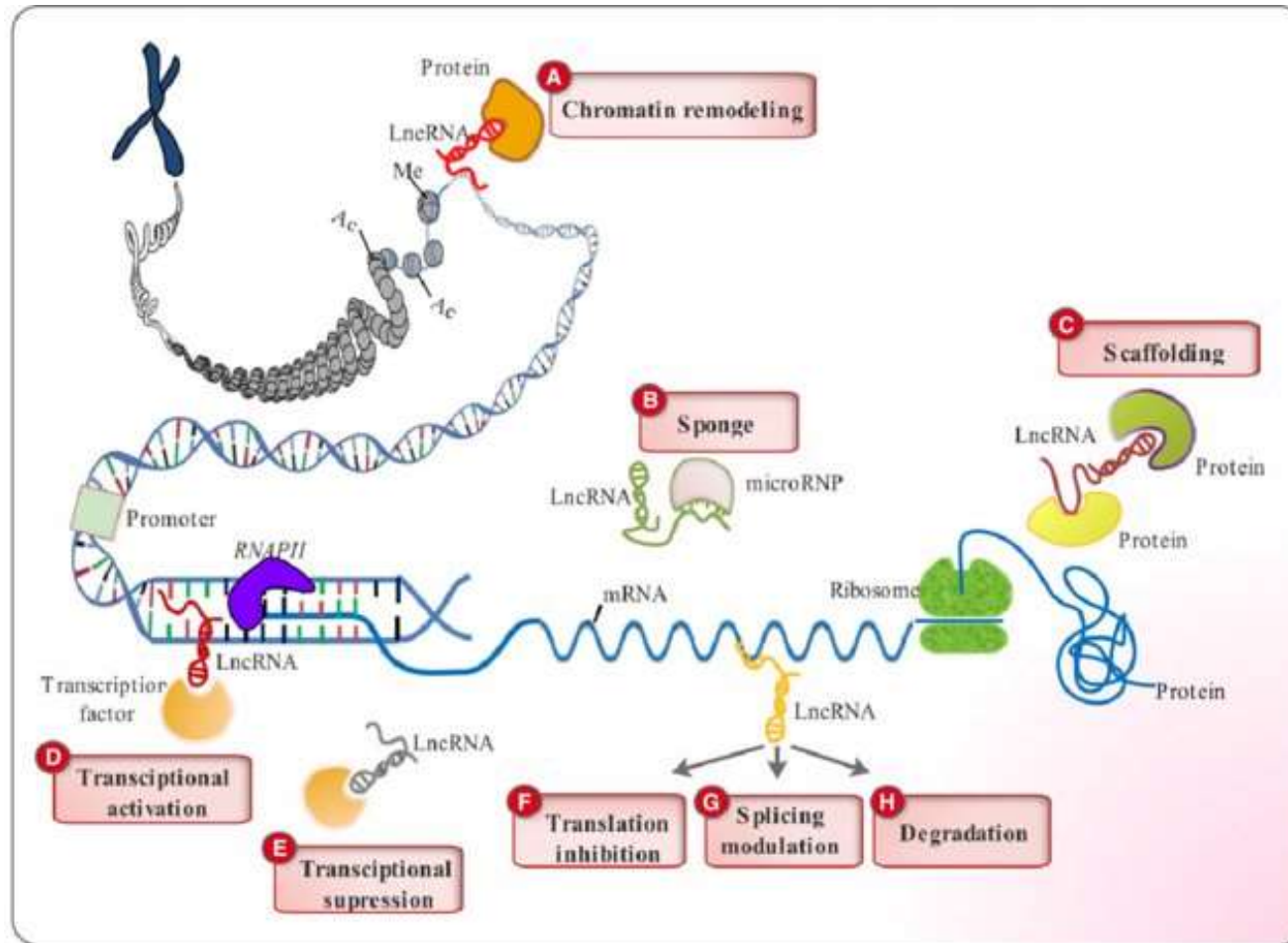


Trends in Genetics

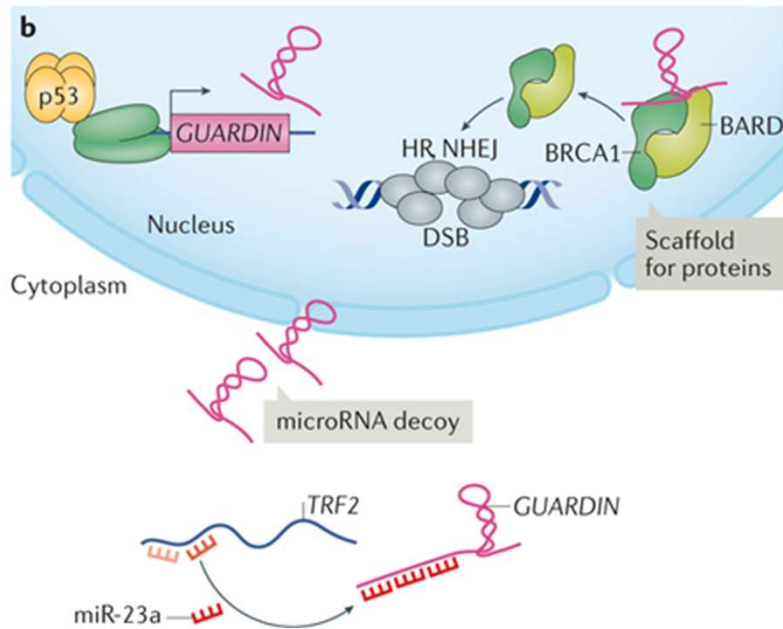
Long noncoding RNA – lncRNA

- Longer than 200 nt; encoded by genes on different chromosomes (in non-coding regions, even in gene introns).
- Usually transcribed by **RNApolIII (III)**, often have a cap at the 5'-end and a poly(A) at the 3'-end, subject to splicing; do not have ORF, cannot be translated.
- Often tissue-specific and developmental stages specific.
- Some lncRNA found in specific DNA locuses.
- Changes in the expression of lncRNAs associated with various diseases (cancer, Alzheimer's, atherosclerosis), can also serve as markers.
- Present in body fluids, possibility of non-invasive diagnostics.
- In humans about 50,000 -100,000 genes for lncRNA, up to 270,000 different transcripts lncRNAs.

Effect of lncRNA on gene expression

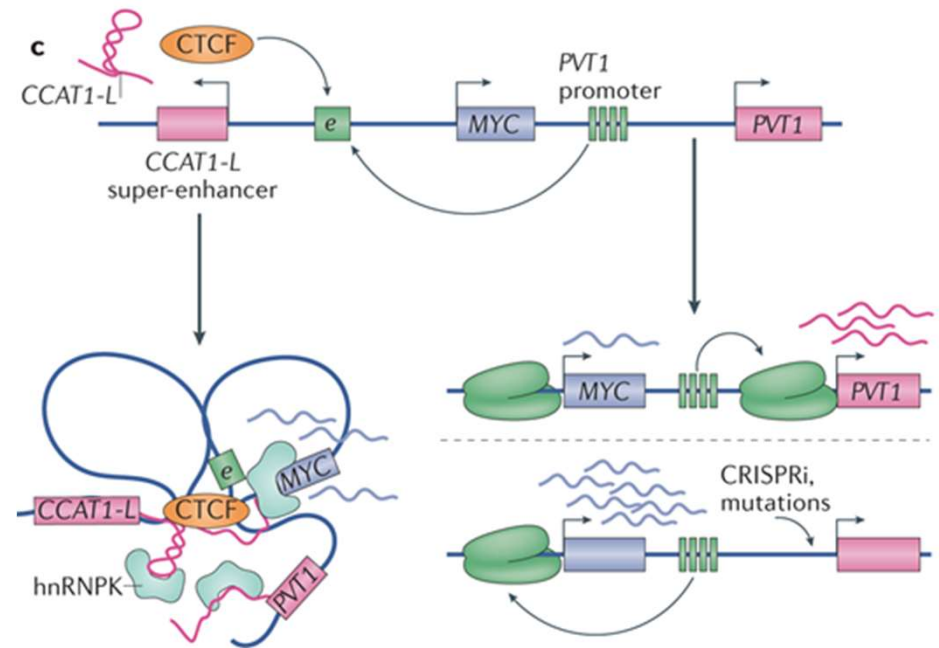


Role of lncRNA in tumorigenesis



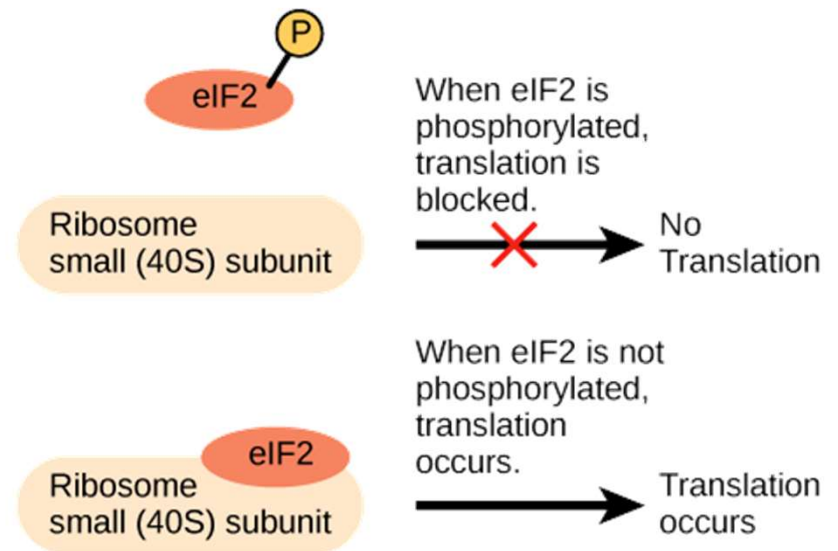
- DNA damage and apoptosis regulation.
- Genome integrity – telomeric repeats and double-strand break.

- Myc transcription.

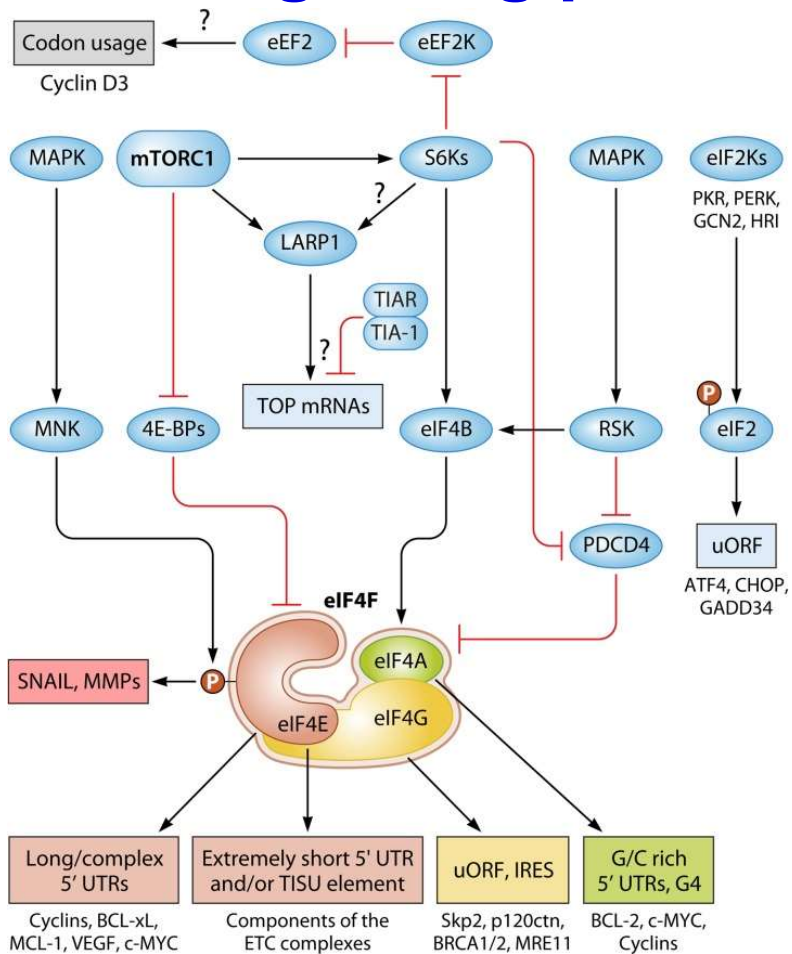


Regulation of translation

- Translation of mRNA involves many “helper” proteins, which make sure the ribosome is correctly positioned.
- Elongation initiation factor-2 (eIF-2) binds to a part of the ribosome called the small subunit.
- When **eIF-2 is phosphorylated, it's turned "off,"** - it undergoes a shape change and can no longer play its role in initiation, so translation cannot begin. When **eIF-2 is not phosphorylated, in contrast, it's "on"** and can carry out its role in initiation, allowing translation to proceed.
- In this way, phosphorylation of eIF-2 acts as a switch, turning translation on or off.

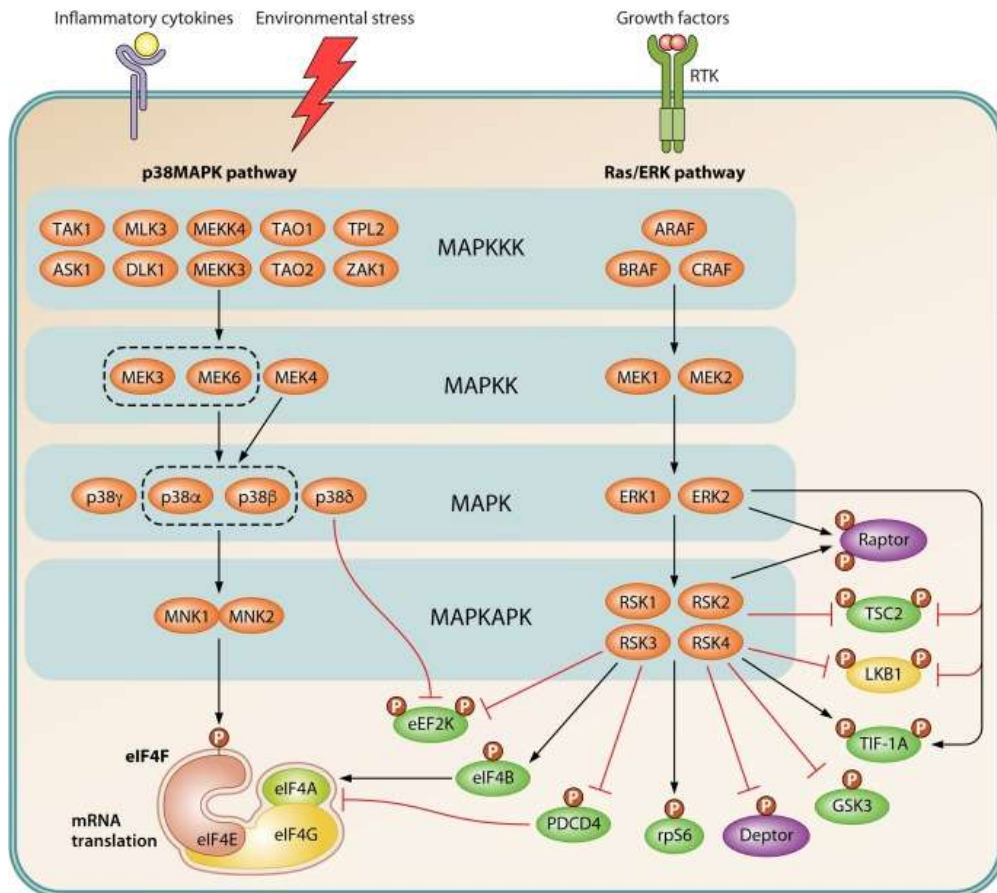


Signaling pathways in regulation of translation



- The mTOR and MAPK pathways affect the translome by modulating the expression of specific subsets of mRNAs.
- Phosphorylation of the 4E-BPs by mTOR leads to their dissociation from eIF4E, which stimulates the interaction of eIF4E with eIF4G and assembly of the eIF4F complex.
- Phosphorylation of eIF4E also seems to bolster the translation of mRNAs encoding proteins involved in tumor dissemination.
- Also eIF4A promotes the translation of mRNAs with G/C-rich 5' UTR sequences, such as the 12-nucleotide guanine quartet (CGG)₄ motif, which can form RNA G-quadruplex structures.

Signaling pathways in regulation of translation



- The Ras/ERK and p38MAPK pathways are activated by a wide range of stimuli, including cytokines, growth factors, and diverse environmental stresses.
- MNK interacts with eIF4G and phosphorylates eIF4E on Ser²⁰⁹, a site that increases its oncogenic potential and facilitates the translation of specific mRNAs.
- RSK phosphorylates rpS6, eIF4B, PDCD4, and eEF2K, which are important regulators of translation.
- ERK and RSK also collaborate in the regulation of ribosome biogenesis by promoting TIF-1A phosphorylation.

THANK YOU FOR YOUR ATTENTION.

