# Molecular Biology Lecture 1

## Gene Expression

***Dr. J Kamau***

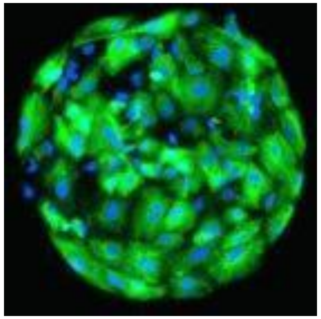
# Disease/Syndromes associated …..

- Cancer, Autoimmunity,
- Neurological disorders, diabetes
- Cardiovascular diseases
- Obesity

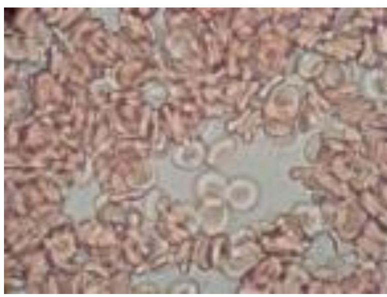
# Questions to Ponder…..

- How do your cells “know” what kind of cell they are?
- How do your cells “know” when to make a particular protein? When to stop making it?
- How does the environment affect your cells?
- ANSWER: Gene Expression

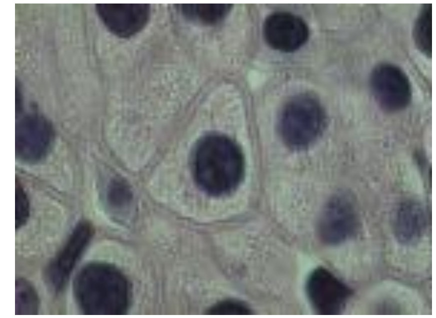
# What makes cells from the same individual look different?



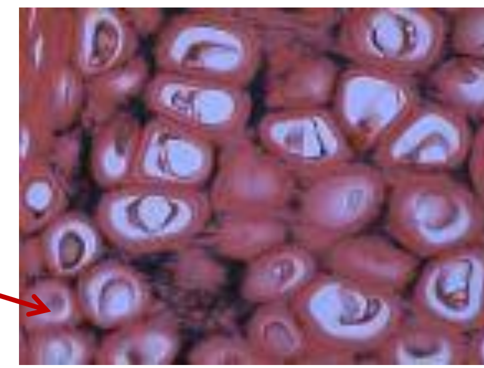
Stem Cells



Red Blood Cells



Liver Cells



Cartilage Cells

DNA sequence in each cell is the same, but different cell types have different “GENE EXPRESSION PATTERNS”

Neuron and lymphocyte
Different morphology, same genome

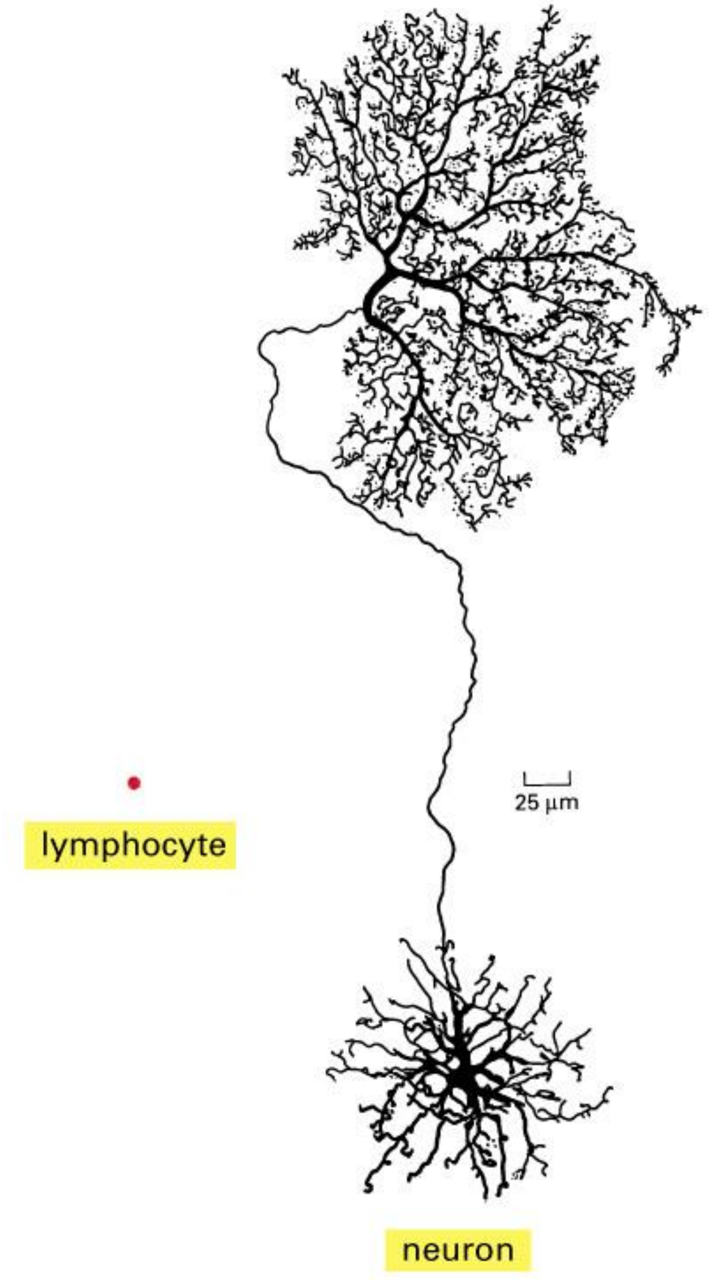
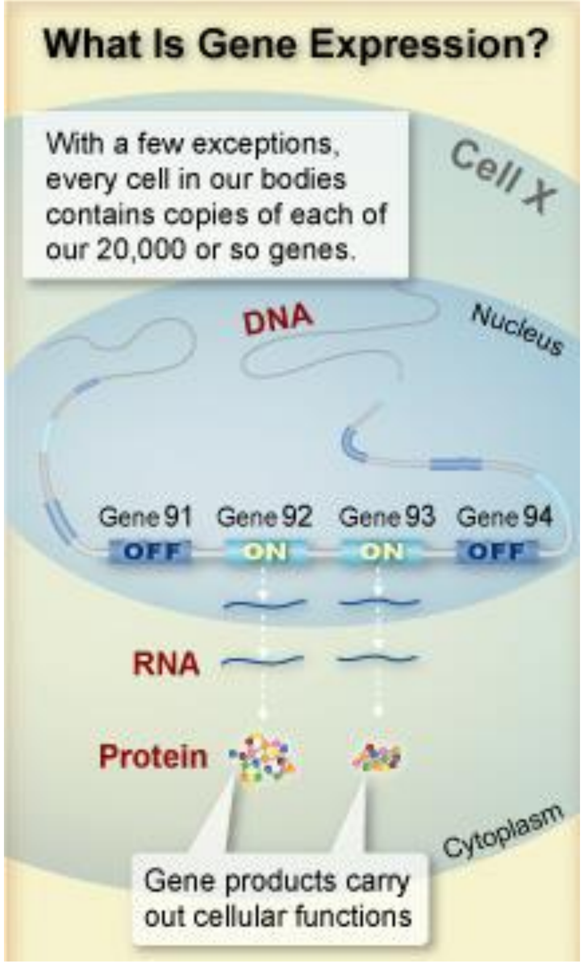



Figure 7–1. Molecular Biology of the Cell, 4th Edition.




- When a gene is "on" and its protein or RNA product is being made, here the gene is being EXPRESSED.
- The on and off states of all of a cell's genes is known as a GENE EXPRESSION PROFILE.
- Each cell type has a unique gene expression profile.

| Insulin | DNA? | Protein? |
|---|---|---|
| Muscle Cell | ✓ | X |
| Pancreatic Cell | ✓ | ✓ |

# Overview

- Prokaryotes and eukaryotes alter gene expression in response to their changing environment (chicken soup, milk or salads; cold, heat, pressure, altitude?)
- In multicellular eukaryotes, gene expression regulates development and is responsible for differences in cell types
- RNA molecules play many roles in regulating gene expression in eukaryotes

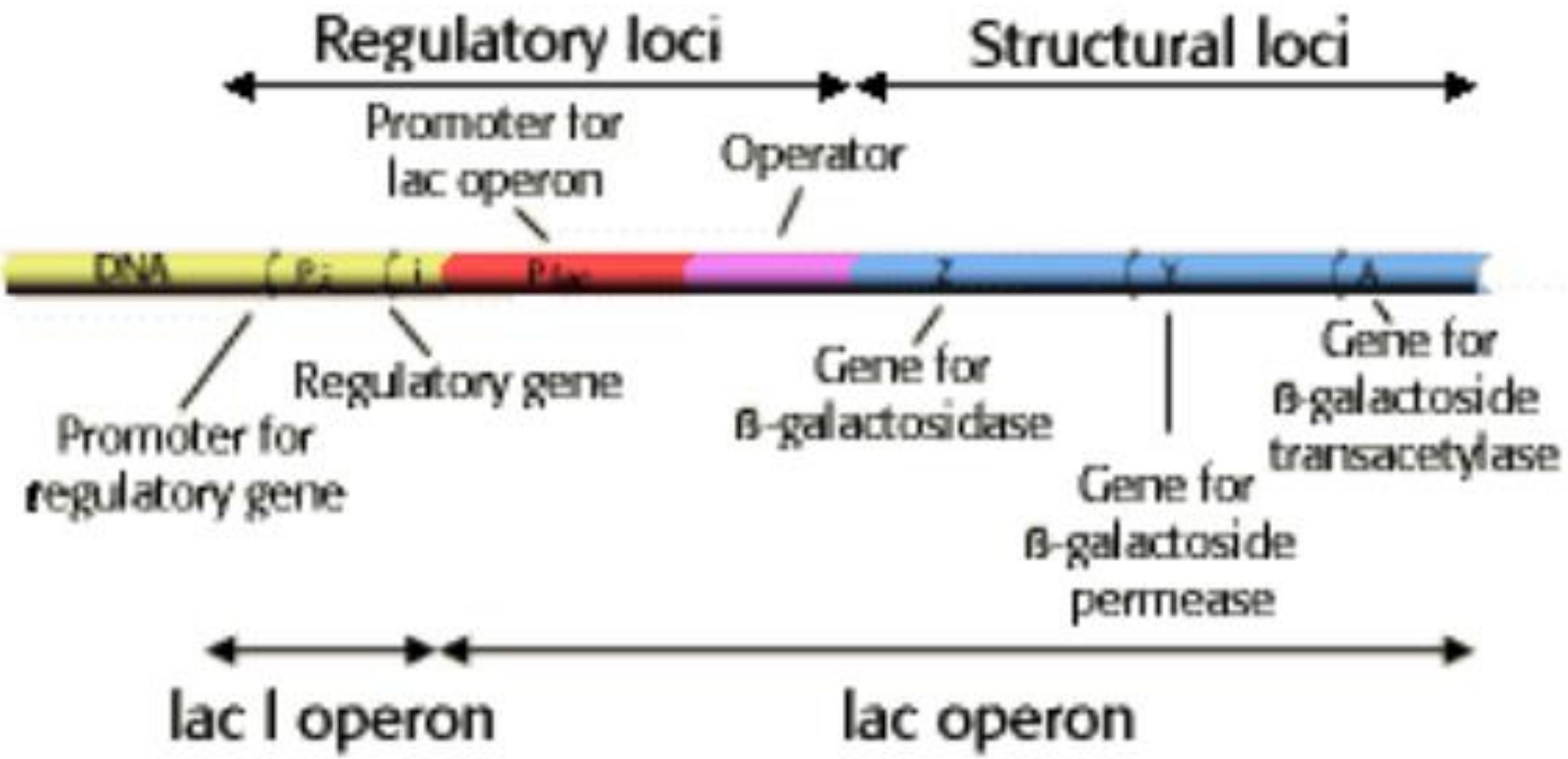
# Bacteria:
# Often respond to environmental change by regulating transcription

- Natural selection has favored bacteria that produce only the products needed by that cell
- A cell can regulate the production of enzymes by feedback inhibition or by gene regulation
- Gene expression in bacteria is controlled by the operon model

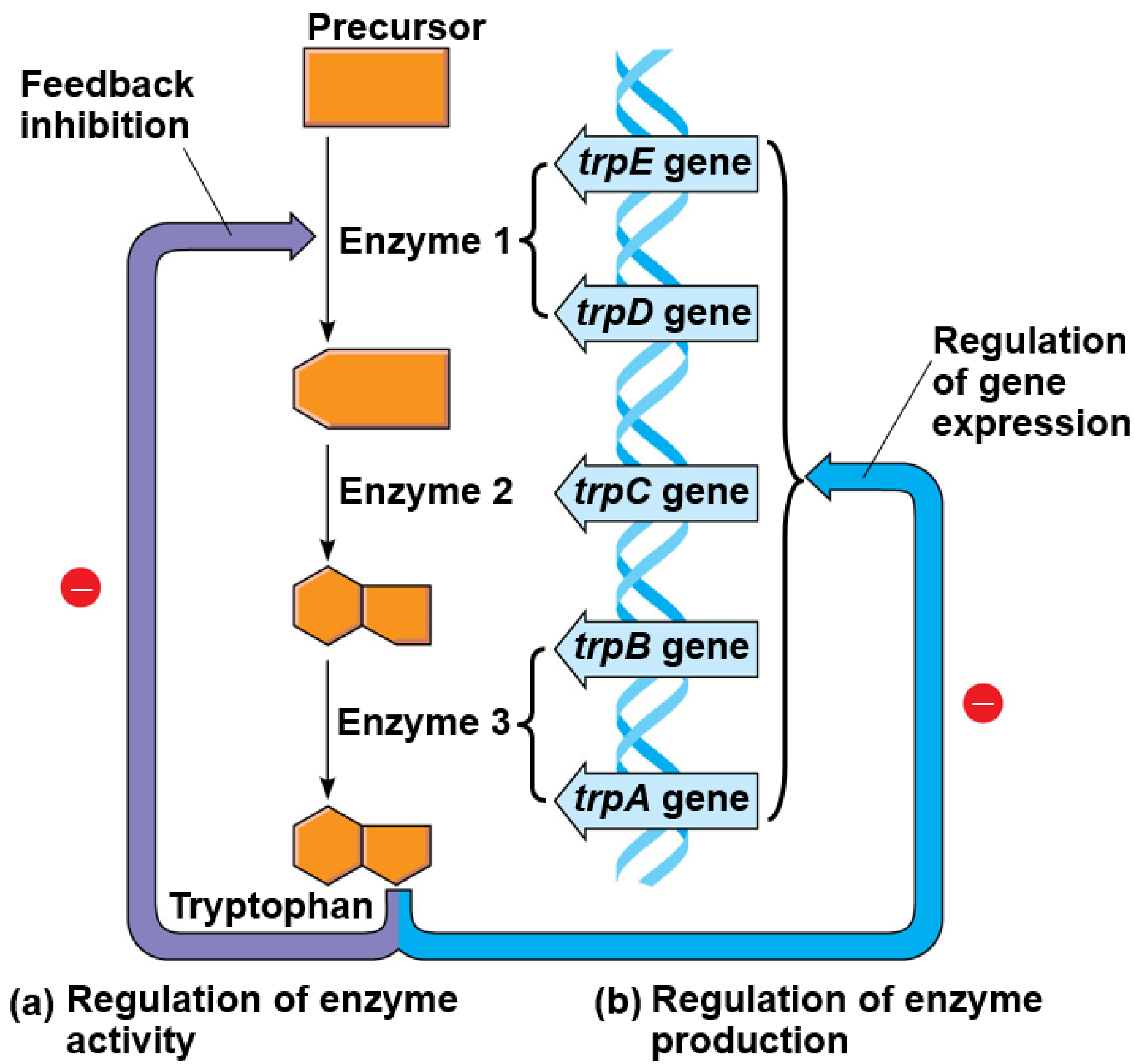
# Operons: The Basic Concept

- A cluster of functionally related genes can be under coordinated control by a single "on-off switch"
- The regulatory "switch" is a segment of DNA called an **operator** usually positioned within the promoter
- An **operon** is the entire stretch of DNA that includes the operator, the promoter, and the genes that they control

# Operon Model



# Tryptophan Synthesis



- The operon can be switched off by a protein **repressor**
- The repressor prevents gene transcription by binding to the operator and blocking RNA polymerase
- The repressor is the product of a separate **regulatory gene**

- The repressor can be in an active or inactive form, depending on the presence of other molecules

- A **corepressor** is a molecule that cooperates with a repressor protein to switch an operon off

- For example, *E. coli* can synthesize the amino acid tryptophan

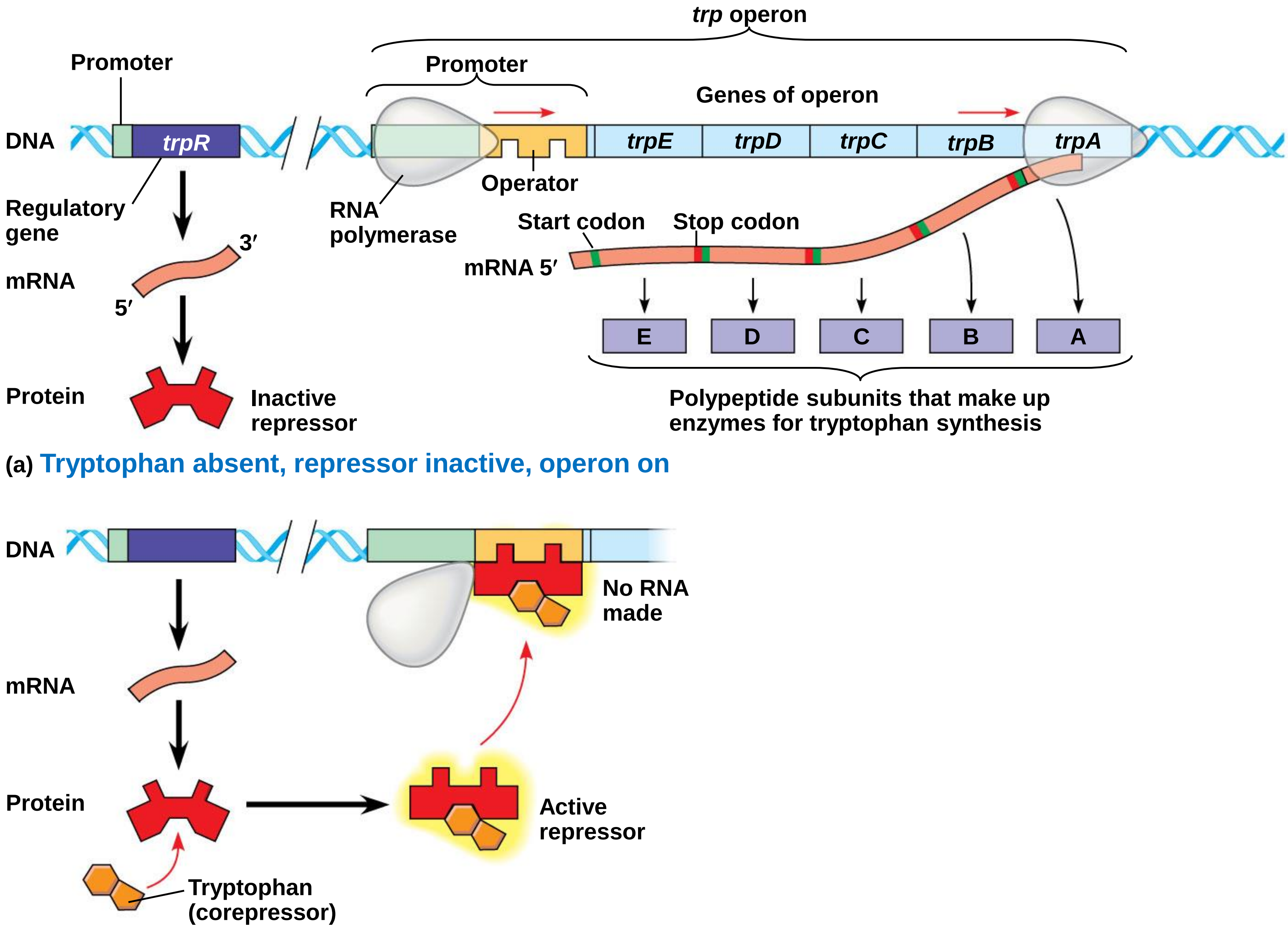
# Two types of Negative Gene Regulation

- A repressible operon is one that is usually on; binding of a repressor to the operator shuts off transcription- *trp* operon is a repressible operon

- An inducible operon is one that is usually off; a molecule called an inducer inactivates the repressor and turns on transcription- *lac* operon

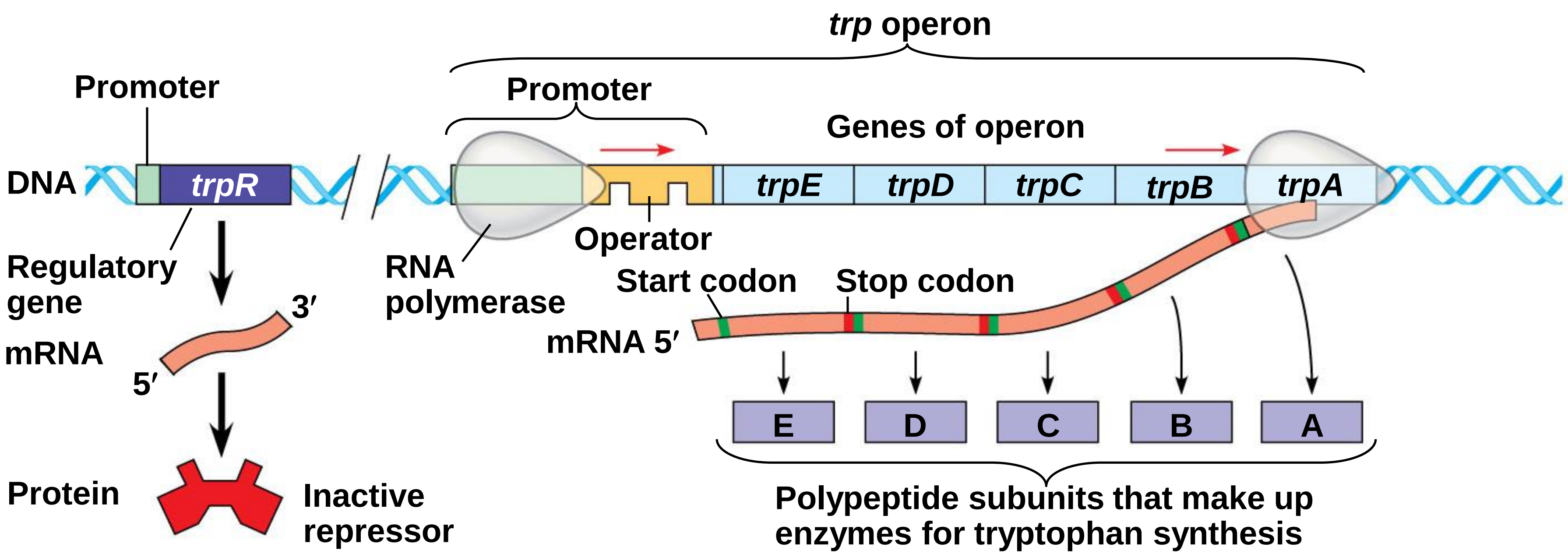
## *trp* operon (repressible)

- By default the *trp* operon is on and the genes for tryptophan synthesis are transcribed
- When tryptophan is present, it binds to the *trp* repressor protein, which turns the operon off
- The repressor is active only in the presence of its corepressor tryptophan; thus the *trp* operon is turned off (repressed) if tryptophan levels are high

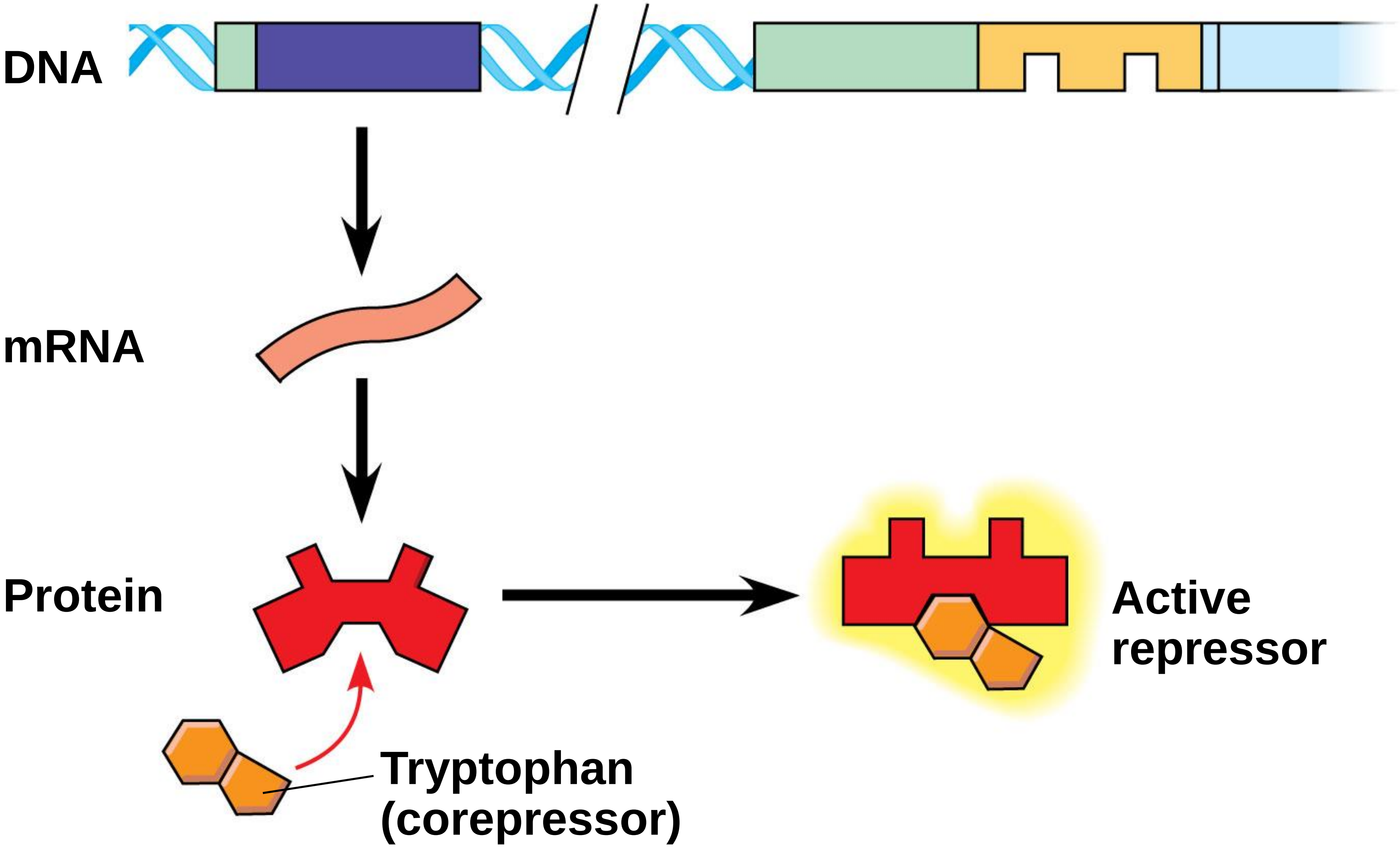



(a) Tryptophan absent, repressor inactive, operon on

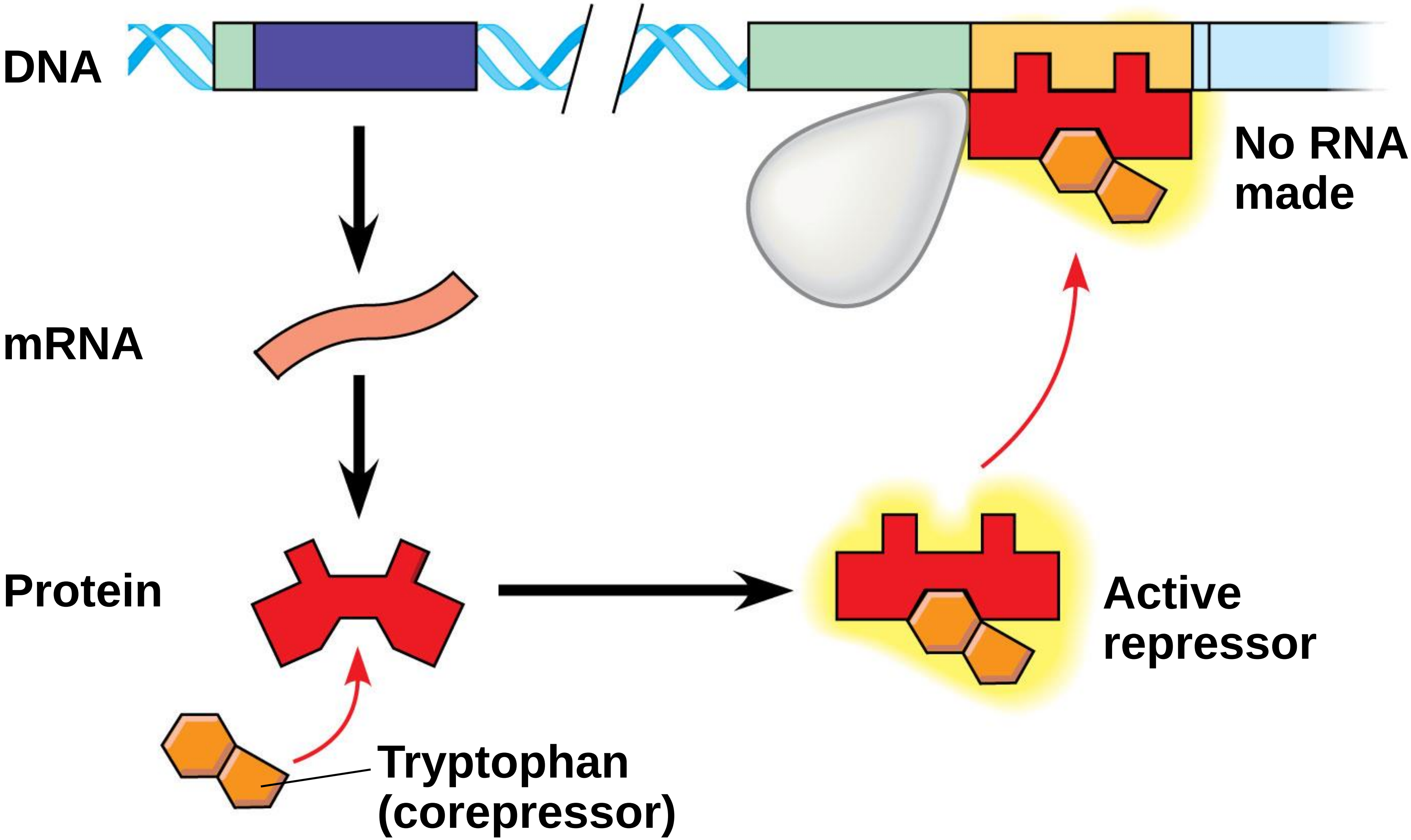
(b) Tryptophan present, repressor active, operon off




**(a) Tryptophan absent, repressor inactive, operon on**



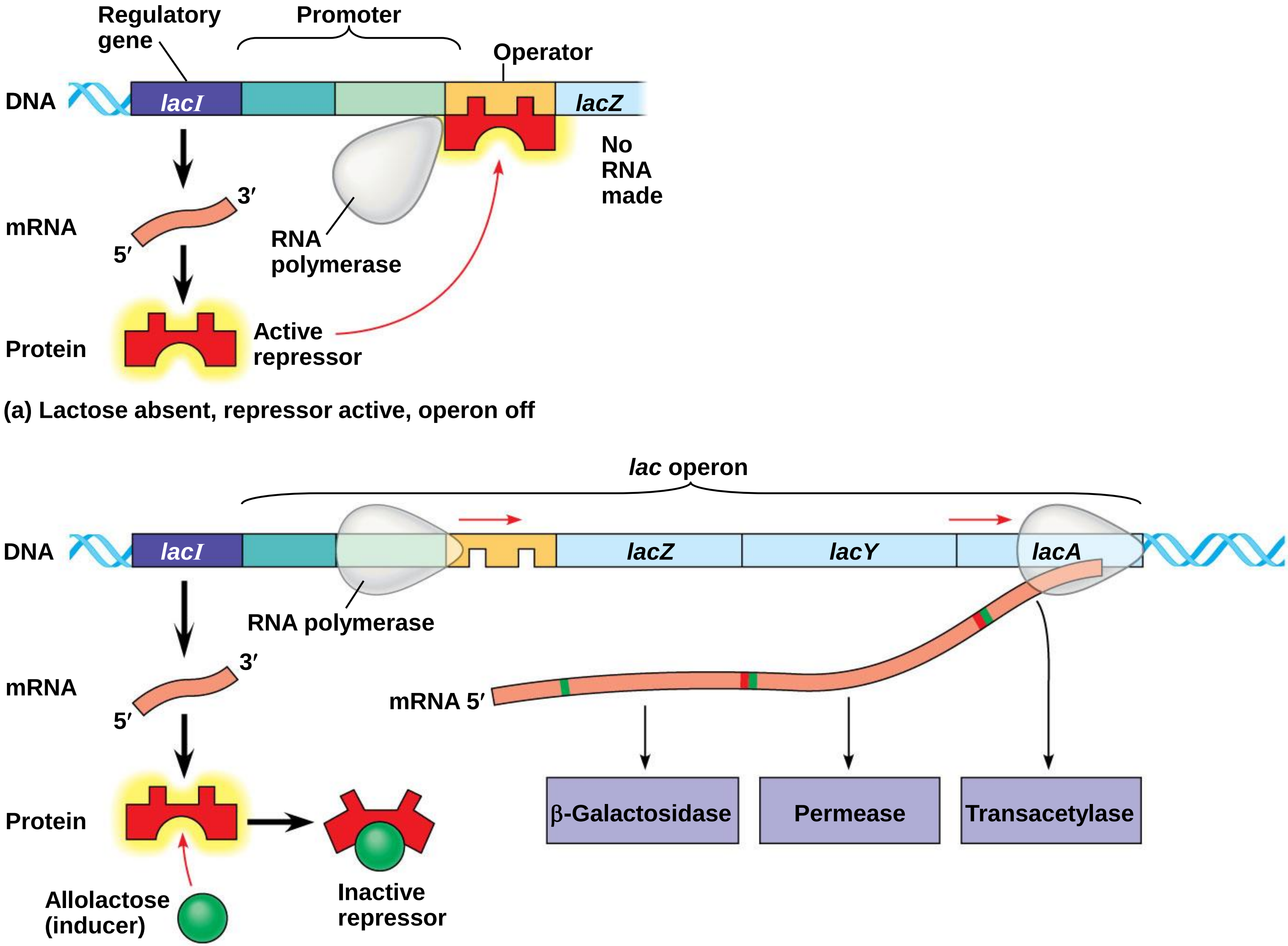

(b) Tryptophan present, repressor active, operon off




(b) Tryptophan present, repressor active, operon off

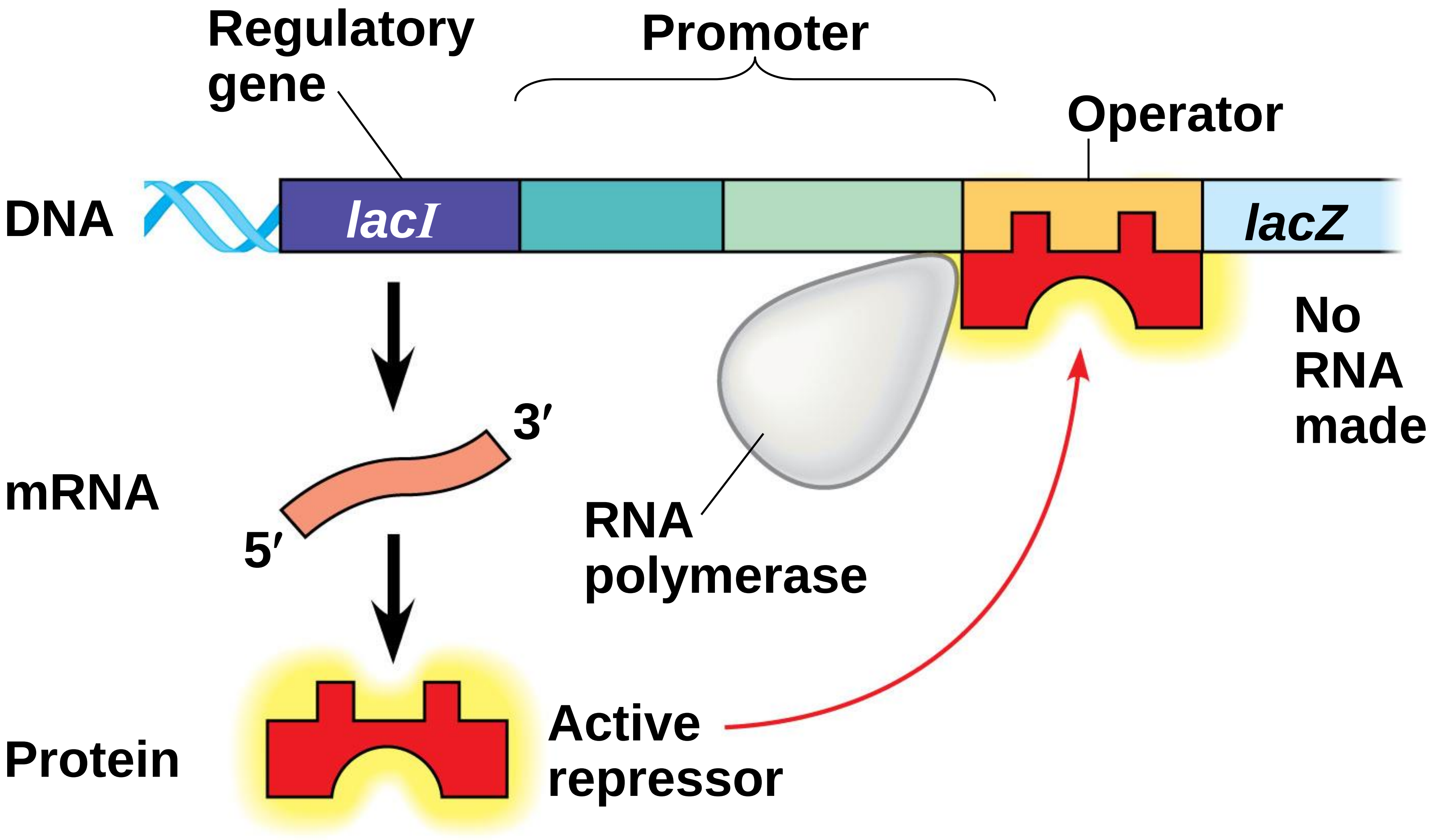
# *lac* operon (Inducible)

- The *lac* operon is an inducible operon and contains genes that code for enzymes used in the hydrolysis and metabolism of lactose
- By itself, the *lac* repressor is active and switches the *lac* operon off
- A molecule called an **inducer** inactivates the repressor to turn the *lac* operon on

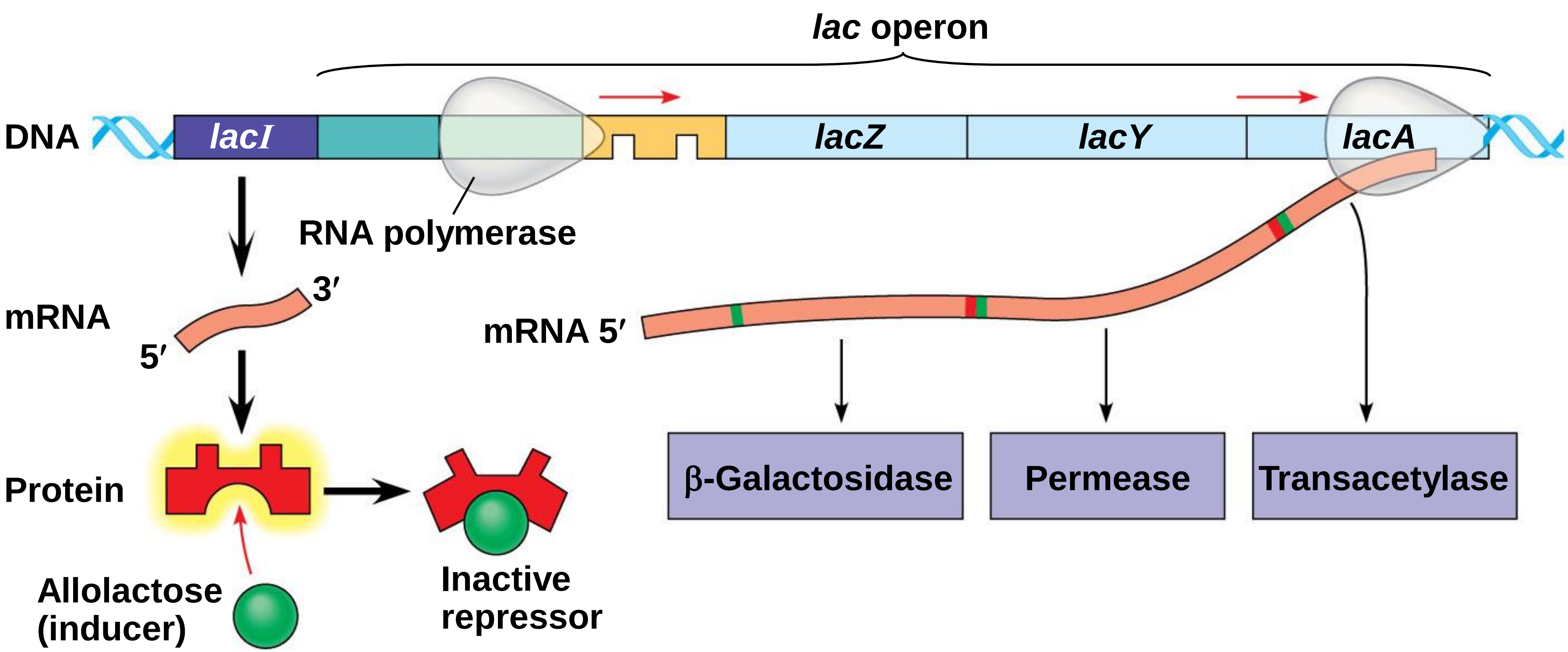



(a) Lactose absent, repressor active, operon off

(b) Lactose present, repressor inactive, operon on




(a) Lactose absent, repressor active, operon off




(b) Lactose present, repressor inactive, operon on

- Inducible enzymes usually function in catabolic pathways; their synthesis is induced by a chemical signal

- Repressible enzymes usually function in anabolic pathways; their synthesis is repressed by high levels of the end product

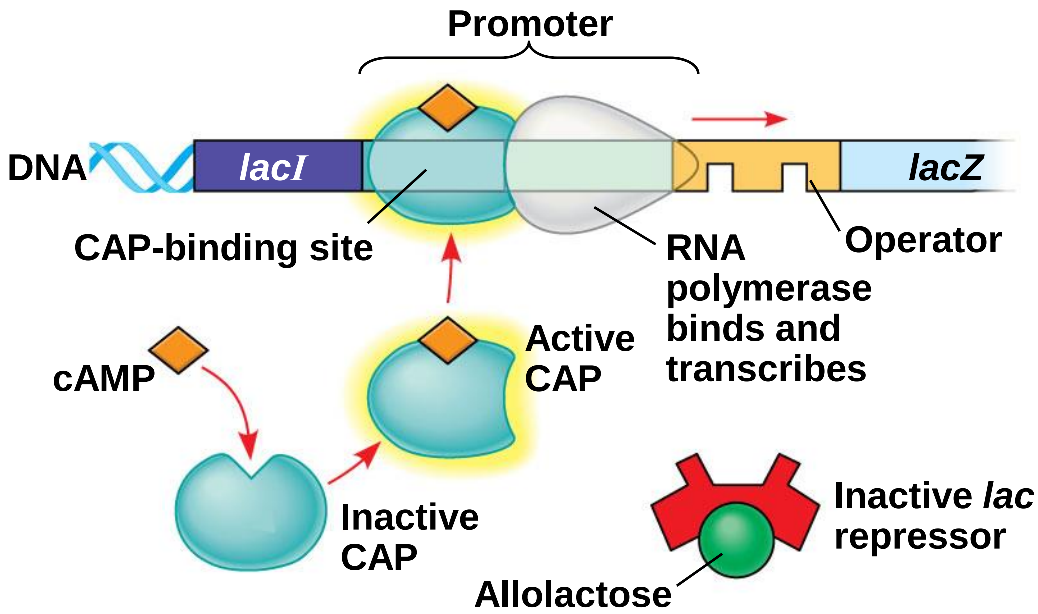
- Regulation of the *trp* and *lac* operons involves negative control of genes because operons are switched off by the active form of the repressor

# Positive Gene Regulation

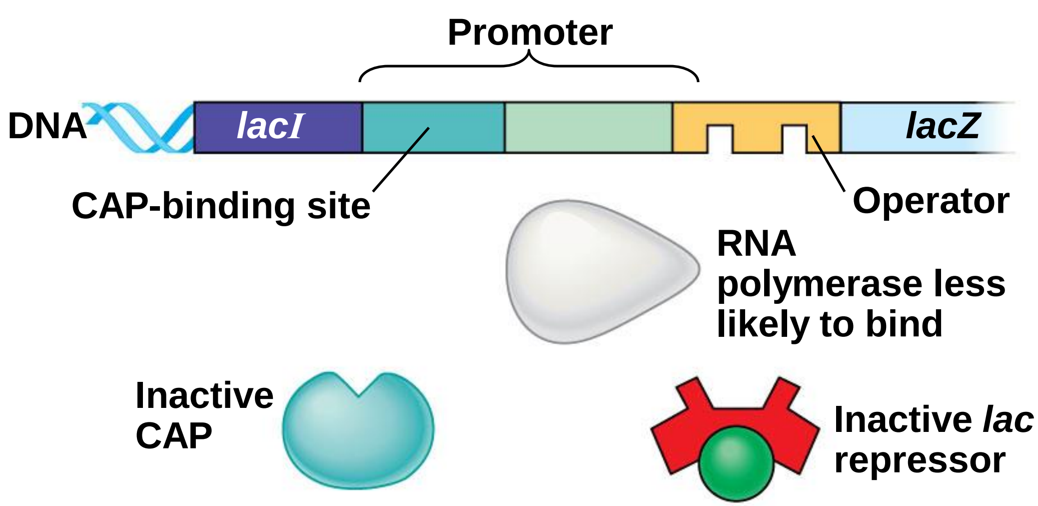
- Some operons are also subject to positive control through a stimulatory protein, such as catabolite activator protein (CAP), an **activator** of transcription
- When glucose is scarse (a preferred food source of *E. coli*), CAP is activated by binding with **cyclic AMP (cAMP)**
- Activated CAP attaches to the promoter of the *lac* operon and increases the affinity of RNA polymerase, thus accelerating transcription

- When glucose levels increase, CAP detaches from the *lac* operon, and transcription returns to a normal rate

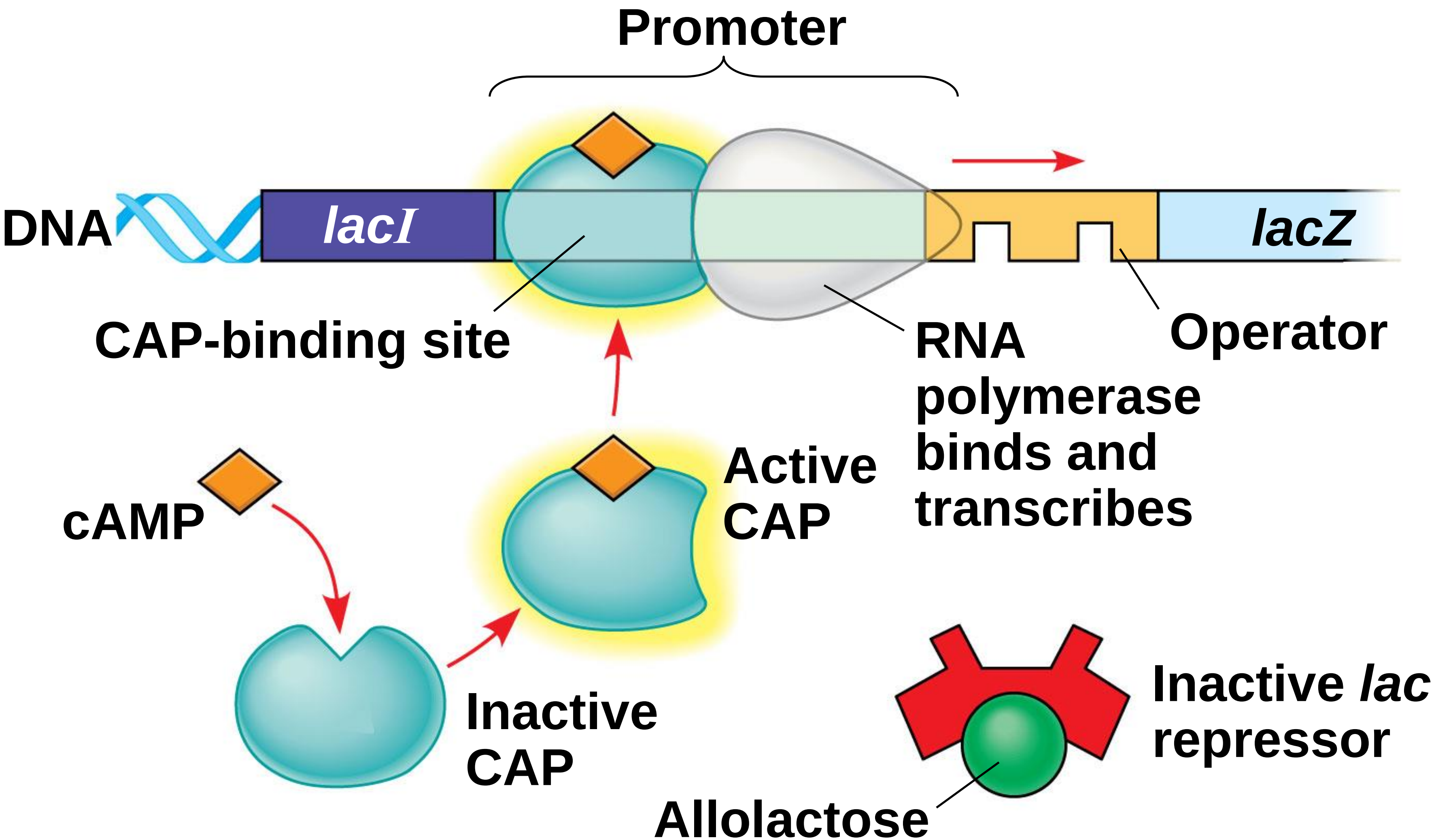
- CAP helps regulate other operons that encode enzymes used in catabolic pathways



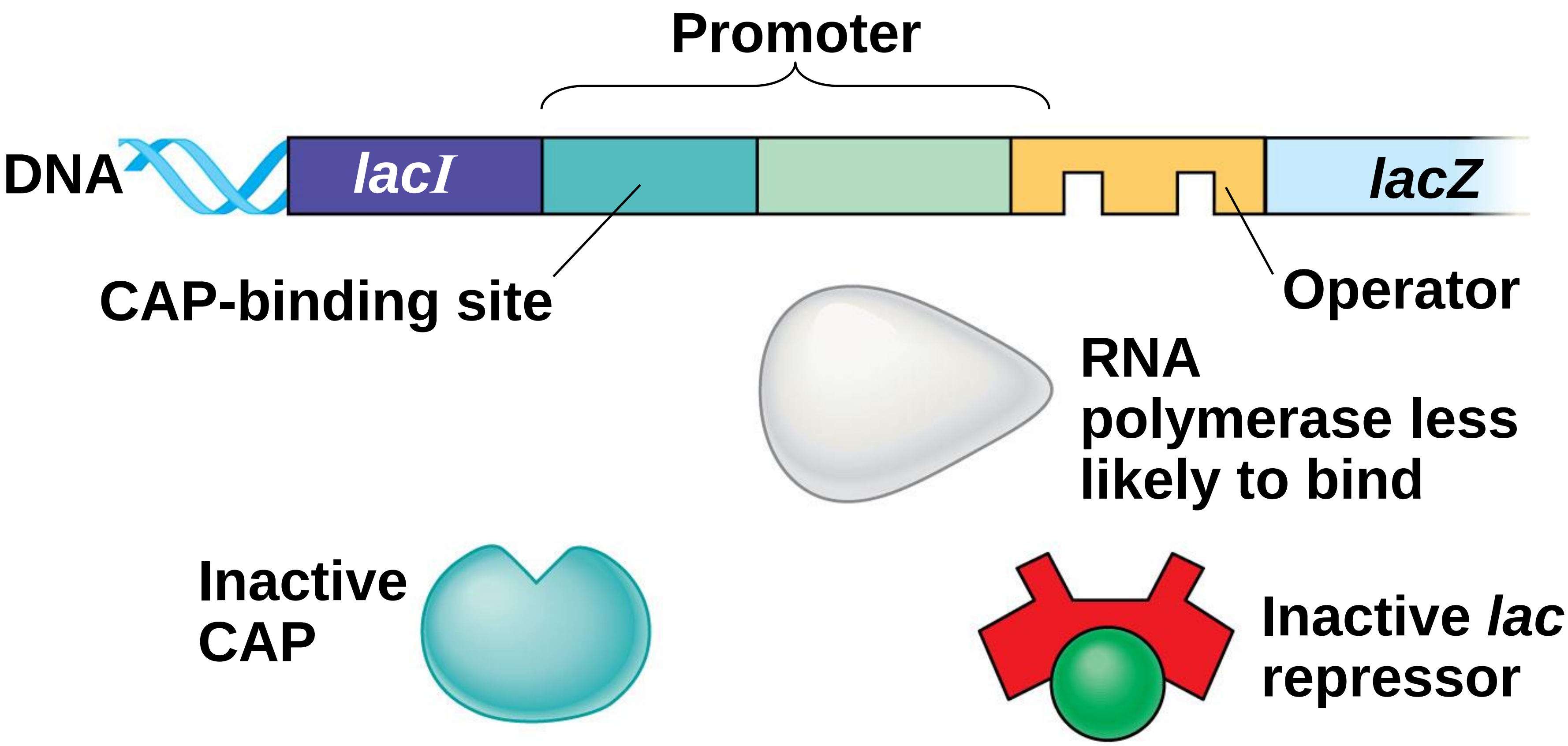

**(a) Lactose present, glucose scarce (cAMP level high): abundant *lac* mRNA synthesized**




**(b) Lactose present, glucose present (cAMP level low): little *lac* mRNA synthesized**




(a) Lactose present, glucose scarce (cAMP level high): abundant *lac* mRNA synthesized

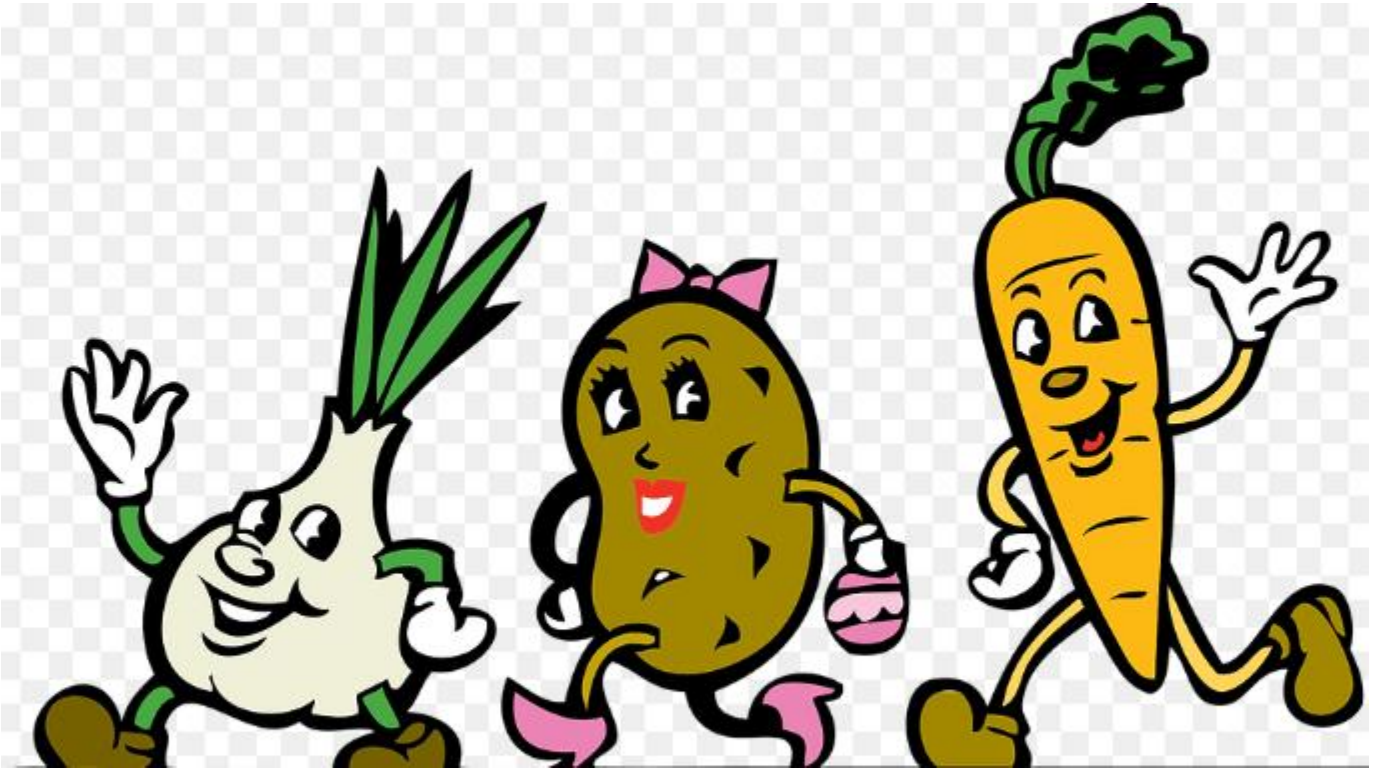



**(b) Lactose present, glucose present (cAMP level low): little *lac* mRNA synthesized**

# Review Questions on Bacterial Gene expression

1. With the use of a diagram, provide an overview of the general regulation strategies available to a bacterial cell.
2. Compare and contrast repressible and inducible operons.
3. What is catabolite repression and how does it work?
4. What elements make up the lac operon and what roles do they play?
5. Describe the process of repression in the *trp* operon.

# Eukaryotic Gene Regulation



# Eukaryotic gene expression is regulated at many stages

- All organisms must regulate which genes are expressed at any given time

- In multicellular organisms regulation of gene expression is essential for cell specialization

# Differential Gene Expression

- Almost all the cells in an organism are **genetically identical**
- **Differences between cell types** result from **differential gene expression.**
- Abnormalities in gene expression can lead to diseases including cancer

# Control of gene expression in Eukaryotic

1) Chromatin modifications (structural reg)
   i. DNA methylation
   ii. Histone acetylation

2) Control of transcription

3) Alternative splicing

4) Degradation of mRNA

5) Blockage of translation

**Six Steps** at which eukaryotic gene expression are controlled

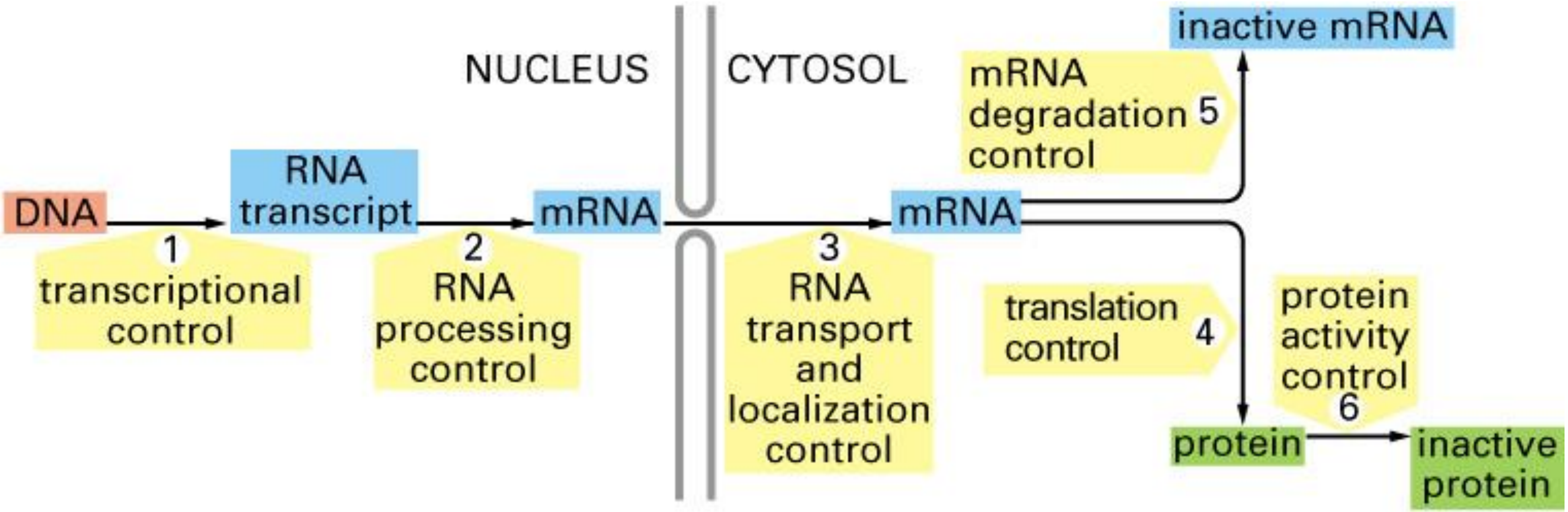
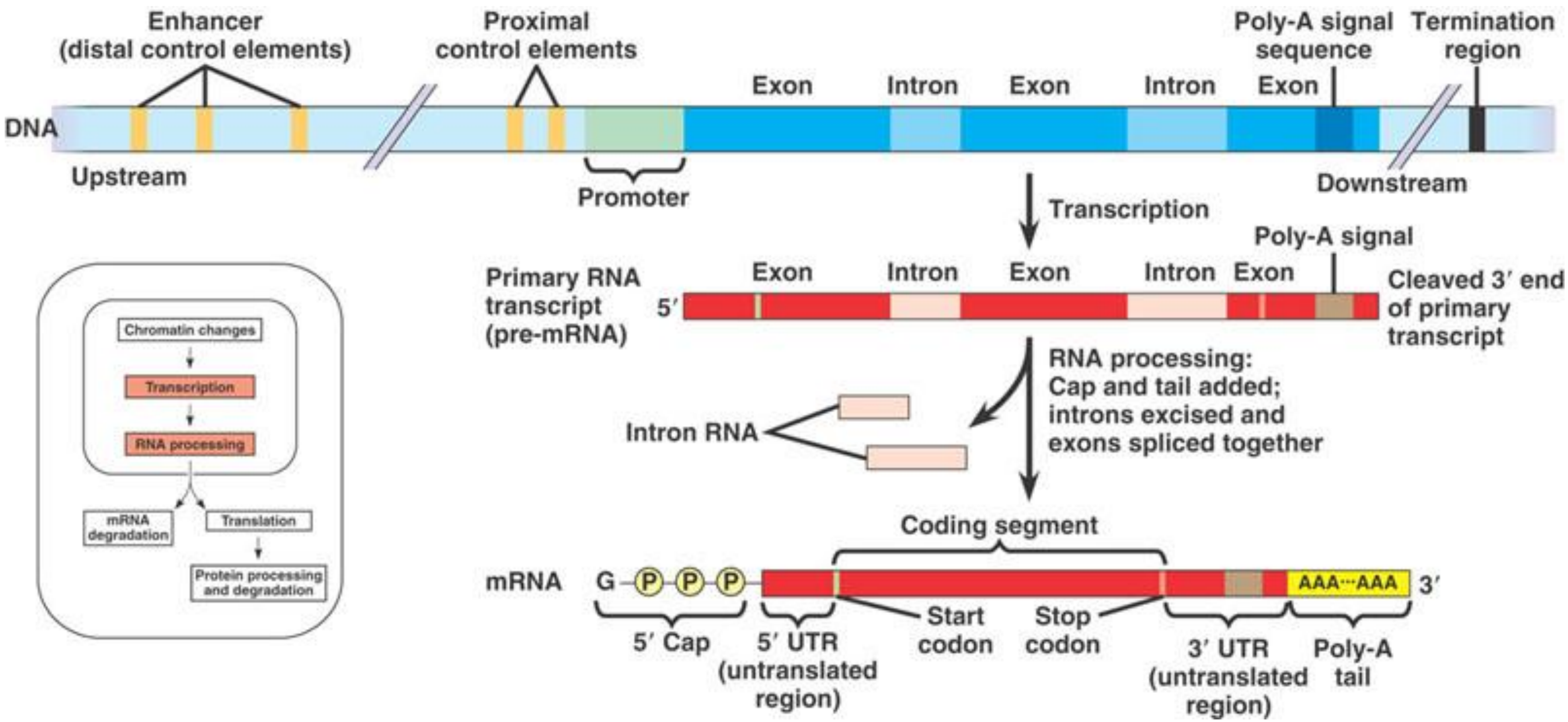
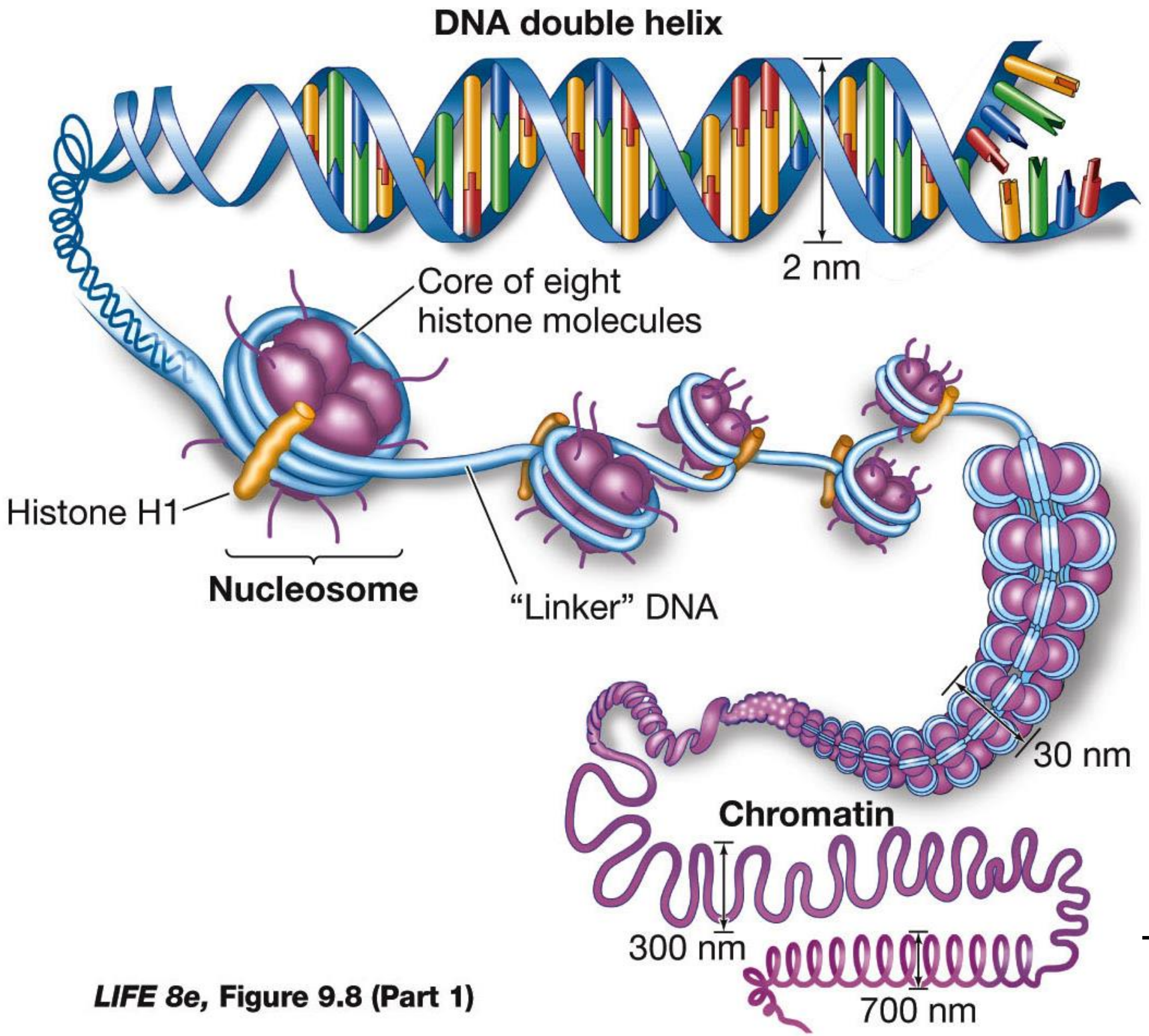



Figure 7–5. Molecular Biology of the Cell, 4th Edition.

# A eukaryotic gene with its control elements and transcript

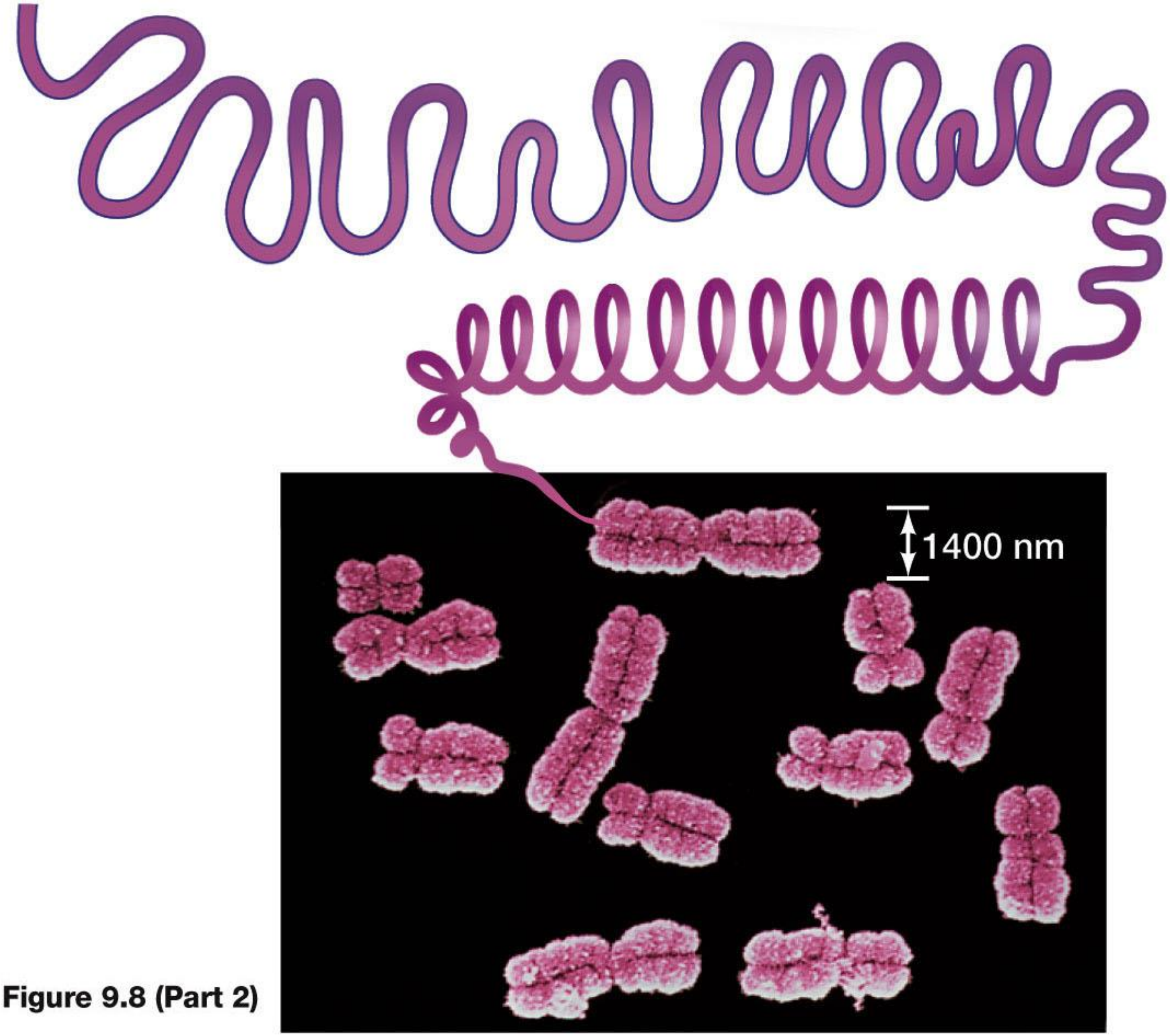


# Structural Regulation: Eukaryote DNA packing




*LIFE 8e*, **Figure 9.8 (Part 1)**

To next figure



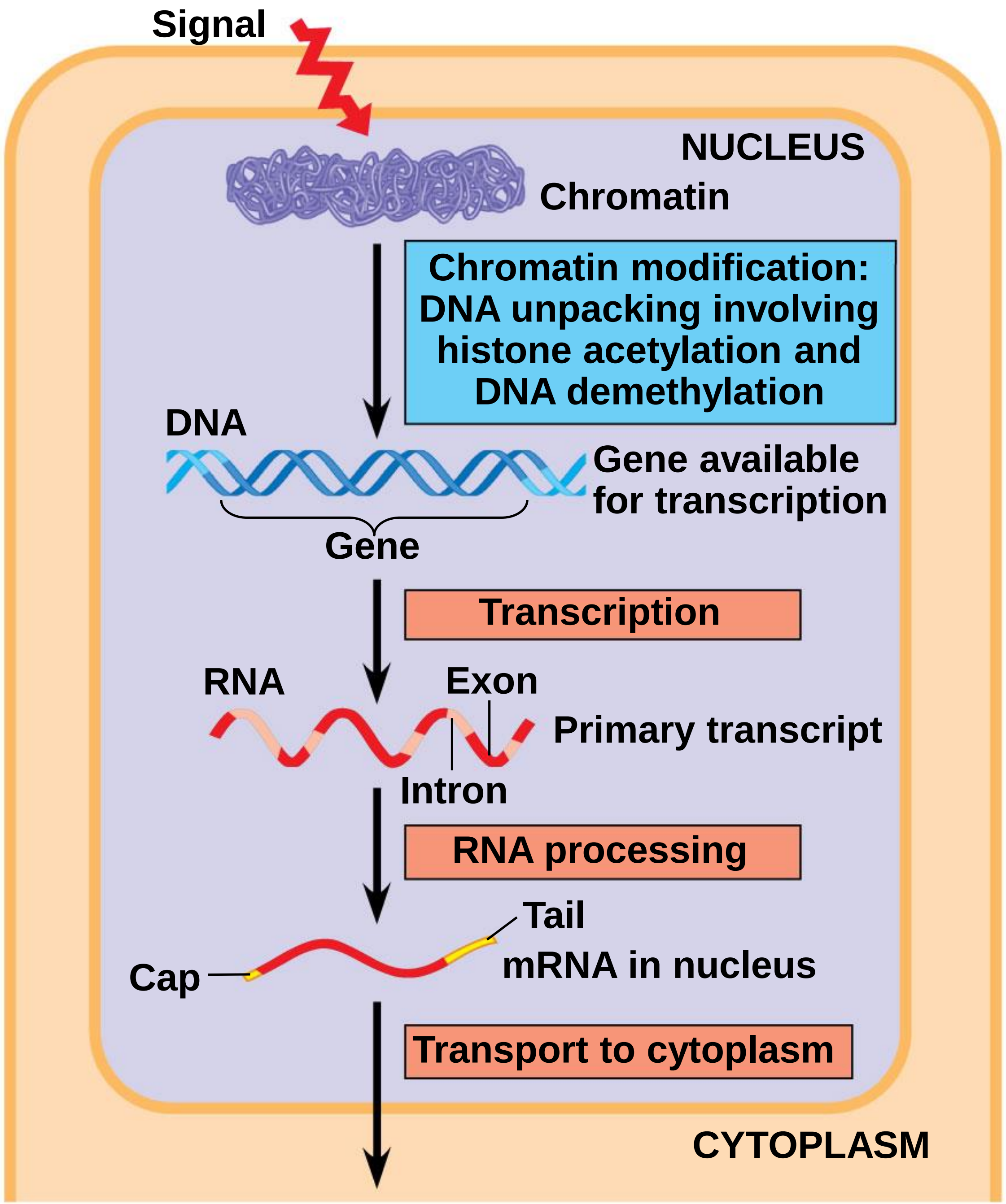

Metaphase chromosomes

*LIFE 8e*, Figure 9.8 (Part 2)

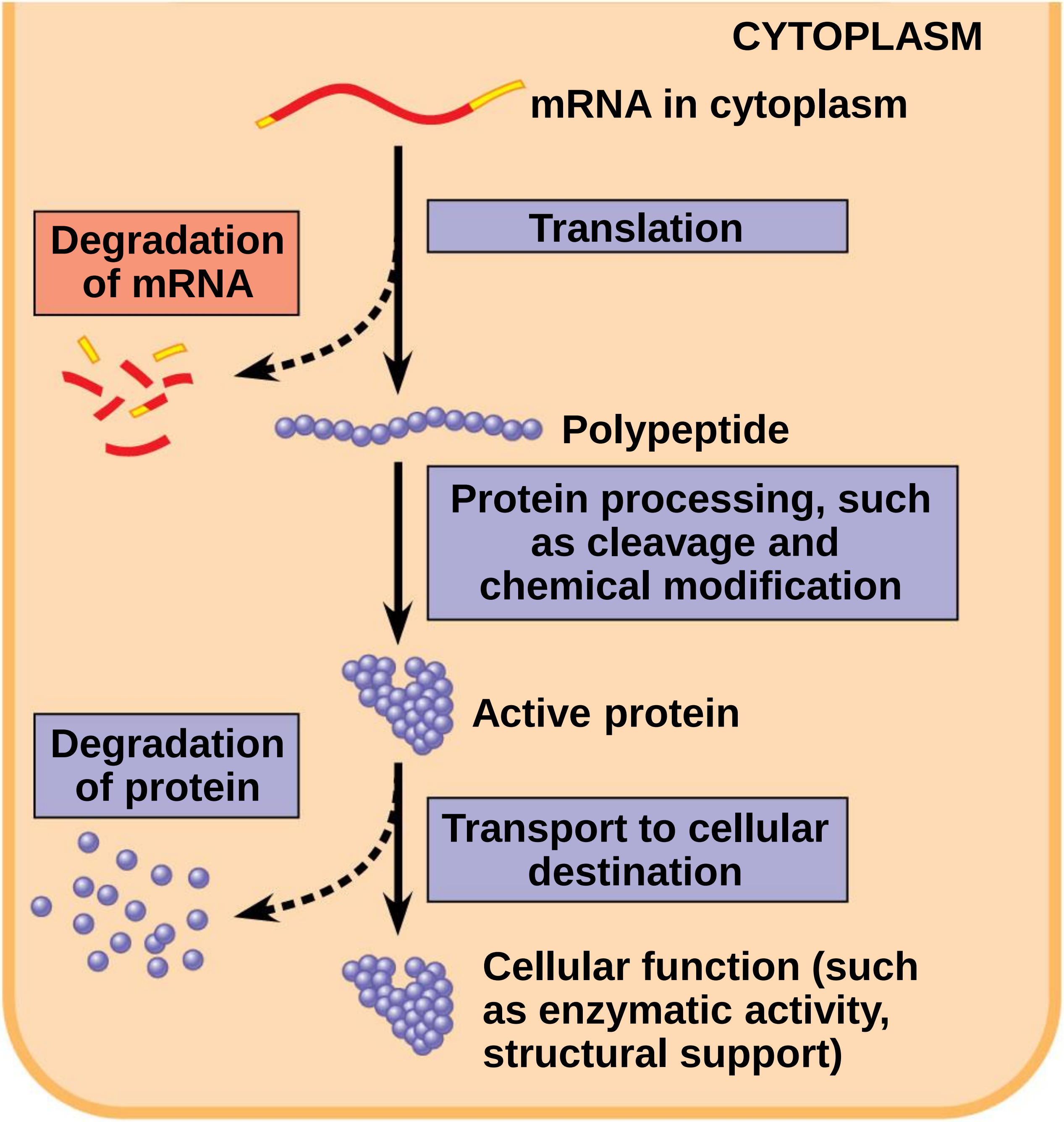
# Regulation of Chromatin Structure

- Genes within highly packed heterochromatin are usually not expressed. Why?

- Chemical modifications to histones and DNA of chromatin influence both chromatin structure and gene expression. How?

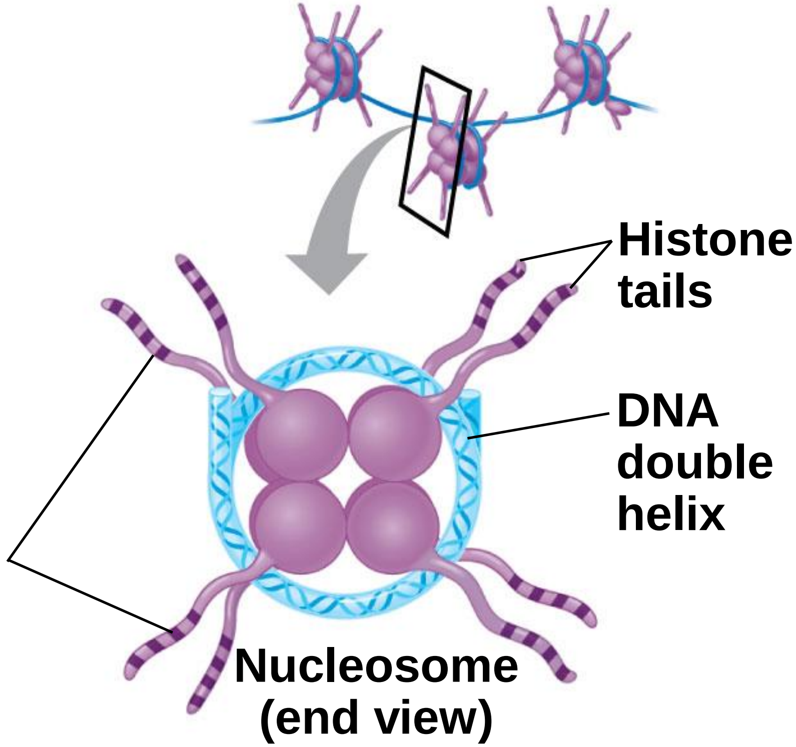


Signal
NUCLEUS
Chromatin
Chromatin modification: DNA unpacking involving histone acetylation and DNA demethylation
DNA
Gene available for transcription
Gene
Transcription
RNA
Exon
Primary transcript
Intron
RNA processing
Tail
mRNA in nucleus
Cap
Transport to cytoplasm
CYTOPLASM

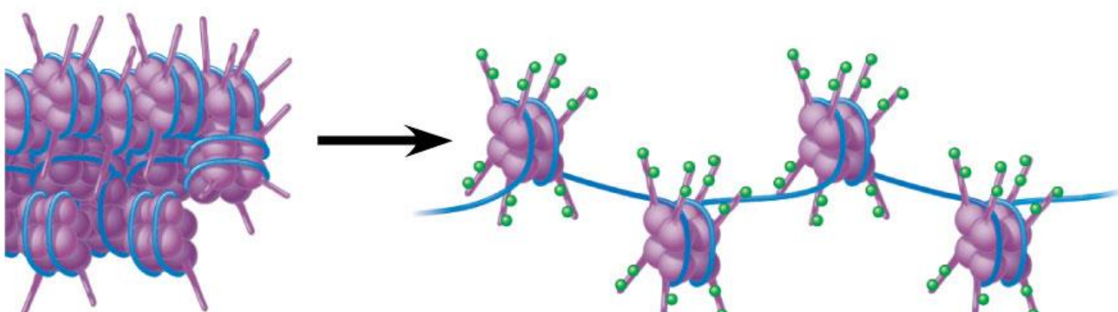

CYTOPLASM
mRNA in cytoplasm
Translation
Degradation of mRNA
Polypeptide
Protein processing, such as cleavage and chemical modification
Active protein
Degradation of protein
Transport to cellular destination
Cellular function (such as enzymatic activity, structural support)

# *Histone Modifications*

- **Methyl groups:** (methylation) condense chromatin.
- **Histone acetylation:** acetyl groups are attached to positively charged lysine's in histone tails. This loosens chromatin structure, thereby promoting the initiation of transcription
- **Phosphate groups:** (phosphorylation) next to a methylated amino acid loosens chromatin




(a) Histone tails protrude outward from a nucleosome




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

# The ***histone code hypothesis***

- Proposes that specific combinations of modifications, as well as
- The order in which they occur, help determine chromatin configuration and influence transcription

# *DNA Methylation*

- **DNA methylation**, the addition of methyl groups to certain bases in DNA, is associated with reduced transcription in some species
- DNA methylation can cause long-term inactivation of genes in cellular differentiation
- In genomic imprinting, methylation regulates expression of either the maternal or paternal alleles of certain genes at the start of development- case of X*Y chromosome

# *Epigenetic Inheritance*

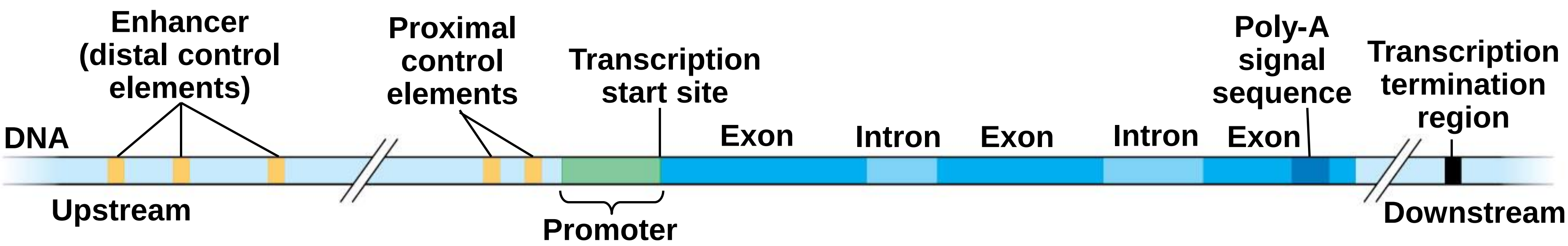
- Although the chromatin modifications just discussed do not alter DNA sequence, they may be passed to future generations of cells
- The inheritance of traits transmitted by mechanisms not directly involving the nucleotide sequence is called **epigenetic inheritance**

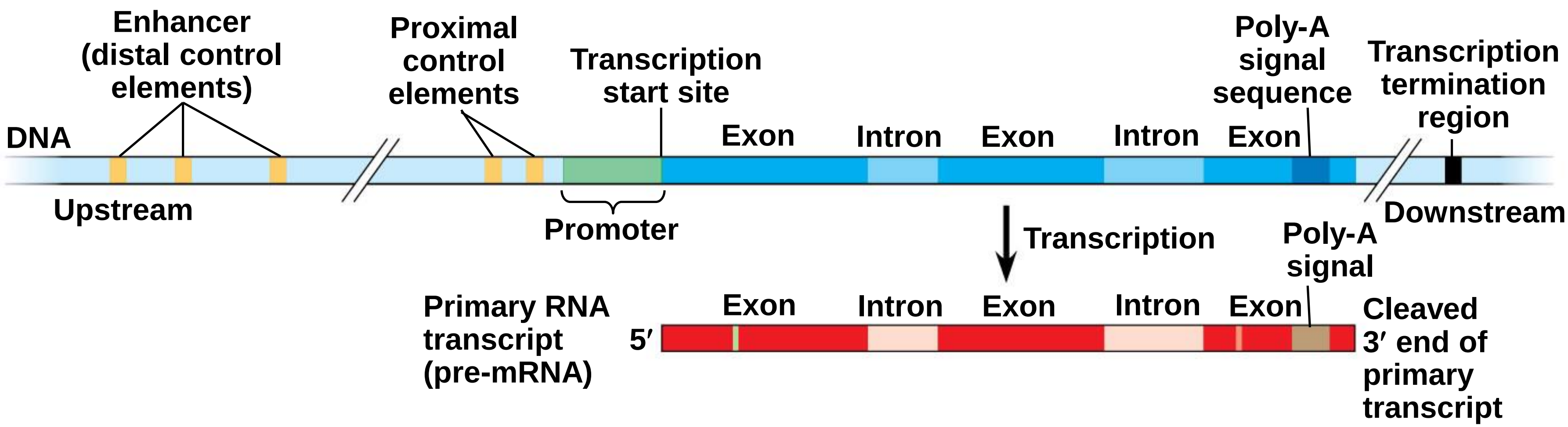
# Regulation of Transcription Initiation / Transcriptional regulation

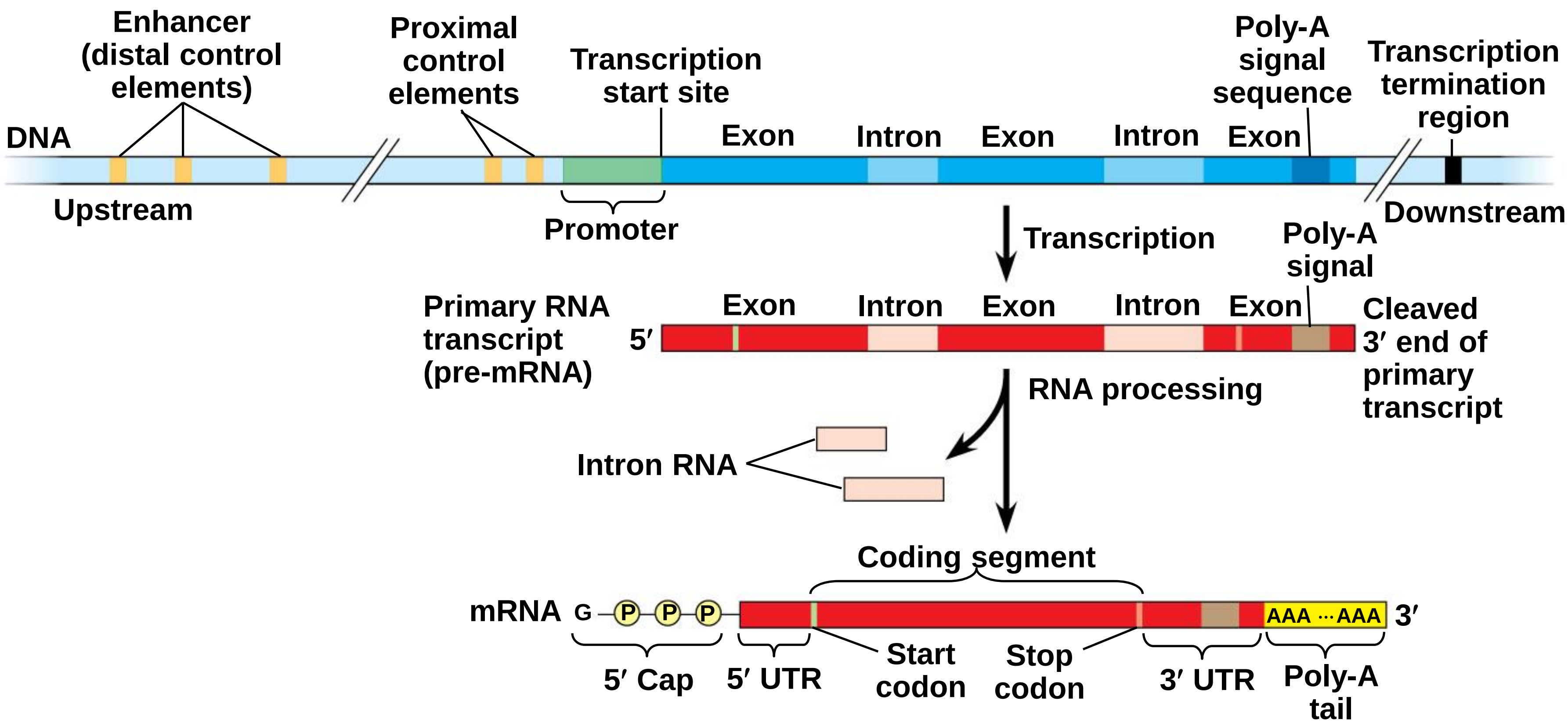
- Chromatin-modifying enzymes provide initial control of gene expression by making a region of DNA either more or less able to bind the transcription machinery

# *Organization of a Typical Eukaryotic Gene*

- Associated with most eukaryotic genes are multiple **control elements**, segments of noncoding DNA that serve as binding sites for transcription factors that help regulate transcription
- Control elements and the transcription factors they bind are critical to the precise regulation of gene expression in different cell types


Enhancer
(distal control
elements)
Proximal
control
elements
Transcription
start site
Poly-A
signal
sequence
Transcription
termination
region
DNA
Exon
Intron
Exon
Intron
Exon
Upstream
Promoter
Downstream


Enhancer
(distal control
elements)
Proximal
control
elements
Transcription
start site
Poly-A
signal
sequence
Transcription
termination
region
DNA
Exon
Intron
Exon
Intron
Exon
Upstream
Promoter
Downstream
Transcription
Poly-A
signal
Primary RNA
transcript
(pre-mRNA)
5′
Exon
Intron
Exon
Intron
Exon
Cleaved
3′ end of
primary
transcript



Enhancer (distal control elements)
Proximal control elements
Transcription start site
Poly-A signal sequence
Transcription termination region
DNA
Exon
Intron
Exon
Intron
Exon
Upstream
Promoter
Downstream
Transcription
Poly-A signal
Primary RNA transcript (pre-mRNA)
5′
Exon
Intron
Exon
Intron
Exon
Cleaved 3′ end of primary transcript
RNA processing
Intron RNA
Coding segment
mRNA
G
P
P
P
AAA ··· AAA
3′
5′ Cap
5′ UTR
Start codon
Stop codon
3′ UTR
Poly-A tail

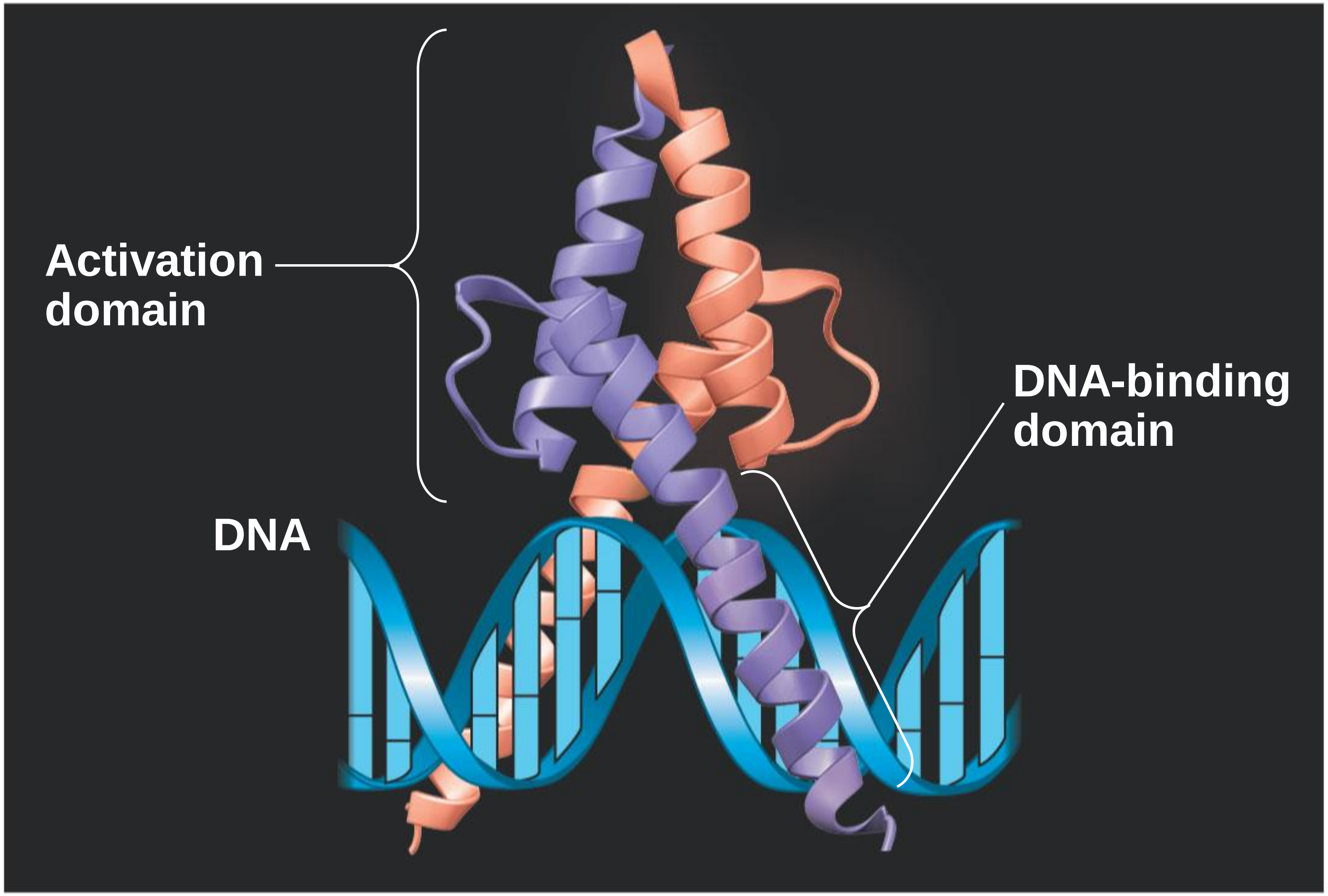
# *The Roles of Transcription Factors*

- To initiate transcription, eukaryotic RNA polymerase requires the assistance of proteins called **transcription factors**
- General transcription factors are essential for the transcription of all protein-coding genes
- In eukaryotes, high levels of transcription of particular genes depend on control elements interacting with specific transcription factors

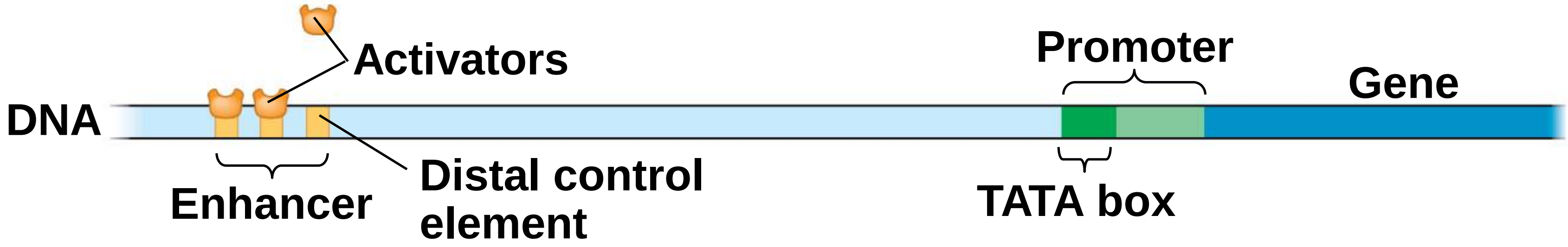
## Enhancers and Specific Transcription Factors

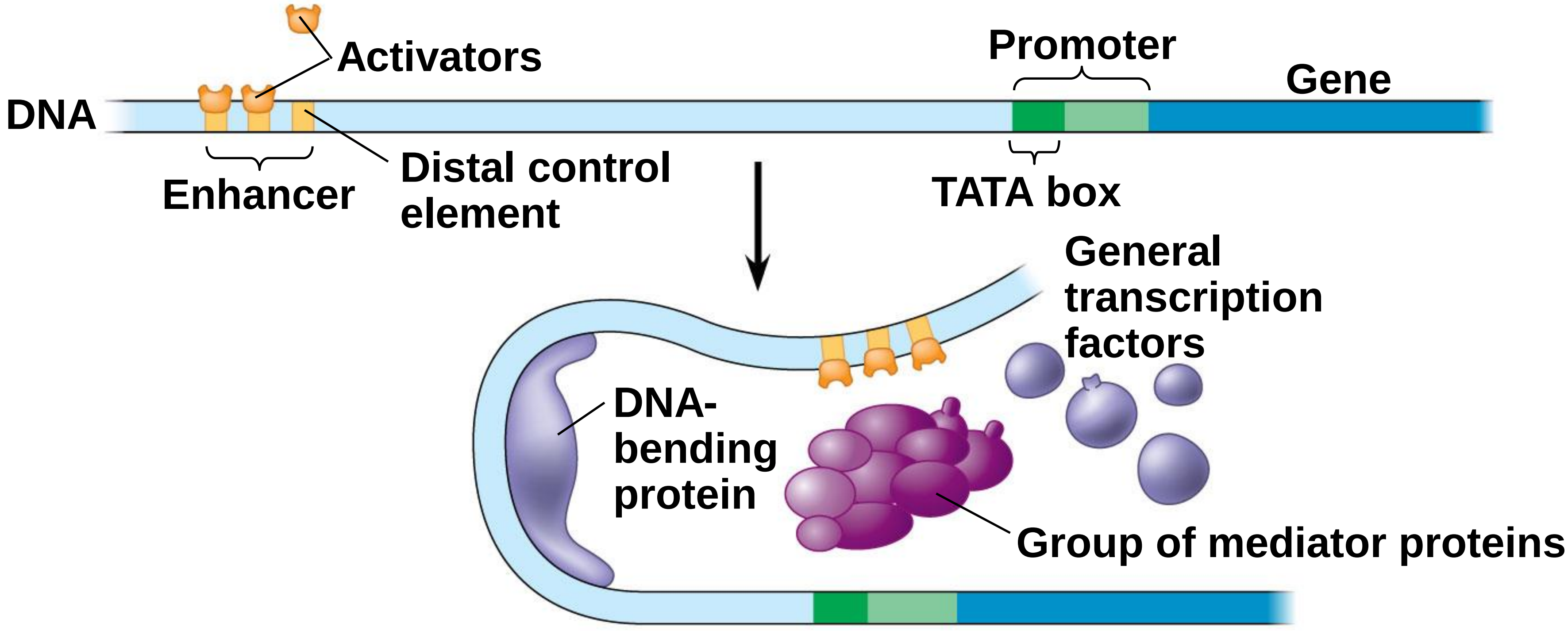
- Proximal control elements are located close to the promoter
- Distal control elements, groupings of which are called **enhancers**, may be far away from a gene or even located in an intron

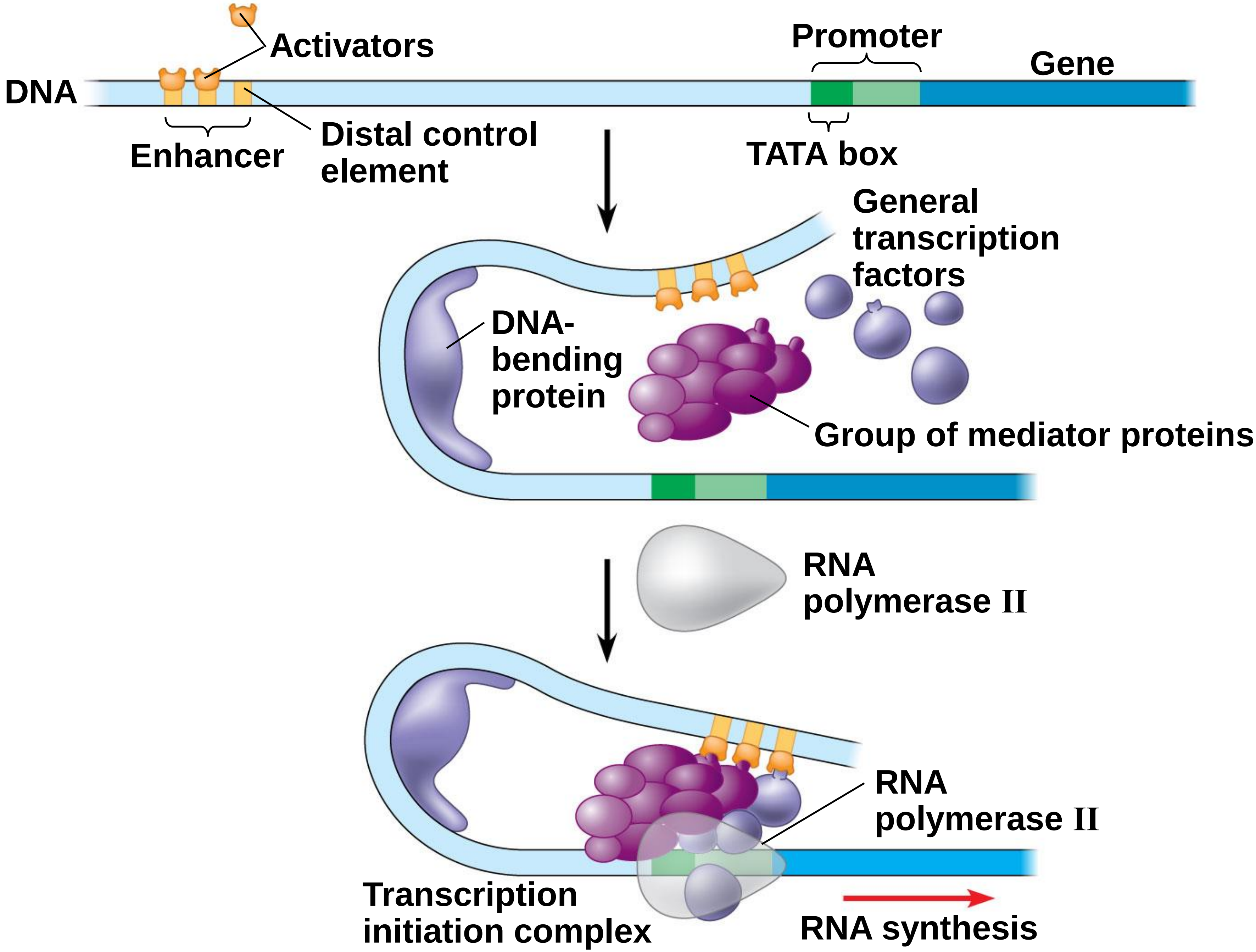
- An activator is a protein that binds to an enhancer and stimulates transcription of a gene
- Activators have two domains, one that binds DNA and a second that activates transcription
- Bound activators facilitate a sequence of protein-protein interactions that result in transcription of a given gene


Activation domain
DNA-binding domain
DNA

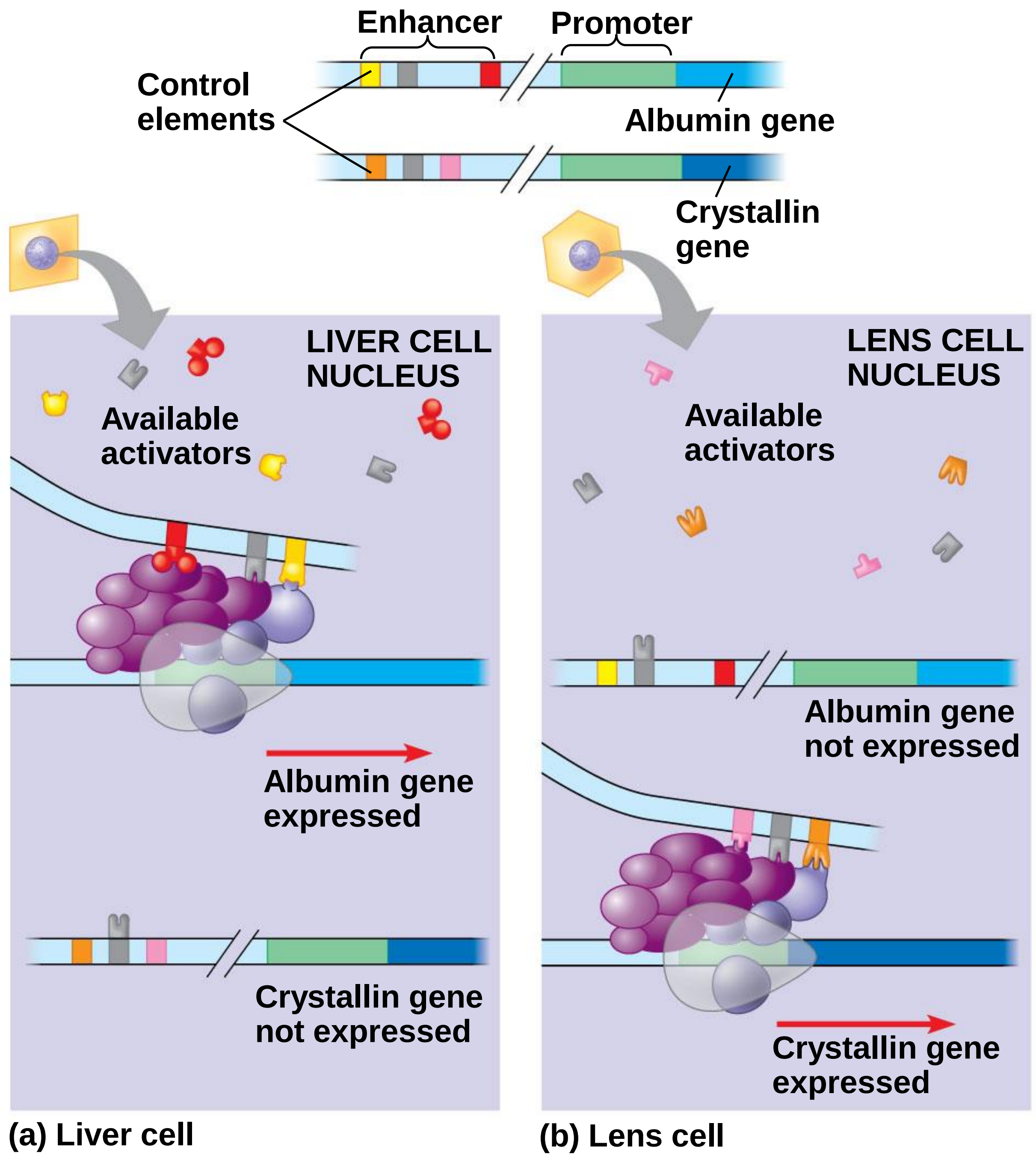
- Some transcription factors function as repressors, inhibiting expression of a particular gene by a variety of methods
- Some activators and repressors act indirectly by influencing chromatin structure to promote or **silence transcription**

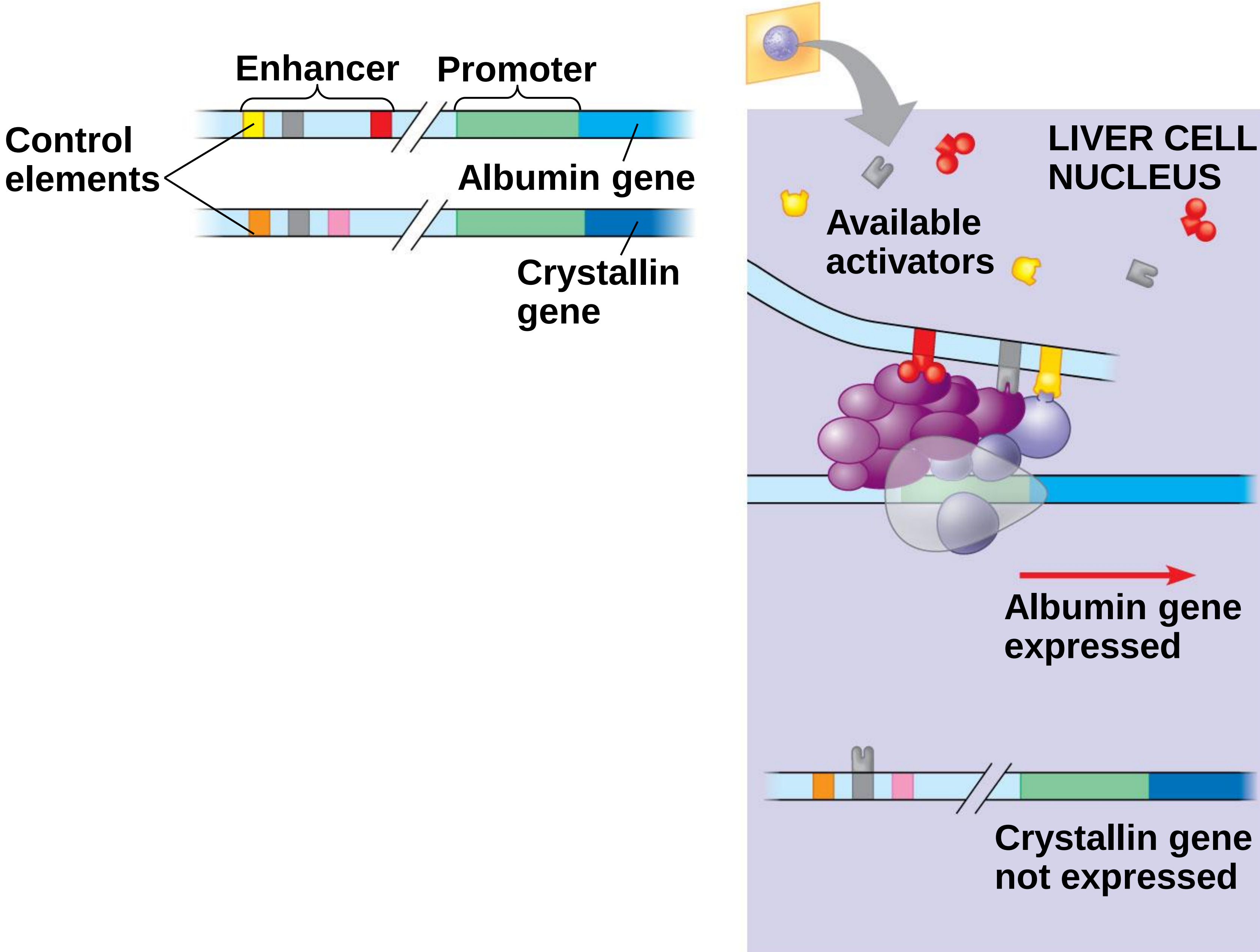

Activators
Promoter
Gene
DNA
Enhancer
Distal control element
TATA box


Activators
Promoter
Gene
DNA
Enhancer
Distal control element
TATA box
General transcription factors
DNA-bending protein
Group of mediator proteins

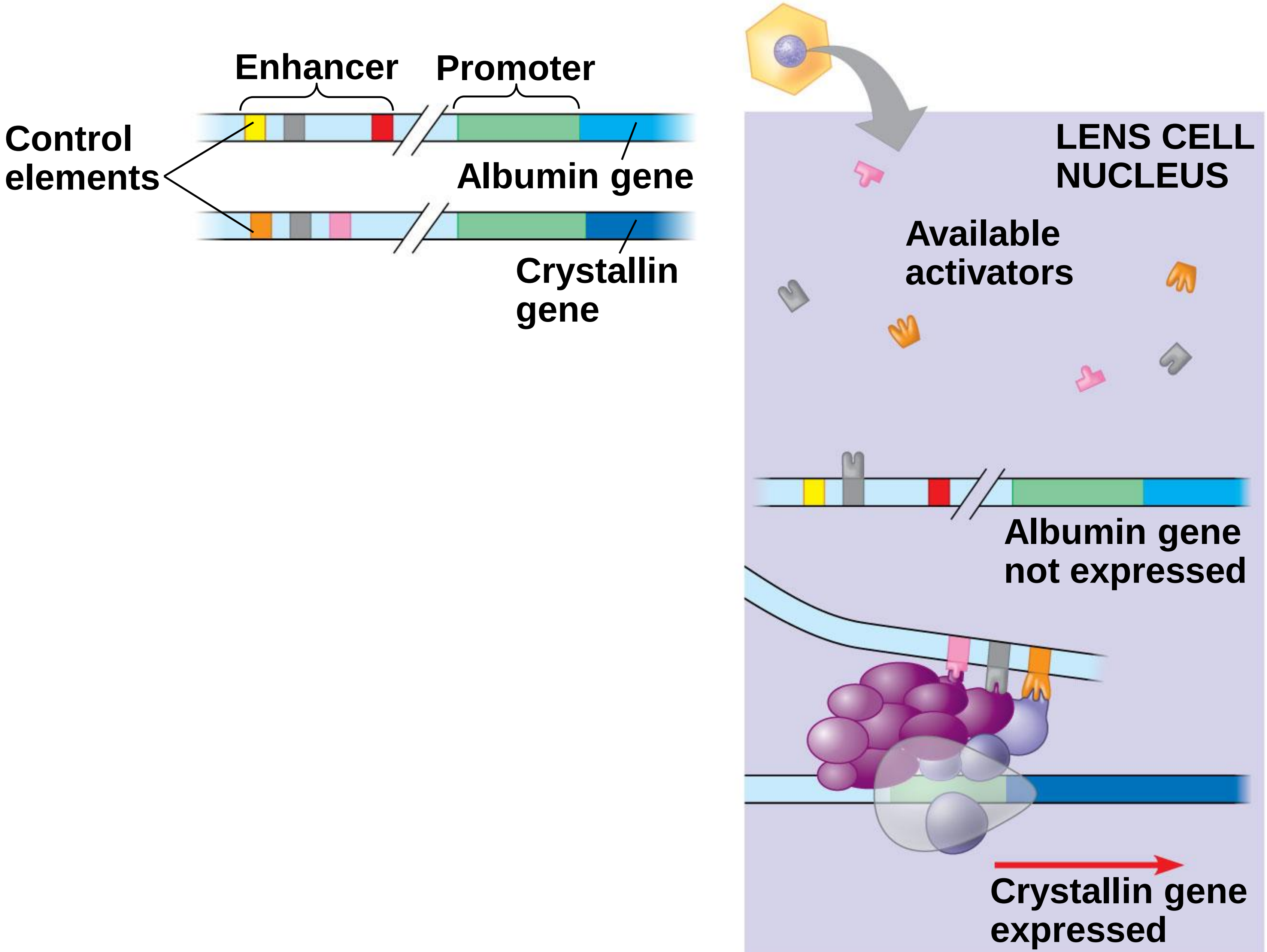


Activators
DNA
Promoter
Gene
Enhancer
Distal control element
TATA box
General transcription factors
DNA-bending protein
Group of mediator proteins
RNA polymerase II
RNA polymerase II
Transcription initiation complex
RNA synthesis


Enhancer
Promoter
Control elements
Albumin gene
Crystallin gene
LIVER CELL NUCLEUS
Available activators
Albumin gene expressed
Crystallin gene not expressed
LENS CELL NUCLEUS
Available activators
Albumin gene not expressed
Crystallin gene expressed
(a) Liver cell
(b) Lens cell


Enhancer
Promoter
Control elements
Albumin gene
Crystallin gene
LIVER CELL NUCLEUS
Available activators
Albumin gene expressed
Crystallin gene not expressed

(a) Liver cell


Enhancer
Promoter
Control elements
Albumin gene
Crystallin gene
LENS CELL NUCLEUS
Available activators
Albumin gene not expressed
Crystallin gene expressed
(b) Lens cell

# *Coordinately Controlled Genes in Eukaryotes*

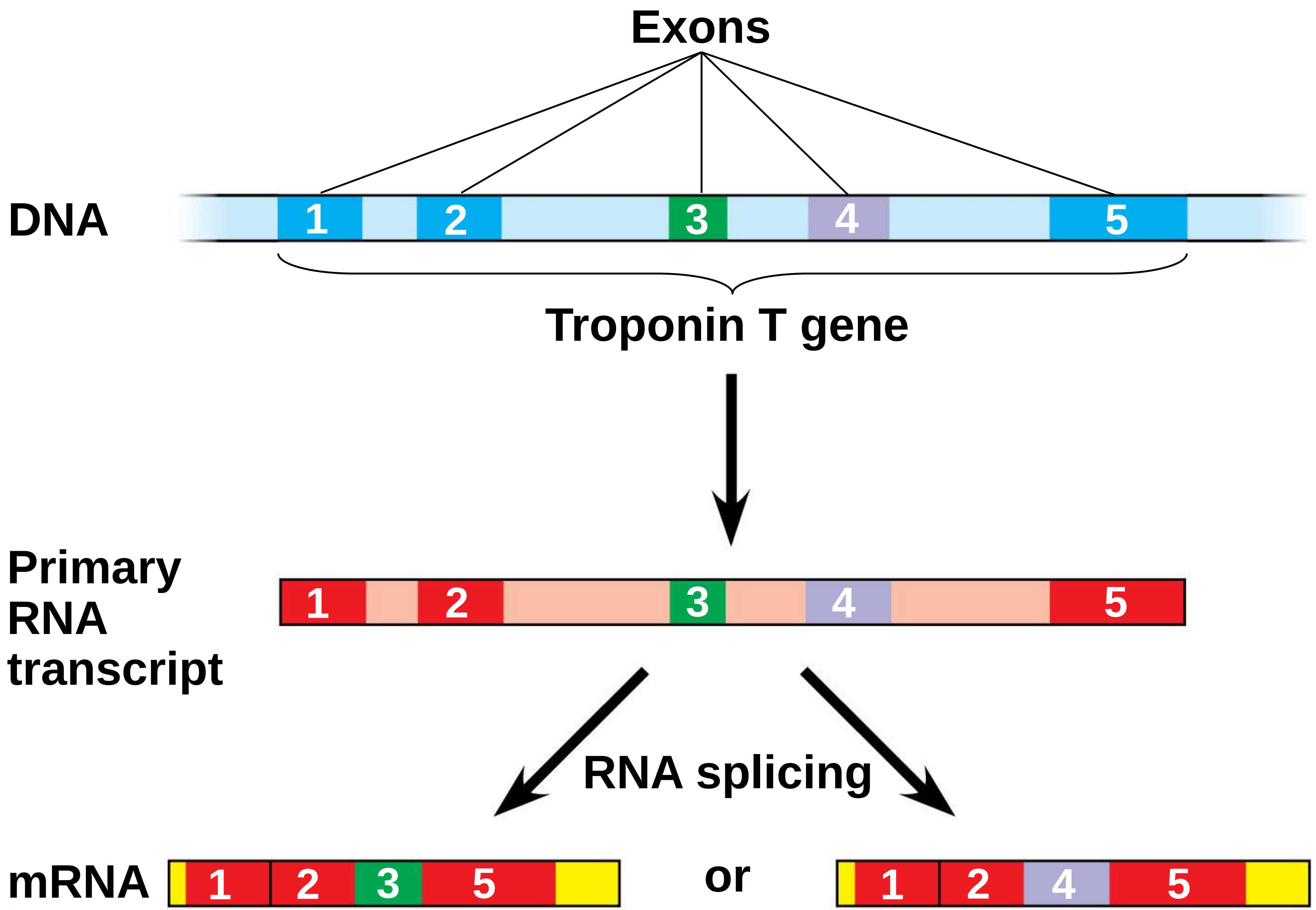
- Unlike the genes of a prokaryotic operon, each of the co-expressed eukaryotic genes has a promoter and control elements
- These genes can be scattered over different chromosomes, but each has the same combination of control elements
- Copies of the activators recognize specific control elements and promote simultaneous transcription of the genes

# Mechanisms of Post-Transcriptional Regulation

- Transcription alone does not account for gene expression
- Regulatory mechanisms can operate at various stages after transcription
- Such mechanisms allow a cell to fine-tune gene expression rapidly in response to environmental changes

# *RNA Processing/Modification*

- In **alternative RNA splicing**, different mRNA molecules are produced from the same primary transcript, depending on which RNA segments are treated as exons and which as introns


Exons
DNA
1
2
3
4
5
Troponin T gene
Primary RNA transcript
1
2
3
4
5
RNA splicing
mRNA
1
2
3
5
or
1
2
4
5

# *mRNA Degradation*

- The life span of mRNA molecules in the cytoplasm is a key to determining protein synthesis
- Eukaryotic mRNA is more long lived than prokaryotic mRNA
- Nucleotide sequences that influence the lifespan of mRNA in eukaryotes reside in the untranslated region (UTR) at the 3′ end of the molecule

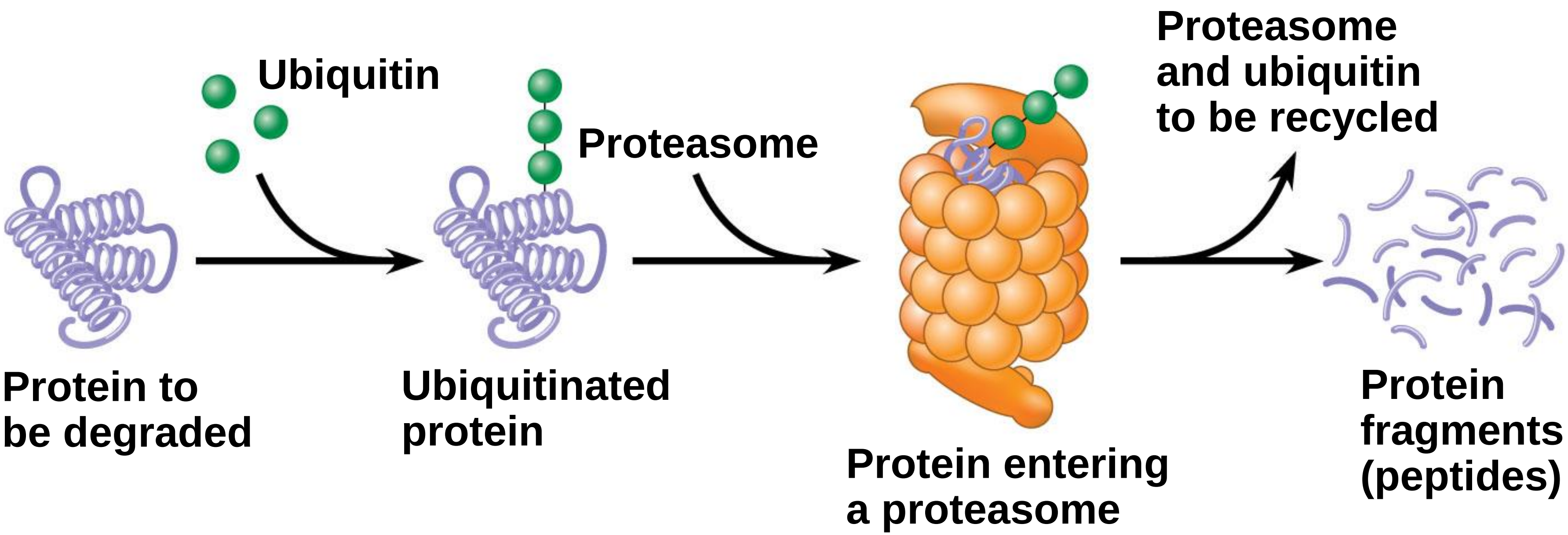
# *Initiation of Translation*

- The initiation of translation of selected mRNAs can be blocked by regulatory proteins that bind to sequences or structures of the mRNA

# *Protein Processing and Degradation*

- After translation, various types of protein processing, including cleavage and the addition of chemical groups, are subject to control (protein modification)

- **Proteasomes** are giant protein complexes that bind protein molecules and degrade them

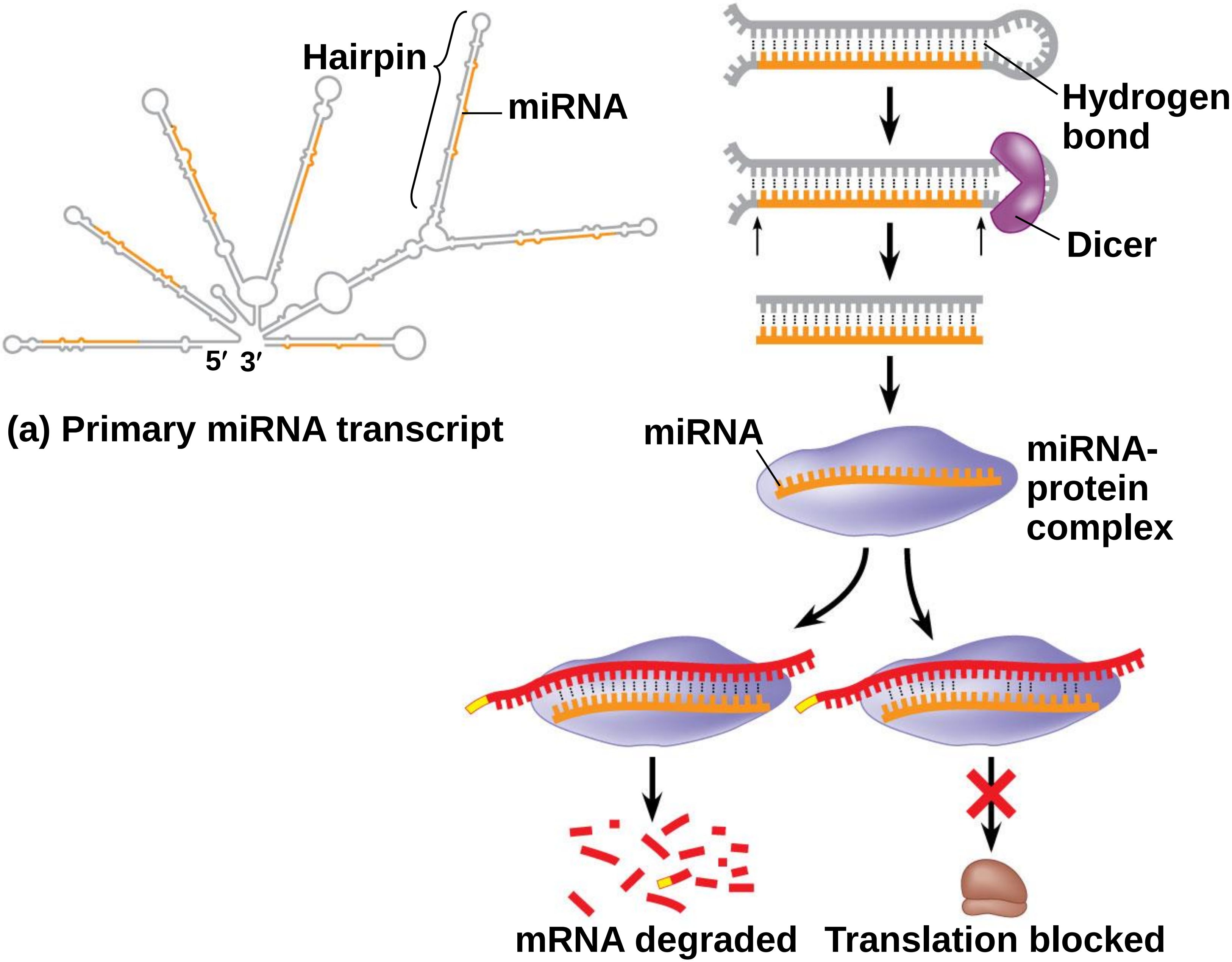

Ubiquitin
Proteasome
Proteasome
and ubiquitin
to be recycled
Protein to
be degraded
Ubiquitinated
protein
Protein entering
a proteasome
Protein
fragments
(peptides)

# Noncoding RNAs play multiple roles in controlling gene expression

- Only a small fraction of DNA codes for proteins, and a very small fraction of the non-protein-coding DNA consists of genes for RNA such as rRNA and tRNA
- A significant amount of the genome may be transcribed into noncoding RNAs (ncRNAs)
- Noncoding RNAs regulate gene expression at two points: mRNA translation and chromatin configuration

# Effects on mRNAs by MicroRNAs and Small Interfering RNAs

- **MicroRNAs** (**miRNAs**) are small single-stranded RNA molecules that can bind to mRNA
- These can degrade mRNA or block its translation

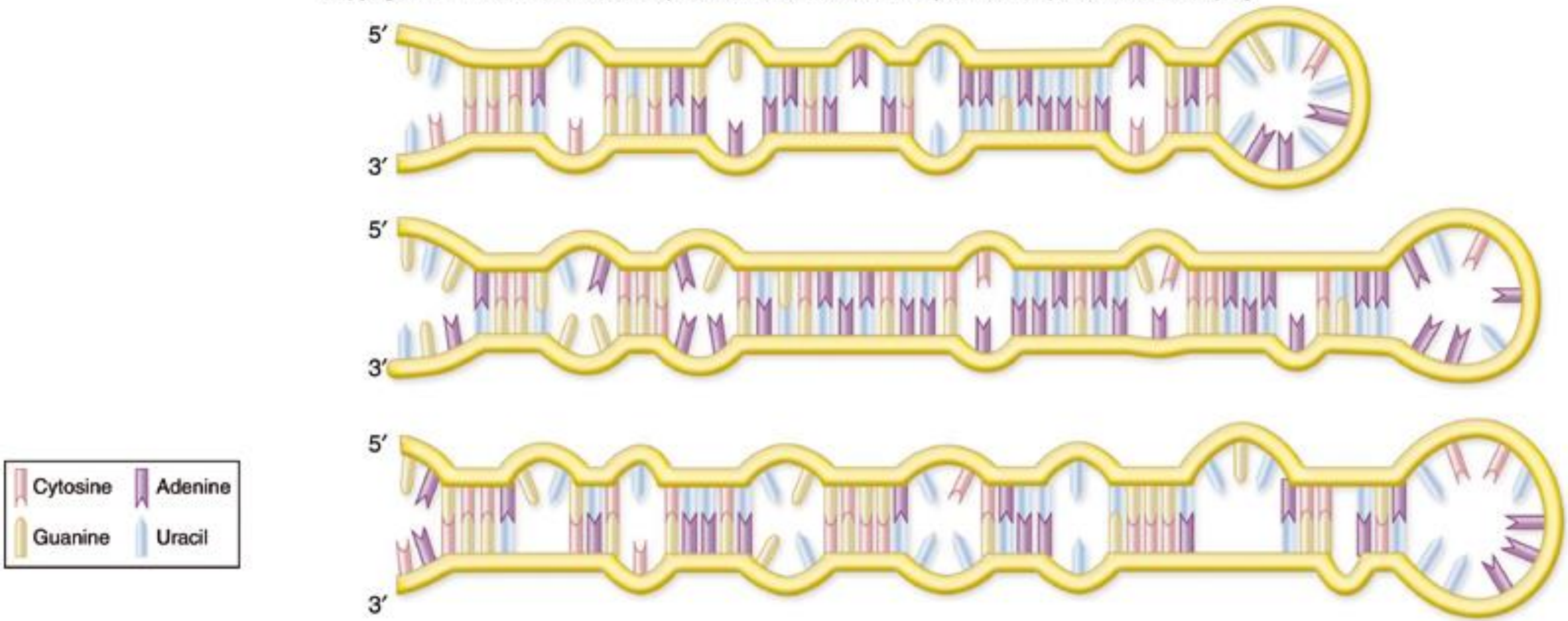



(a) Primary miRNA transcript

(b) Generation and function of miRNAs

# RNA interference

- **RNA interference** involves the use of small RNA molecules
- The enzyme **Dicer** chops double stranded RNA into small pieces of RNA
  - **micro-RNAs** (**miRNA**) bind to complementary RNA to prevent translation
  - **small interfering RNAs** (**siRNA**) degrade particular mRNAs before translation


5′
3′
5′
3′
5′
3′
Cytosine
Adenine
Guanine
Uracil

# RNA editing

- **RNA editing** creates mature mRNA that are not truly encoded by the genome.
- For example –
  - apolipoprotein B exists in 2 isoforms
  - one isoform is produced by editing the mRNA to create a stop codon
  - **RNA editing is tissue-specific**

# Disease/Syndromes associated …..

- Cancer, Autoimmunity,
- Neurological disorders, diabetes
- Cardiovascular diseases
- Obesity

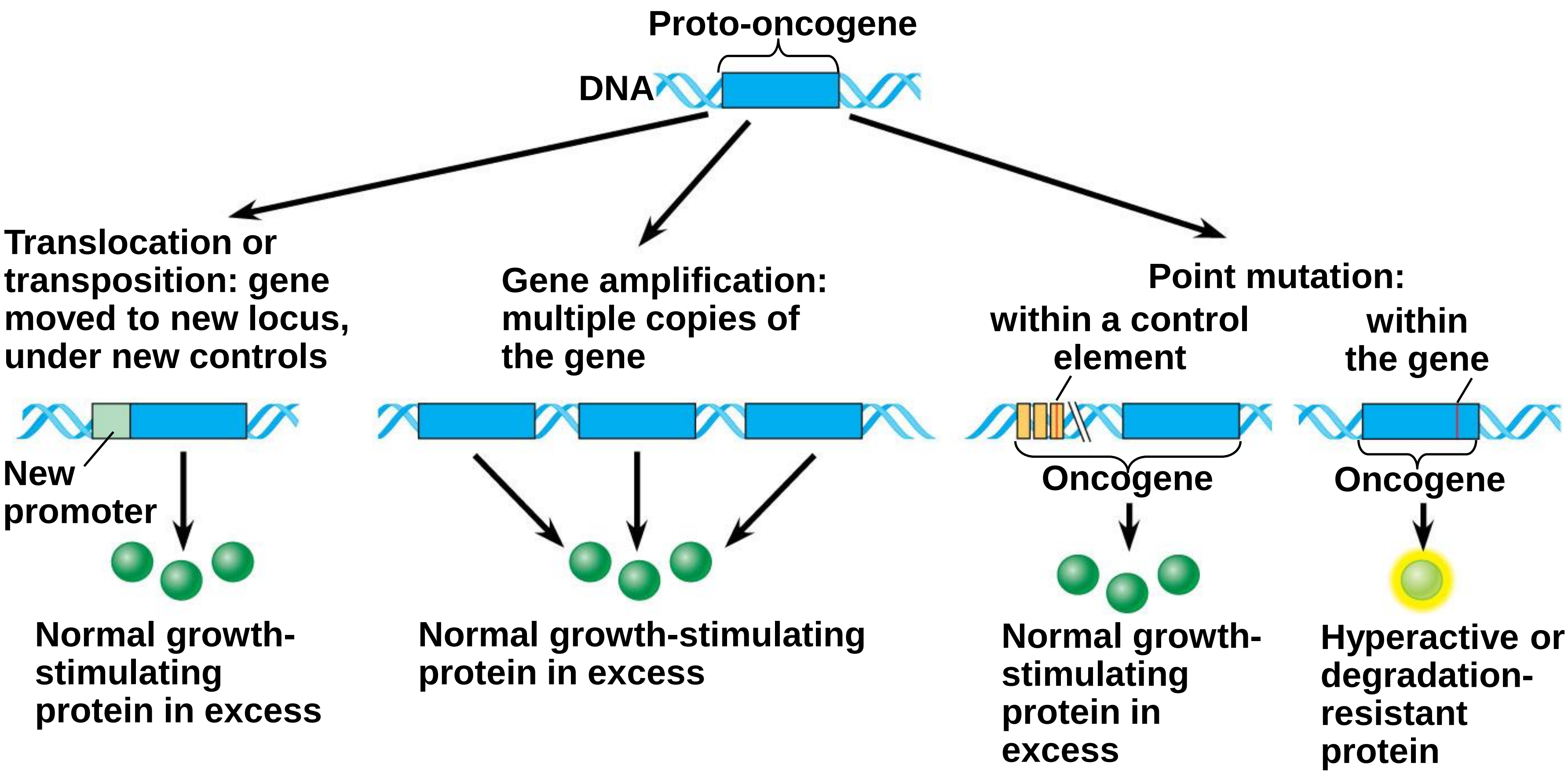
# Cancer results from genetic changes that affect cell cycle control

- The gene regulation systems that go wrong during cancer are the very same systems involved in embryonic development

# Types of Genes Associated with Cancer

- Cancer can be caused by mutations to genes that regulate cell growth and division
- Tumor viruses can cause cancer in animals including humans e.g. human papillomavirus, hepatitis B and hepatitis C virus, Epstein–Barr virus, human T-lymphotropic virus among others

- **Oncogenes** are cancer-causing genes
- **Proto-oncogenes** are the corresponding normal cellular genes that are responsible for normal cell growth and division
- Conversion of a proto-oncogene to an oncogene can lead to abnormal stimulation of the cell cycle


Proto-oncogene
DNA
Translocation or transposition: gene moved to new locus, under new controls
Gene amplification: multiple copies of the gene
Point mutation:
within a control element
within the gene
New promoter
Oncogene
Oncogene
Normal growth-stimulating protein in excess
Normal growth-stimulating protein in excess
Normal growth-stimulating protein in excess
Hyperactive or degradation-resistant protein

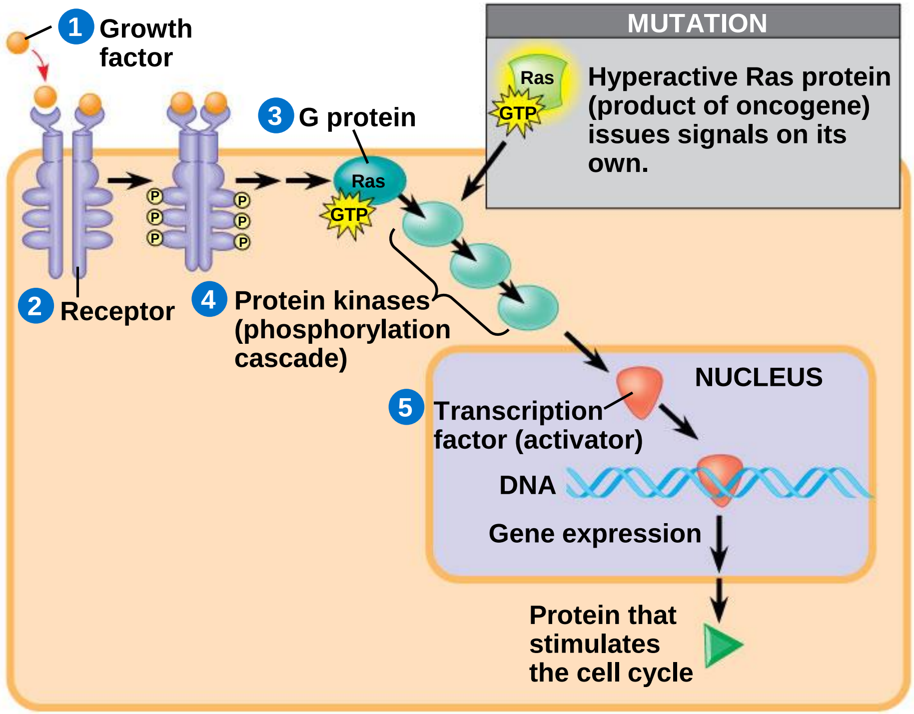
- Proto-oncogenes can be converted to oncogenes by
  - Movement of DNA within the genome: if it ends up near an active promoter, transcription may increase
  - Amplification of a proto-oncogene: increases the number of copies of the gene
  - Point mutations in the proto-oncogene or its control elements: cause an increase in gene expression

# *Tumor-Suppressor Genes*

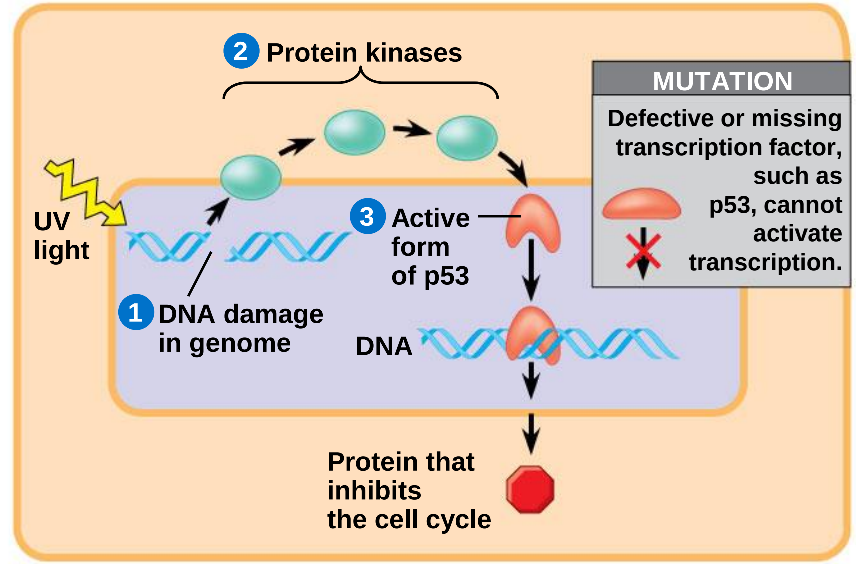
- **Tumor-suppressor genes** help prevent uncontrolled cell growth
- Mutations that decrease protein products of tumor-suppressor genes may contribute to cancer onset
- Tumor-suppressor proteins
  - Repair damaged DNA
  - Control cell adhesion
  - Inhibit the cell cycle in the cell-signaling pathway

# Interference with Normal Cell-Signaling Pathways

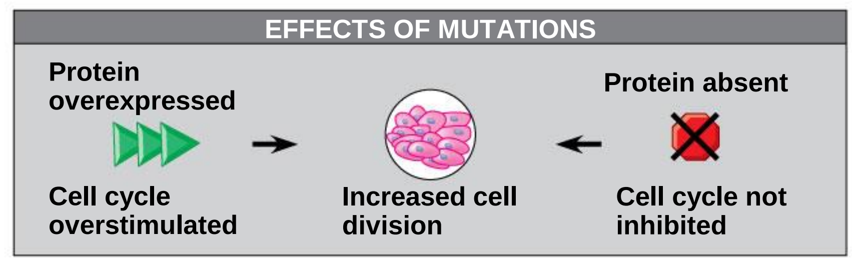
- Mutations in the *ras* proto-oncogene and *p53* tumor-suppressor gene are common in human cancers
- Mutations in the ***ras*** **gene** can lead to production of a hyperactive Ras protein and increased cell division



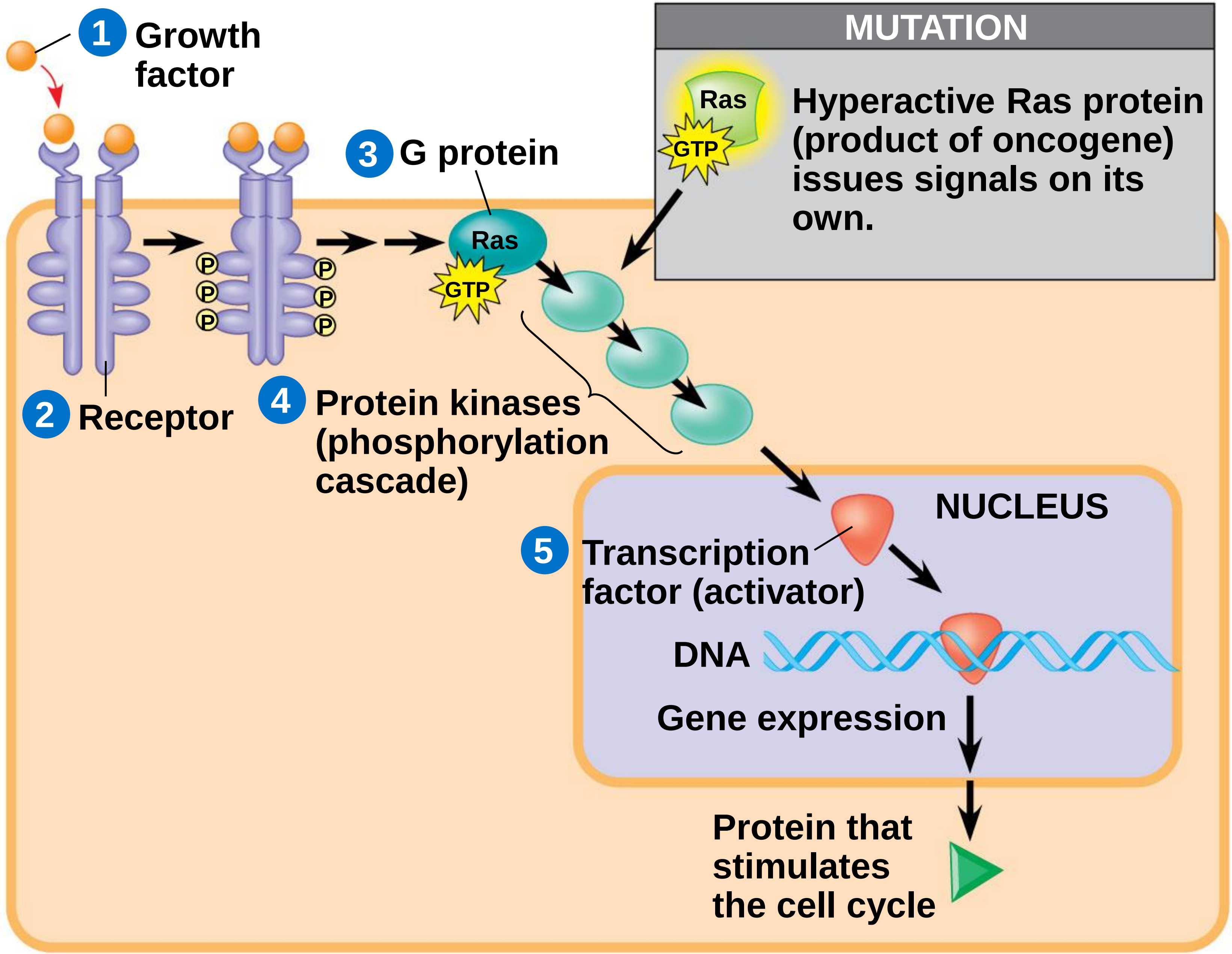

(a) Cell cycle–stimulating pathway



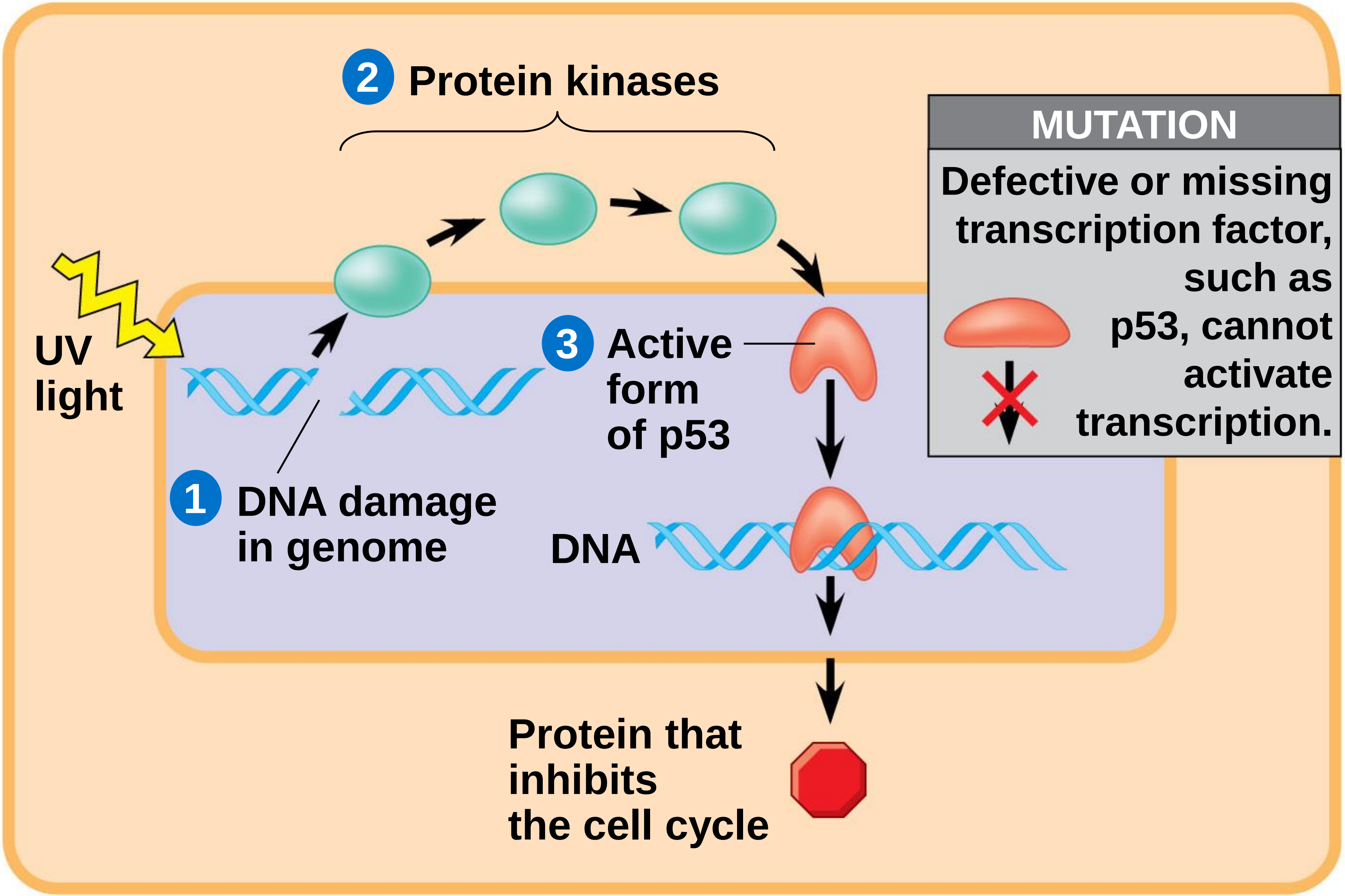

(b) Cell cycle–inhibiting pathway




(c) Effects of mutations

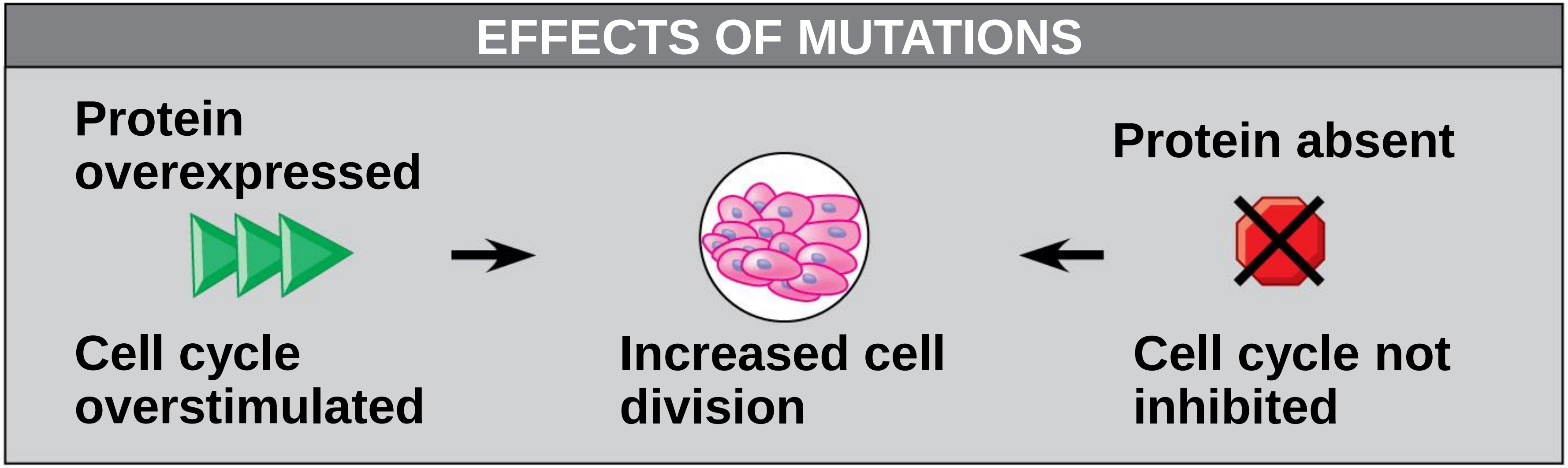



(a) Cell cycle–stimulating pathway




(b) Cell cycle–inhibiting pathway

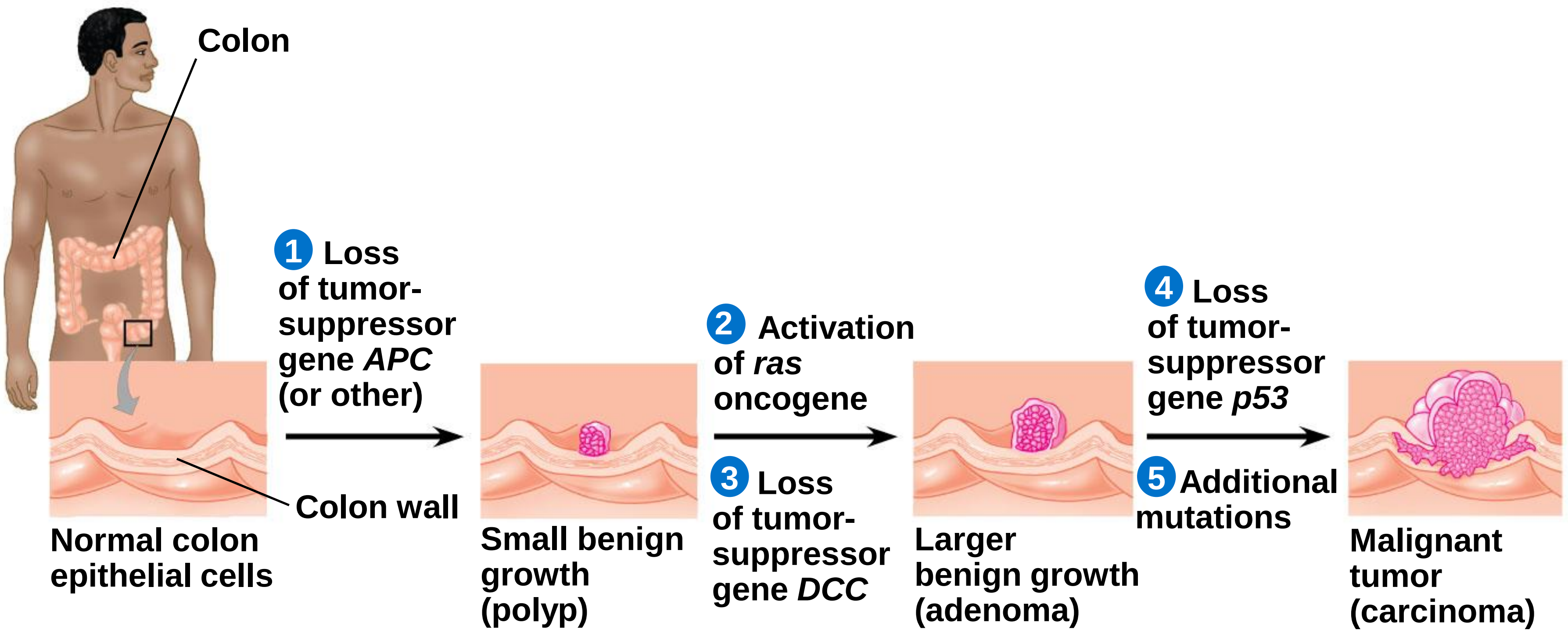
- Suppression of the cell cycle can be important in the case of damage to a cell’s DNA; *p53* prevents a cell from passing on mutations due to DNA damage
- Mutations in the ***p53* gene** prevent suppression of the cell cycle

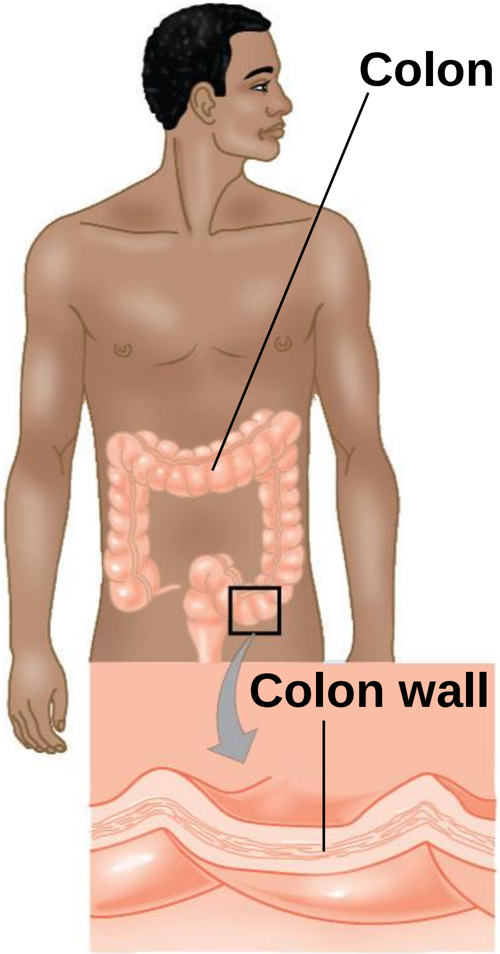



(c) Effects of mutations

# The Multistep Model of Cancer Development

- Multiple mutations are generally needed for full-fledged cancer; thus the incidence increases with age
- At the DNA level, a cancerous cell is usually characterized by at least one active oncogene and the mutation of several tumor-suppressor genes


Colon
1 Loss of tumor-suppressor gene *APC* (or other)
Colon wall
Normal colon epithelial cells
Small benign growth (polyp)
2 Activation of *ras* oncogene
3 Loss of tumor-suppressor gene *DCC*
Larger benign growth (adenoma)
4 Loss of tumor-suppressor gene *p53*
5 Additional mutations
Malignant tumor (carcinoma)



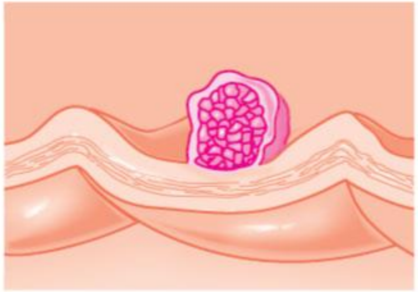

**Normal colon epithelial cells**

1 **Loss of tumor-suppressor gene *APC* (or other)**

**Small benign growth (polyp)**

2 Activation of *ras* oncogene

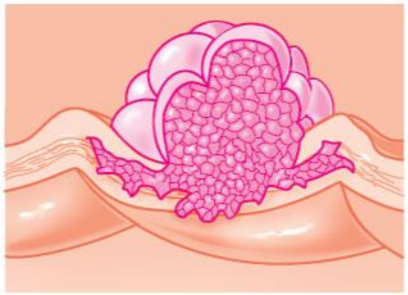
3 Loss of tumor-suppressor gene *DCC*



Larger benign growth (adenoma)

**4** **Loss of tumor-suppressor gene *p53***

**5** **Additional mutations**

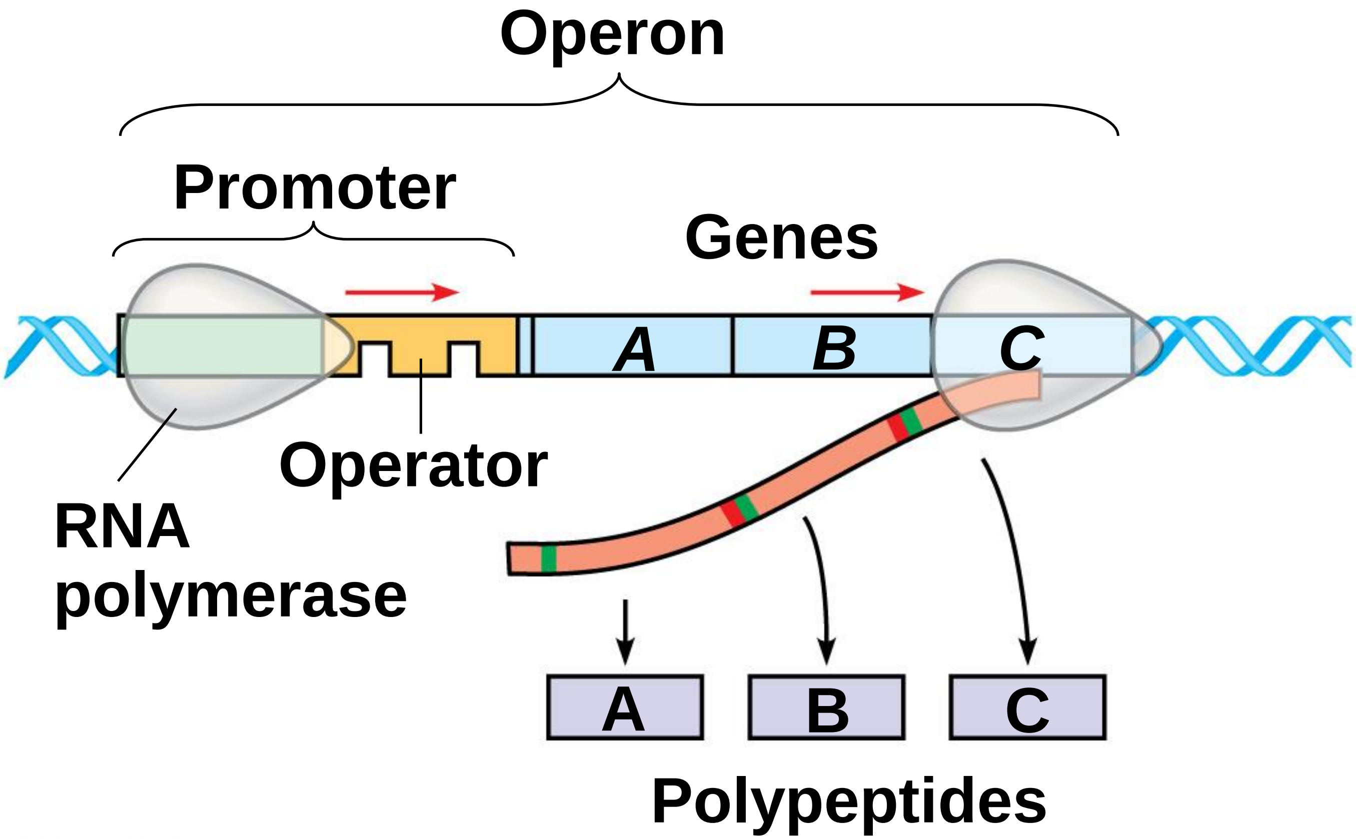


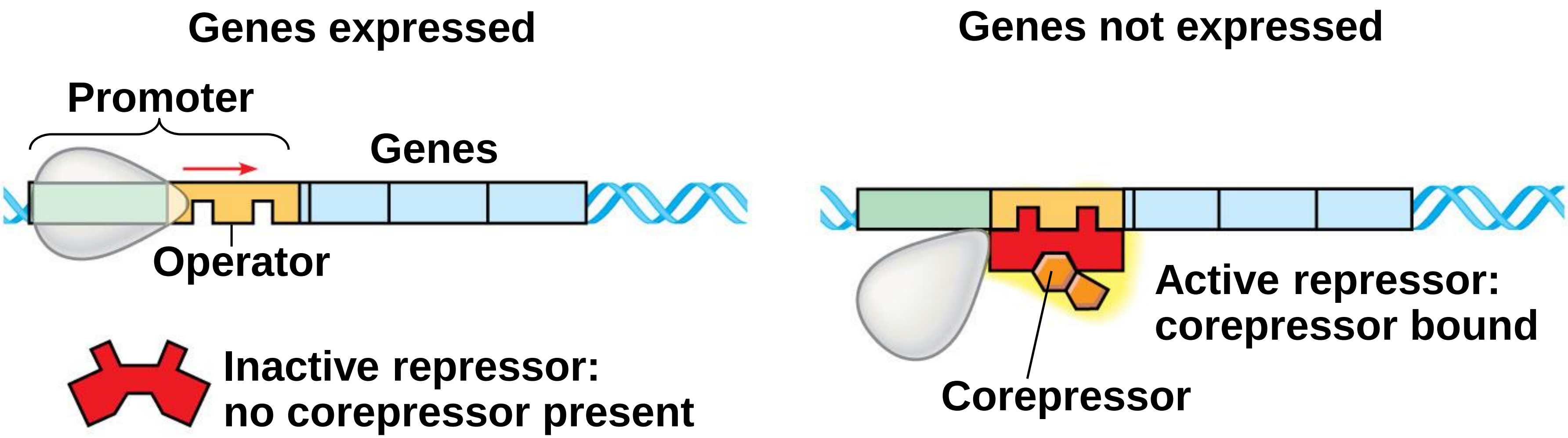
**Malignant tumor (carcinoma)**

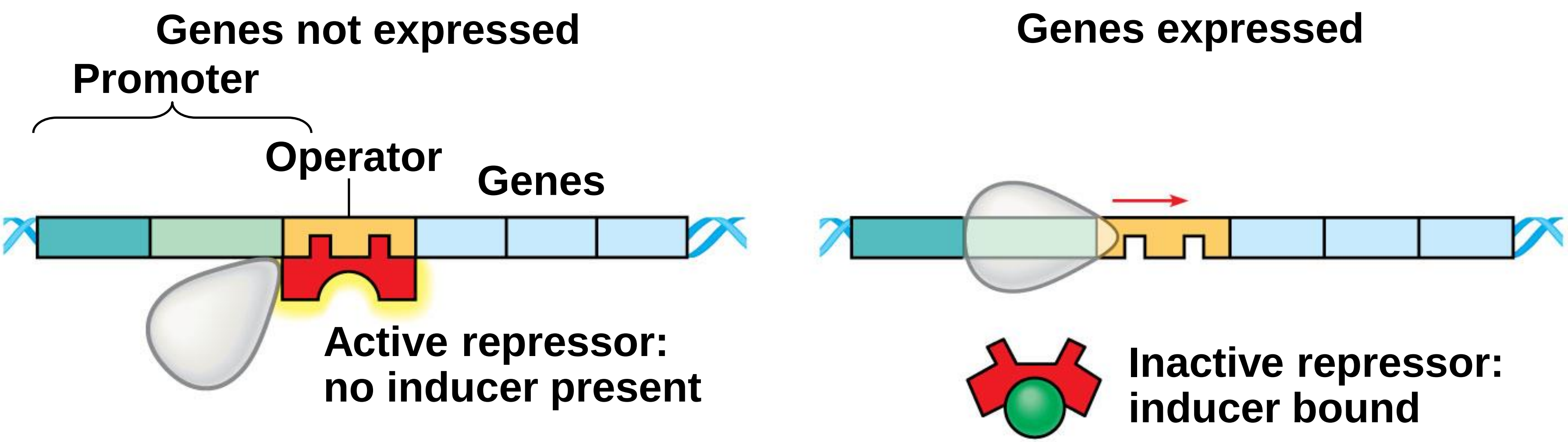
# Inherited Predisposition and Other Factors Contributing to Cancer

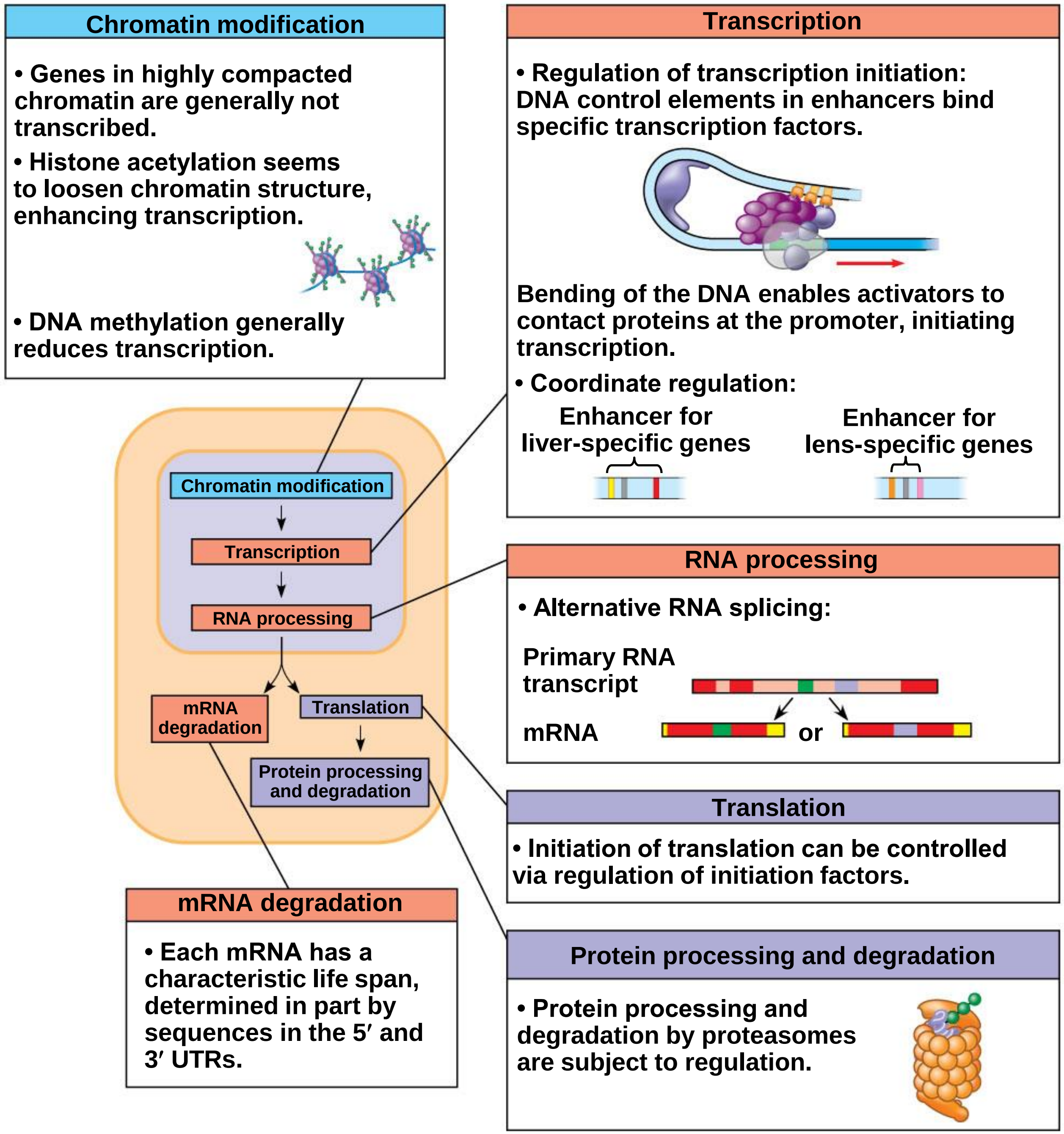
- Individuals can inherit oncogenes or mutant alleles of tumor-suppressor genes
- Inherited mutations in the tumor-suppressor gene *adenomatous polyposis coli* are common in individuals with colorectal cancer
- Mutations in the *BRCA1* or *BRCA2* gene are found in at least half of inherited breast cancers, and tests using DNA sequencing can detect these mutations

# Summary of gene regulation

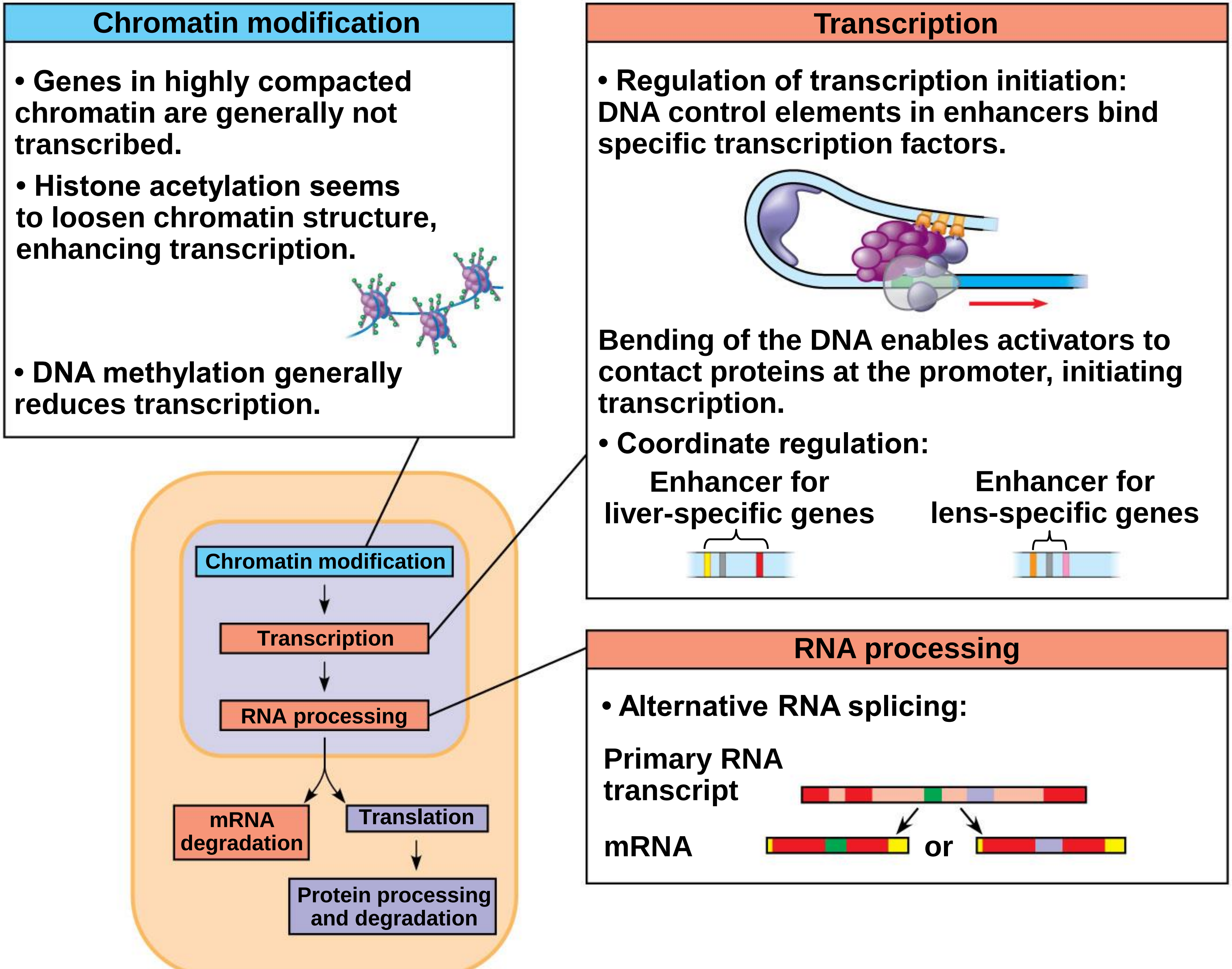


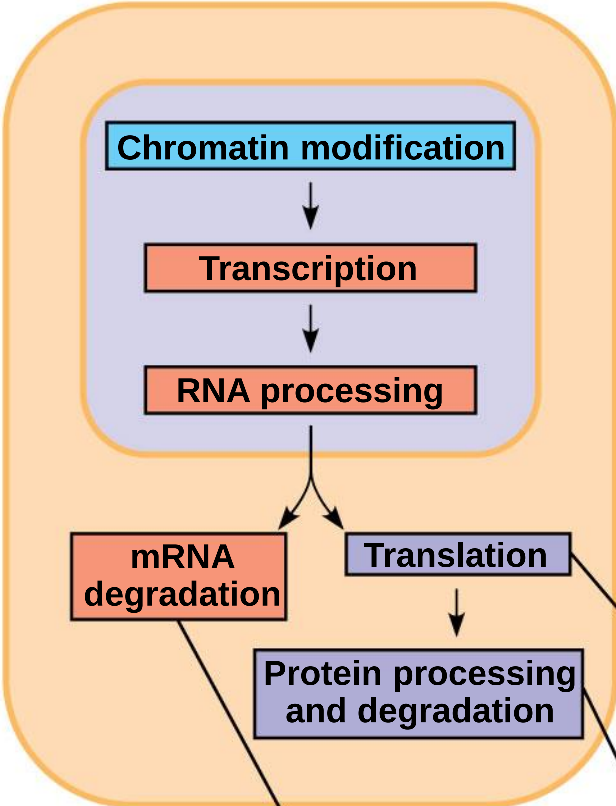

Genes expressed
Promoter
Genes
Operator
Inactive repressor:
no corepressor present
Genes not expressed
Active repressor:
corepressor bound
Corepressor


Genes not expressed
Promoter
Operator
Genes
Active repressor:
no inducer present
Genes expressed
Inactive repressor:
inducer bound



Chromatin modification
• Genes in highly compacted chromatin are generally not transcribed.
• Histone acetylation seems to loosen chromatin structure, enhancing transcription.
• DNA methylation generally reduces transcription.
Transcription
• Regulation of transcription initiation: DNA control elements in enhancers bind specific transcription factors.
Bending of the DNA enables activators to contact proteins at the promoter, initiating transcription.
• Coordinate regulation:
Enhancer for liver-specific genes
Enhancer for lens-specific genes
Chromatin modification
Transcription
RNA processing
mRNA degradation
Translation
Protein processing and degradation
RNA processing
• Alternative RNA splicing:
Primary RNA transcript
mRNA
or
Translation
• Initiation of translation can be controlled via regulation of initiation factors.
mRNA degradation
• Each mRNA has a characteristic life span, determined in part by sequences in the 5′ and 3′ UTRs.
Protein processing and degradation
• Protein processing and degradation by proteasomes are subject to regulation.


Chromatin modification
• Genes in highly compacted chromatin are generally not transcribed.
• Histone acetylation seems to loosen chromatin structure, enhancing transcription.
• DNA methylation generally reduces transcription.
Transcription
• Regulation of transcription initiation: DNA control elements in enhancers bind specific transcription factors.
Bending of the DNA enables activators to contact proteins at the promoter, initiating transcription.
• Coordinate regulation:
Enhancer for liver-specific genes
Enhancer for lens-specific genes
Chromatin modification
Transcription
RNA processing
mRNA degradation
Translation
Protein processing and degradation
RNA processing
• Alternative RNA splicing:
Primary RNA transcript
mRNA
or




## Translation

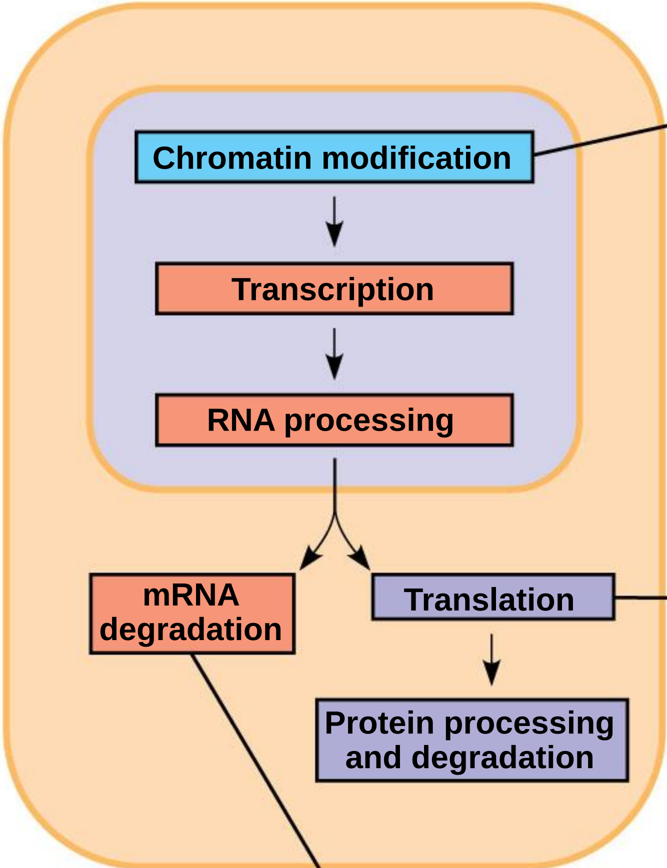
- Initiation of translation can be controlled via regulation of initiation factors.

## mRNA degradation

- Each mRNA has a characteristic life span, determined in part by sequences in the 5′ and 3′ UTRs.

## Protein processing and degradation

- Protein processing and degradation by proteasomes are subject to regulation.

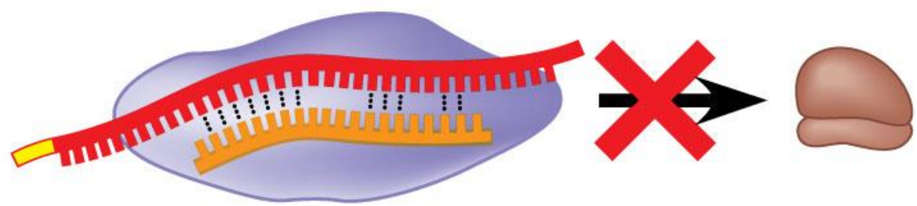



## Chromatin modification

- Small or large noncoding RNAs can promote the formation of heterochromatin in certain regions, blocking transcription.

## Translation

- miRNA or siRNA can block the translation of specific mRNAs.



## mRNA degradation

- miRNA or siRNA can target specific mRNAs for destruction.

# Review Questions

- Differentiate between gene regulation strategies in prokaryotes and eukaryotes
- Gene expression can be regulated at the levels of transcription and translation. Discuss each with respect to energy efficiency of each mechanism.
- What is structural gene regulation regulation and how does it usually work?
- What is posttranscriptional regulation and how does it usually work?