Gene Expression Regulation

Central dogma : collection of reactions that take place for expressing a particular gene (extract a sequence of amino acids from DNA, express a particular gene as a protein that consists of specified amino acids)

Regulation of 2 processes,
1) transcription which is producing RNA from template DNA from one or more particular gene
2)translation which is releasing mRNA from the nucleus aftervbeing processed (capping, poly A tail, and splicing) and convert it to amino acid sequence

*gene expression regulation occur through regulatory proteins.

- Controlling gene expression is often accomplished by controlling transcription initiation. (Transcription factors , inhancers and silencers)
- **Regulatory proteins** bind to DNA to either (block or stimulate). transcription, depending on how they interact with RNA polymerase.
- Prokaryotic organisms regulate gene expression in response to their environment. Ex some nutrients like lactose will increase the rate of gene expression which will catabolize this nutrients, and when these catabolixing enzymes are not needed , the gene expression rate decreases or it's blocked
- Eukaryotic cells regulate gene expression to maintain **homeostasis** in the organism. In order to maintain a stable enternal environment, so the it doesn't exceed the required transcription rate or decrease less than the required transcription rate

Decrease or complete absence of transcription 》 decrease translation rate or complete no production of protein

Increase transcription rate 》 increase protein synthesis rate

Regulatory Proteins

- The regulatory proteins are binding to specific DNA sequences to regulate gene expression. Interact with RNA plymerase either increase or decrease it's activity
- They gain access to the bases of DNA at the **major groove** by possessing **DNA-binding motifs**
- DNA-binding motifs are **regions of regulatory proteins which bind to DNA**

We have 4 different interactions between regulatory proteins and DNA

1- Helix-turn-helix motif
2- Homeodomain motif
3- Zinc finger motif
4- Leucine zipper motif

this interactions control gene expression.

Regulation of gene expression in prokaryotes

activator → ⊕
repressor → ⊖

- Control of transcription initiation can be:
- Positive control – increases transcription when activators bind to DNA
- Negative control – reduces transcription when repressors bind to DNA regulatory regions (operators).

→ first region of promotor in each gene.

- Prokaryotic cells often respond to their environment by changes in gene expression.
- Genes involved in the same metabolic pathway are organized in **operons**.

→ polycistronic genes → collection of gene that participate in the same metabolic pathway (related to one reagent.)

- Some operons are induced when the metabolic pathway is needed.
- Some operons are repressed when the metabolic pathway is no longer needed.

polycistronic gene : production of a single mRNA from multiple genes which will be translated as one polypeptide in the ribosomes which will be then cleaved by specific proteases so that each type of protein is seperated from the other

These genes are usually related to onr reagent, ex if three enzymes are responsible for metabolic pathway, each one encoded for by a polycistronic gene to produce one mRNA for all

- The ***lac* operon** contains genes for the use of **lactose** as an energy source. (it includes genes encoding for three enzymes)

in the absence of glucose (main source of energy, simple and more easy to catabolize than lactose)

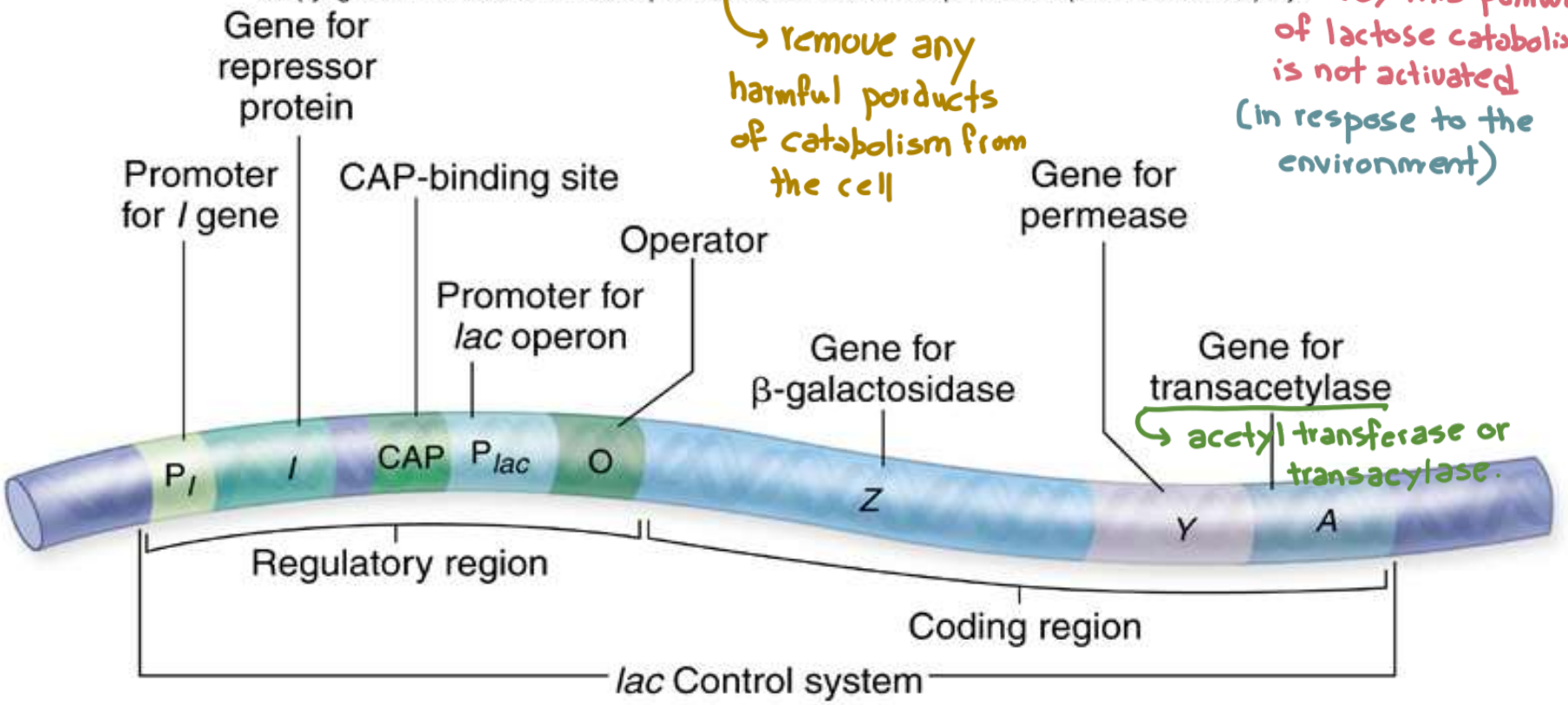
1- lac Z: encodes β-galactosidase which cleaves lactose into glucose and galactose

2- lac Y: encodes lactose permease to transport lactose into the cell from the surrounding environment to the cell

3- lac A: encodes galactoside O- acetyltransferase which plays a role in cell detoxification

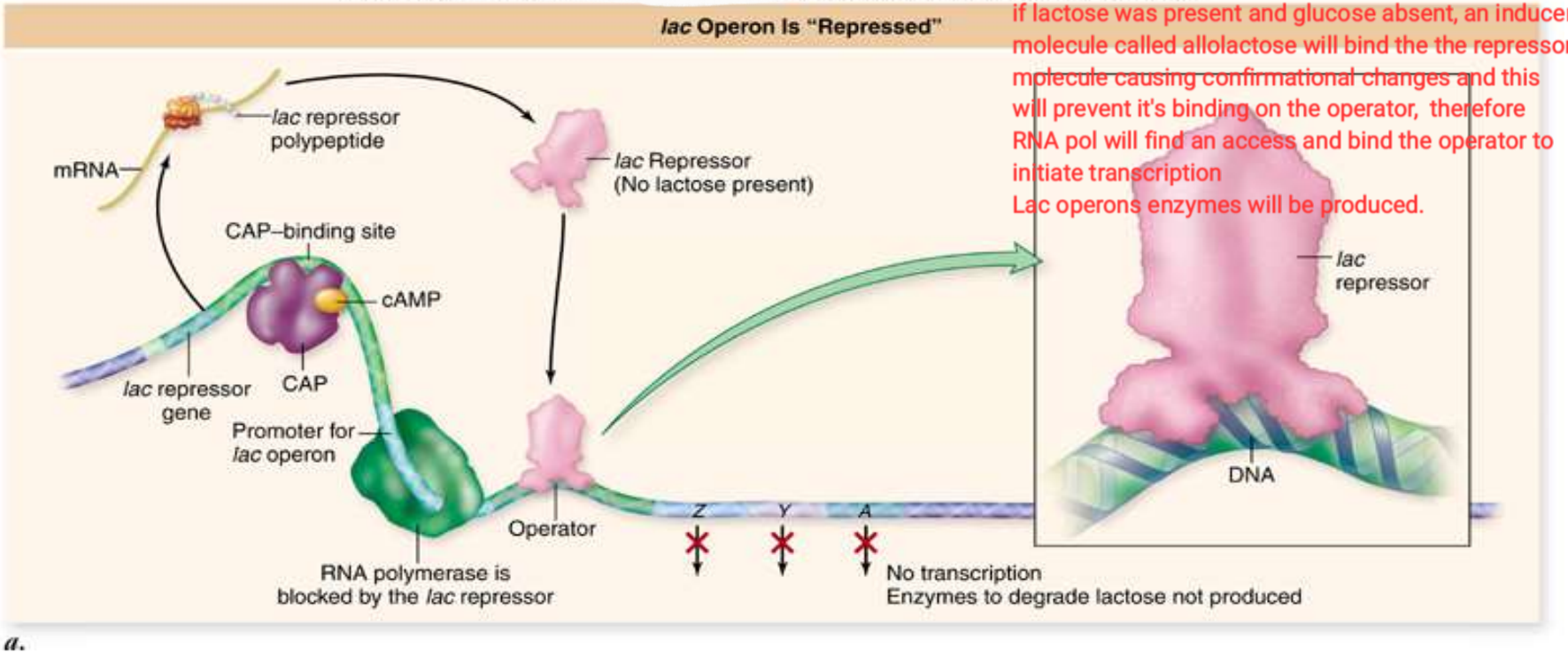
remove any harmful products of catabolism from the cell

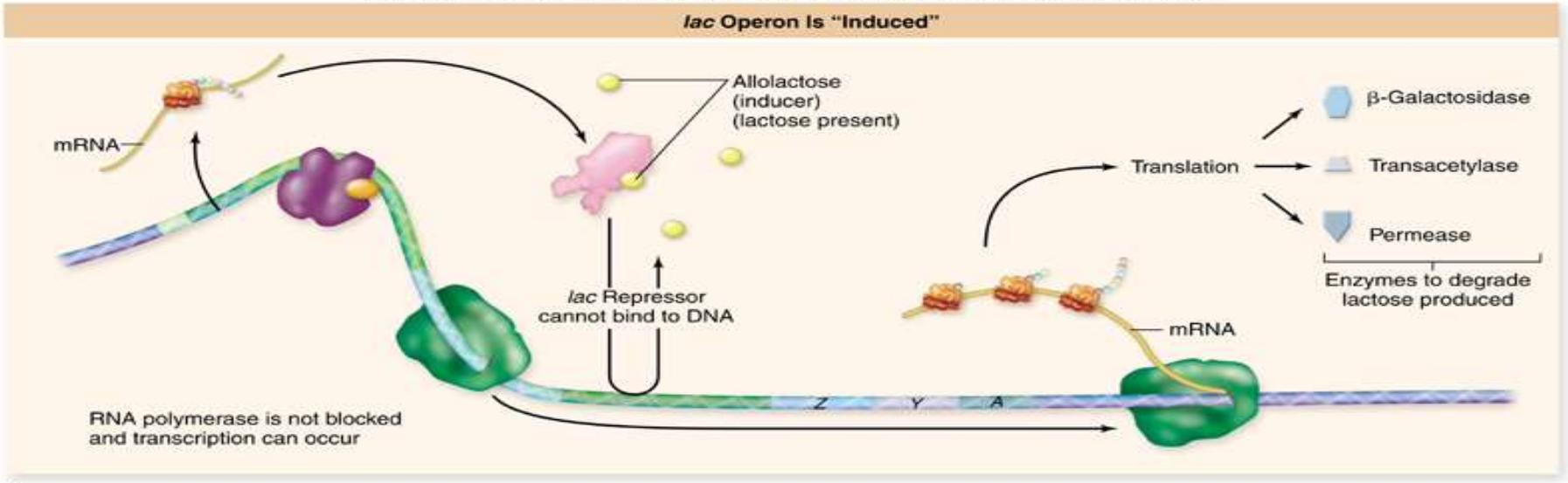
if glucose was found in the surrounding area, this pathway of lactose catabolism is not activated (in respose to the environment)



- Regulatory regions of the operon include the CAP (Catabolite activator protein) binding site, promoter, and the operator.
- The ***lac*** operon is **negatively regulated** by a repressor protein:
  - ***lac*** repressor binds to the operator to block transcription
  - In the presence of lactose, an inducer molecule binds to the repressor protein
  - Repressor can no longer bind to operator
  - Transcription proceeds

Lac operon is negative (inactive) when glucose if present, gene I is responsible for producing repressor molecules, this gene is activated when glucose is present to repress the action of lac operon by binding to the operator (site for RNA pol to bind and initiate transcription) repressor prevents RNA pol binding and there will be no transcription of lac operon enzymes , if lactose was present and glucose absent, an inducer molecule called allolactose will bind the the repressor molecule causing confirmational changes and this will prevent it's binding on the operator, therefore RNA pol will find an access and bind the operator to initiate transcription
Lac operons enzymes will be produced.





b.

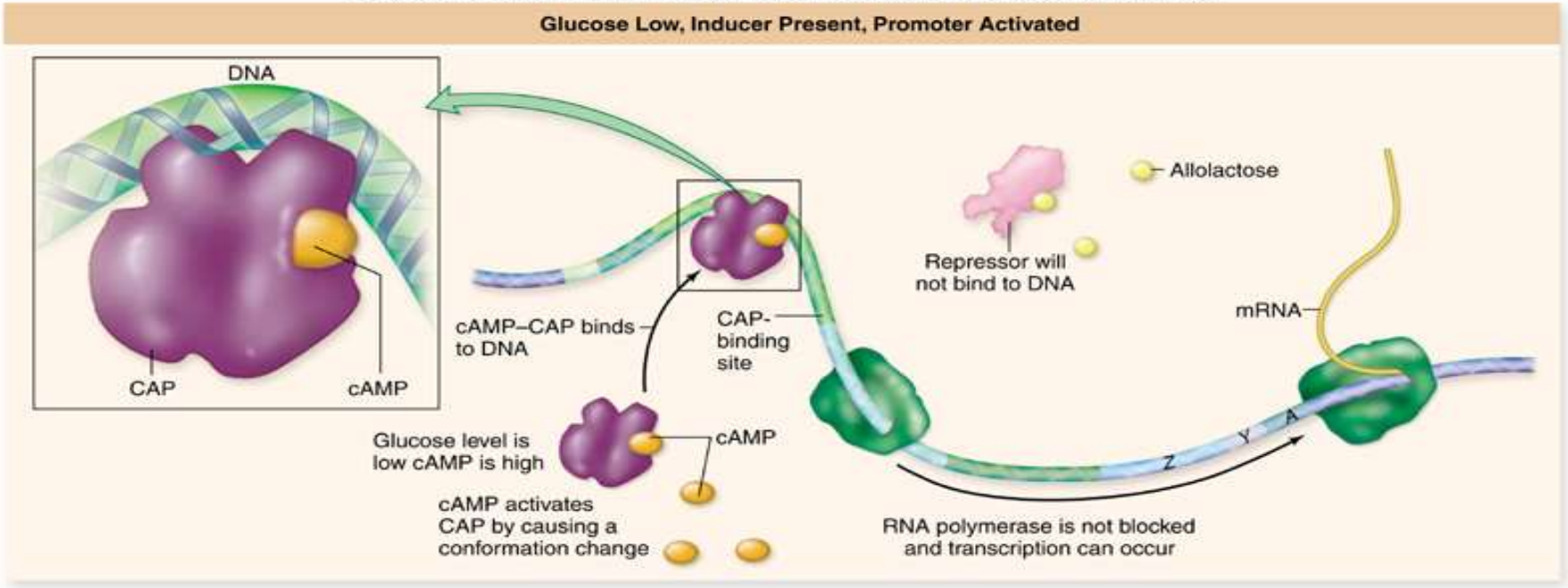
- In the presence of both **glucose** and **lactose**, bacterial cells prefer to use **glucose**.
- Glucose prevents induction of the ***lac* operon**.

  binding of **CAP – cAMP complex** to the CAP binding site is required for induction of the ***lac* operon**

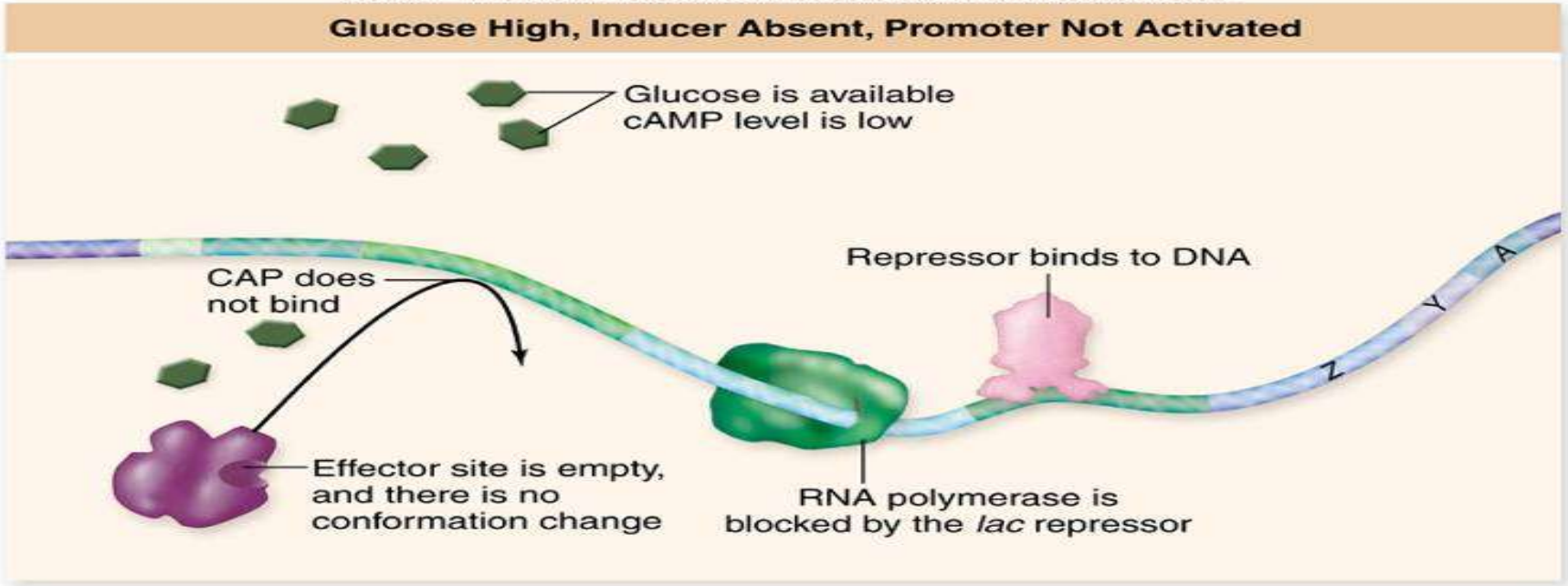
  high glucose levels cause low **cAMP** levels

  high glucose → low **cAMP** → **no induction**

Summary: ***Lac* operon** is active only in time, when the activator **CAP-cAMP** complex is attached to promotor (**no glucose**) and when is not present repressor on operator (**lactose present**).



a.



b.

- The ***trp* operon** encodes genes for the **biosynthesis of tryptophan**.

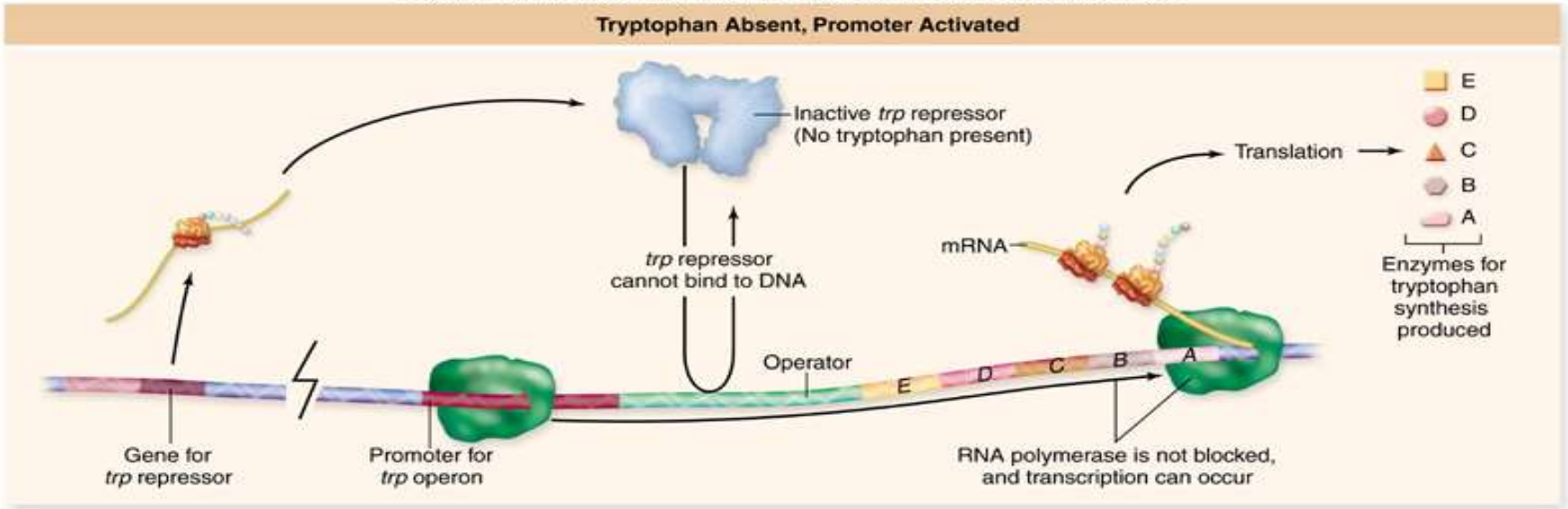
↳ essential amino acid

- The operon is not expressed when the cell contains sufficient amounts of tryptophan.

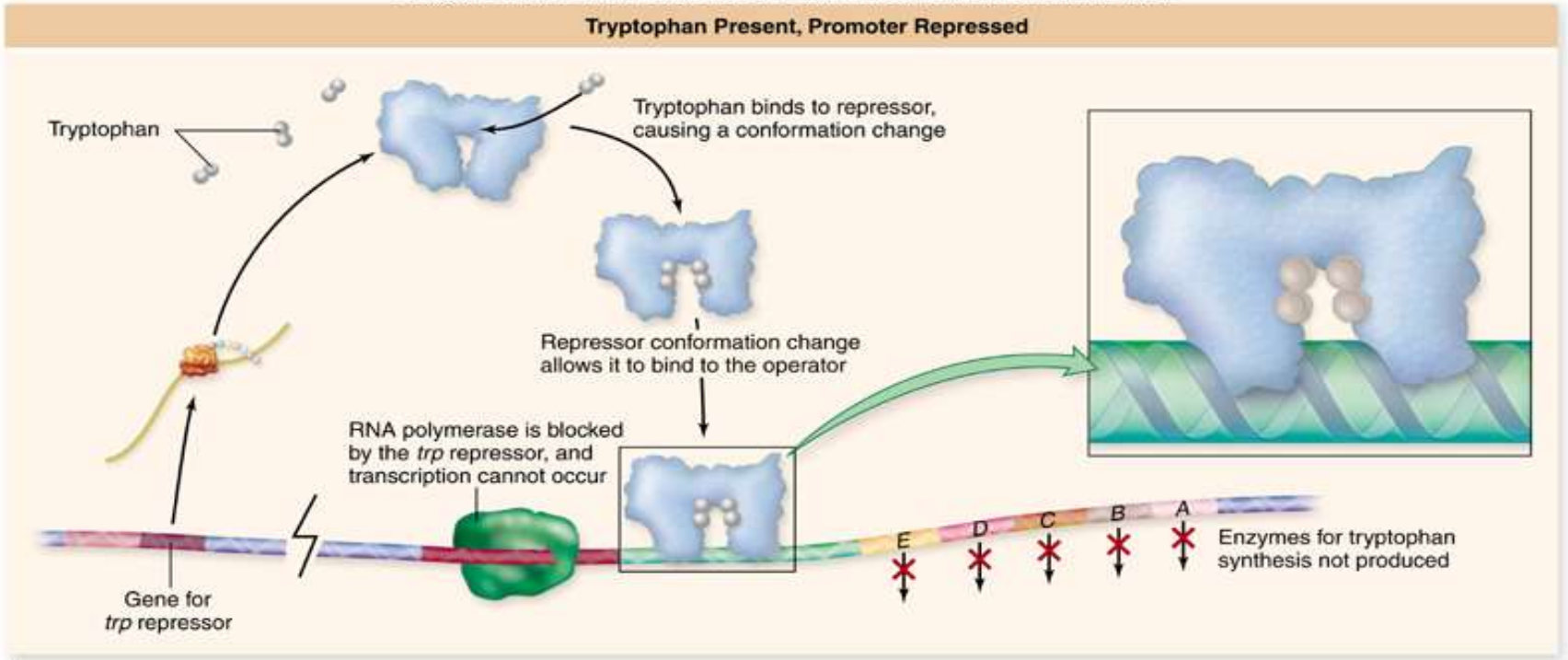
- The operon is expressed when levels of tryptophan are low.

- The ***trp* operon** is **negatively** regulated by the ***trp* repressor protein**
  - ***trp* repressor** binds to the operator to block transcription
  - Binding of repressor to the operator requires a **corepressor** which is tryptophan
  - Low levels of tryptophan prevent the repressor from binding to the operator

Repressor molecules are produced from the gene encoding for tryptophan repressor protein, this repressor will bind to the operator and this will prevent RNA pol binding to the operator 》 decrease or completely block the transcription of tryptophan operon, for this repressor to work it needs something called corepressor (tryptophan), since high levels of tryptophan inhibit the production of tryptophan operon



a.



b.

6 different levels of gene expression regulation in eukaryotes.

Regulation of gene expression in eukaryotes

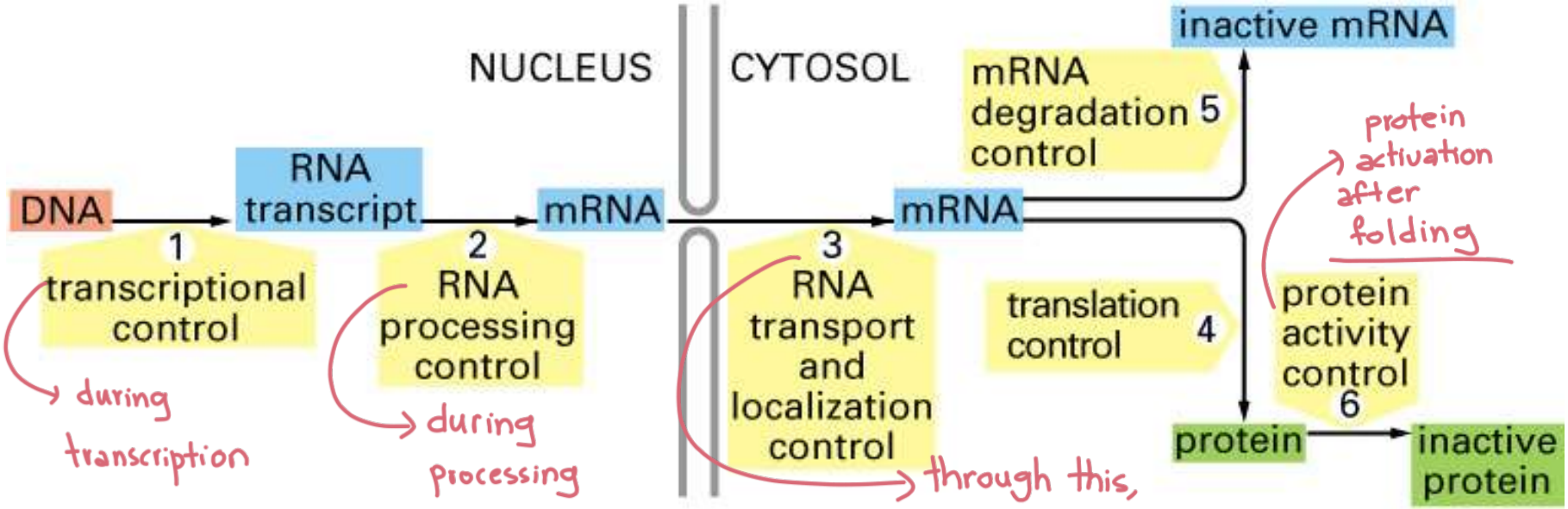


Figure 7–5. Molecular Biology of the Cell, 4th Edition.

*protein activation occur in to steps:

1- precursor protein → mature protein

2- mature protein —folding (3D structure)→ biologically active protein

- Controlling the expression of eukaryotic genes requires **transcription factors.** (2 types)
  - **General transcription factors** are required for transcription initiation (for proper binding of RNA polymerase to the DNA).
  - **Specific transcription factors** increase transcription in certain cells or in response to signals

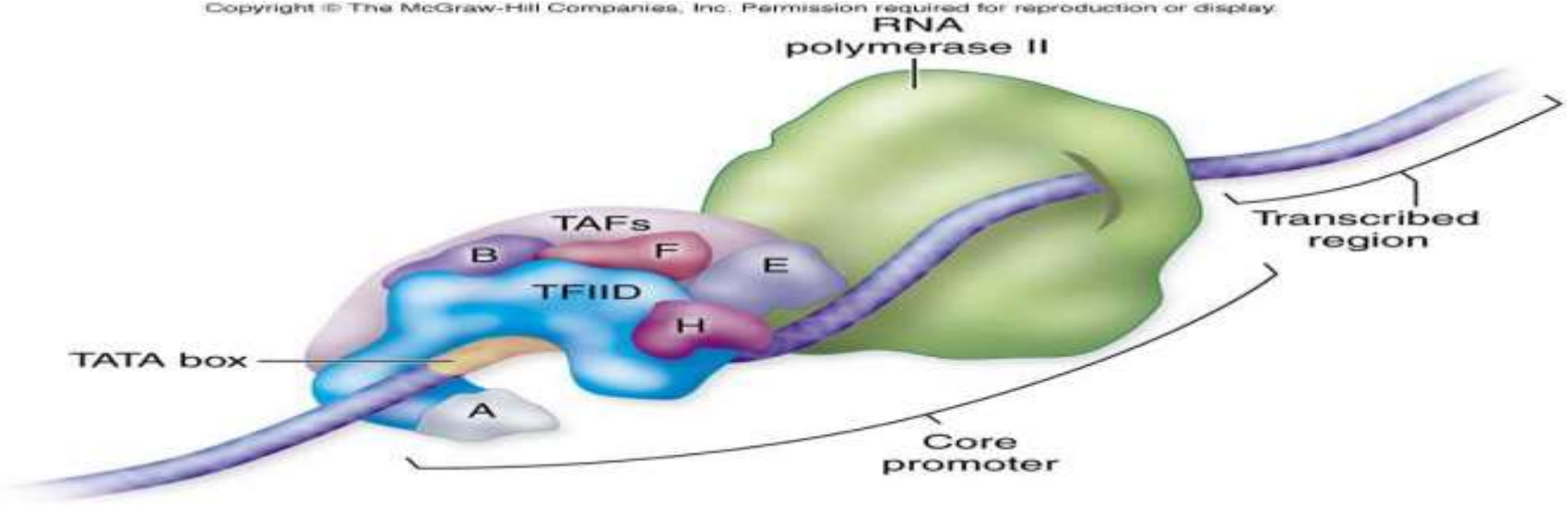
What transcription factors can recognize promoter element?
- TFIID- TBP: recognizes TATA box.
- TFIID- TAF: recognizes non- TATA box elements; CAAT and GC boxes.

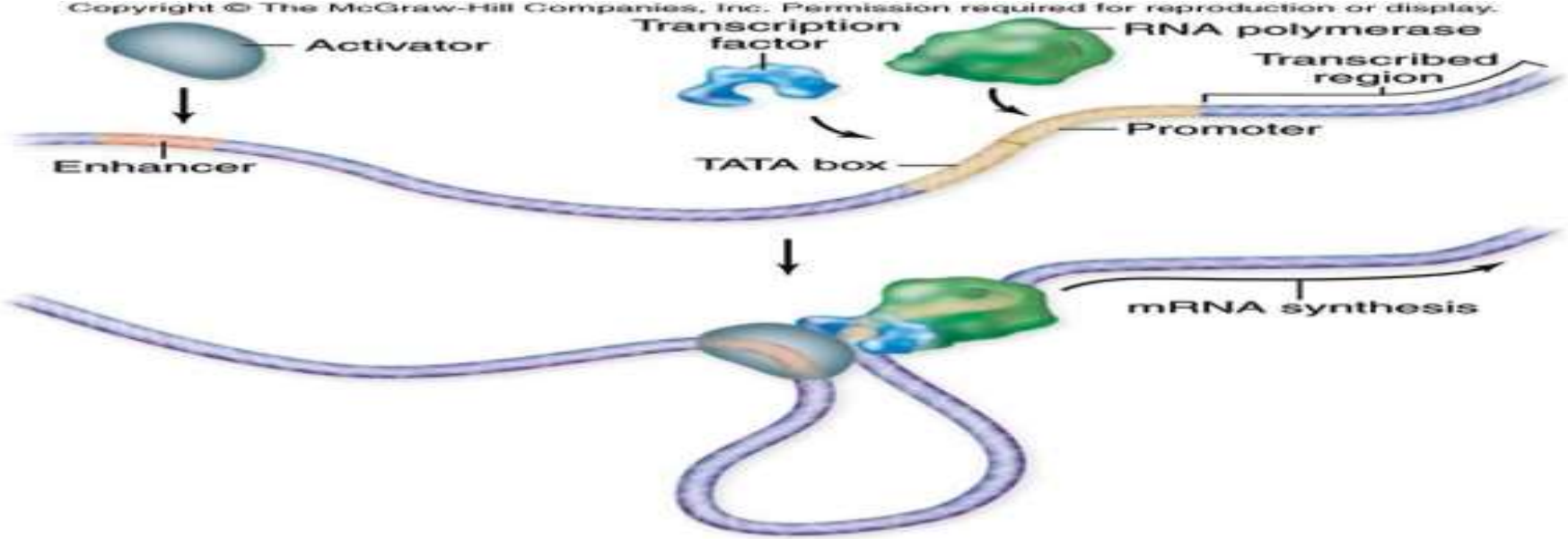
- General transcription factors bind to the promoter region of the gene.
- RNA polymerase II then binds to the promoter to begin transcription at the start site (+1).

- Enhancers are DNA sequences to which specific transcription factors (activators) bind to increase the rate of transcription.

There are two types of transcription factors:
- General transcription factors are required for transcription initiation (for proper binding of RNA polymerase to the DNA). They are 7 types we have talked about them in transcription lecture (TFIID, TFIIB... etc.). Their function: recognition of promoter elements, recruitment of other factors, recruitment of RNA polymerase II, clearance and opening of promoter with helicase activity
- Specific transcription factors increase transcription in certain cells or in response to signals
which tells the cell: produce more of specific gene
Regulation of gene expression

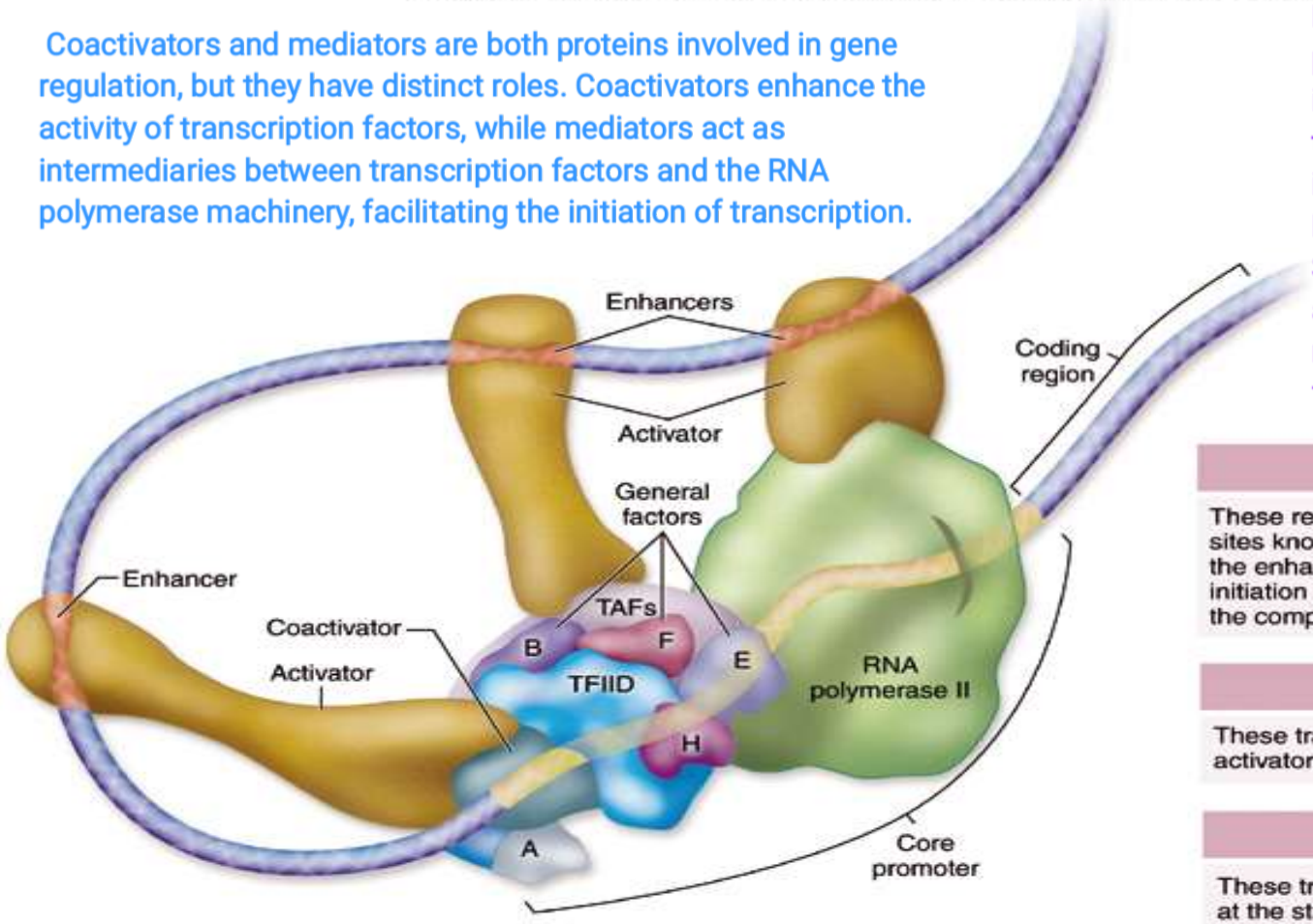
instead of enhancers/ silencers we will use activators/ repressors (which control rate of transcription). They are non-promoter DNA sequence, as a result of binding these proteins to RNA polymerase they can increase the rate of transcription or decrease it, so some of them act as enhancers (activators) and others as silencers (repressors).





- **Coactivators** and **mediators** are also required for the function of transcription factors.
  - Coactivators and mediators bind to transcription factors and bind to other parts of the transcription apparatus

Coactivators and mediators are both proteins involved in gene regulation, but they have distinct roles. Coactivators enhance the activity of transcription factors, while mediators act as intermediaries between transcription factors and the RNA polymerase machinery, facilitating the initiation of transcription.



Enhancers and silencers are non promotor elements that can increase ore decrease the rate of transcription in response to the interaction of the proteins binding RNA pol

Some of these elements can act as Enhancers or silencers in response to binding to certain factors

Activators

These regulatory proteins bind to DNA at distant sites known as enhancers. When DNA folds so that the enhancer is brought into proximity with the initiation complex, the activator proteins interact with the complex to increase the rate of transcription.

Coactivators

These transcription factors transmit signals from activator proteins to the general factors.

General Factors

These transcription factors position RNA polymerase at the start of a protein-coding sequence and then release the polymerase to initiate transcription.

Eukaryotic Chromosome Structure

- Eukaryotic DNA is packaged into chromatin which is directly related to the control of gene expression.

Increase transcription rate 》 euchromatin (less tightness)
Decrease transcription rate 》 heterochromatin (more tightness)

- Chromatin structure begins with the organization of the DNA into nucleosomes.
- Nucleosomes may block RNA polymerase II from gaining access to promoters.

The effects of chromatin structure changes on genes expression

1- DNase I hypersensitivity

DNase I hypersensitive sites: more open chromatin configuration site, upstream of the transcription start site

When DNA hypersensitivity occurs, a breakdown of some hydrogen bonds between nucleotides in both strands of DNA occurs, this allows RNA polymerase to access into promoter. Note that polymerases (DNA and RNA) can't work on double strands.So, these strands has to be opened by DNase then helicase breaks more hydrogen bonds, so that transcription continues

2- Histone modification

- Addition of methyl groups to the histone protein tails
- Addition of acetyl groups to histone proteins

by acetylase enzyme to convert heterochromatin to euchromatin to be accessbile for, RNA polymerase in case of transcription ,DNApolymerase in case of replication

This is called chromatin remodeling. The idea of using acetyl group is because it is an acidic group that detaches histone (a basic molecule that contains basic amino acids like lysine) from DNA (an acidic molecule) by acetylase enzyme which provides acetate group for histone in the place of DNA.

Chromatin Structure

- In eukaryotes, the DNA is organized into nucleosomes: about 200 bp of DNA wrapped around a protein core.

- The protein core consists of 8 histone proteins
- Histones are basic (i.e. alkaline): they contain positively charged amino acids that bind to the negative charges on the DNA (backbone phosphate groups).

- DNA tightly wrapped around histones is inaccessible to RNA polymerase

-Thus, one important event in preparing a gene for transcription is "chromatin remodeling": sliding the nucleosomes along the DNA to expose the promoter region.

after DNA synthesis, it's organized on groups of octamer molecules (histone core) consisting of 8 subunits 2 from each type (H2A,H2B,H3,H4) the DNA is organized and wrapped around this histone core (about 200 bp of DNA) The DNA is about 2 meters, therefore it should be packaged to fit inside the nucleus

Histone Acetylation

- A second event needed for transcription affects large regions of the chromosome instead of individual genes.

- DNA is normally tightly wrapped around the histones and is inaccessible to transcription factors.

-The structure can be loosened by acetylating the histones.
- Acetyl groups are added to lysines, which removes their positive charge.

-The binding of the DNA to the histones is lessened, and the DNA structure opens up, allowing access to transcription factors.

↳ by helicase

- Conversely, deacetylation tightens the chromatin structure, preventing transcription throughout that region of the chromosome

Deacetylation occurs after transcription or replication by deacetylase enzyme so DNA returns to make wrapping around histone.

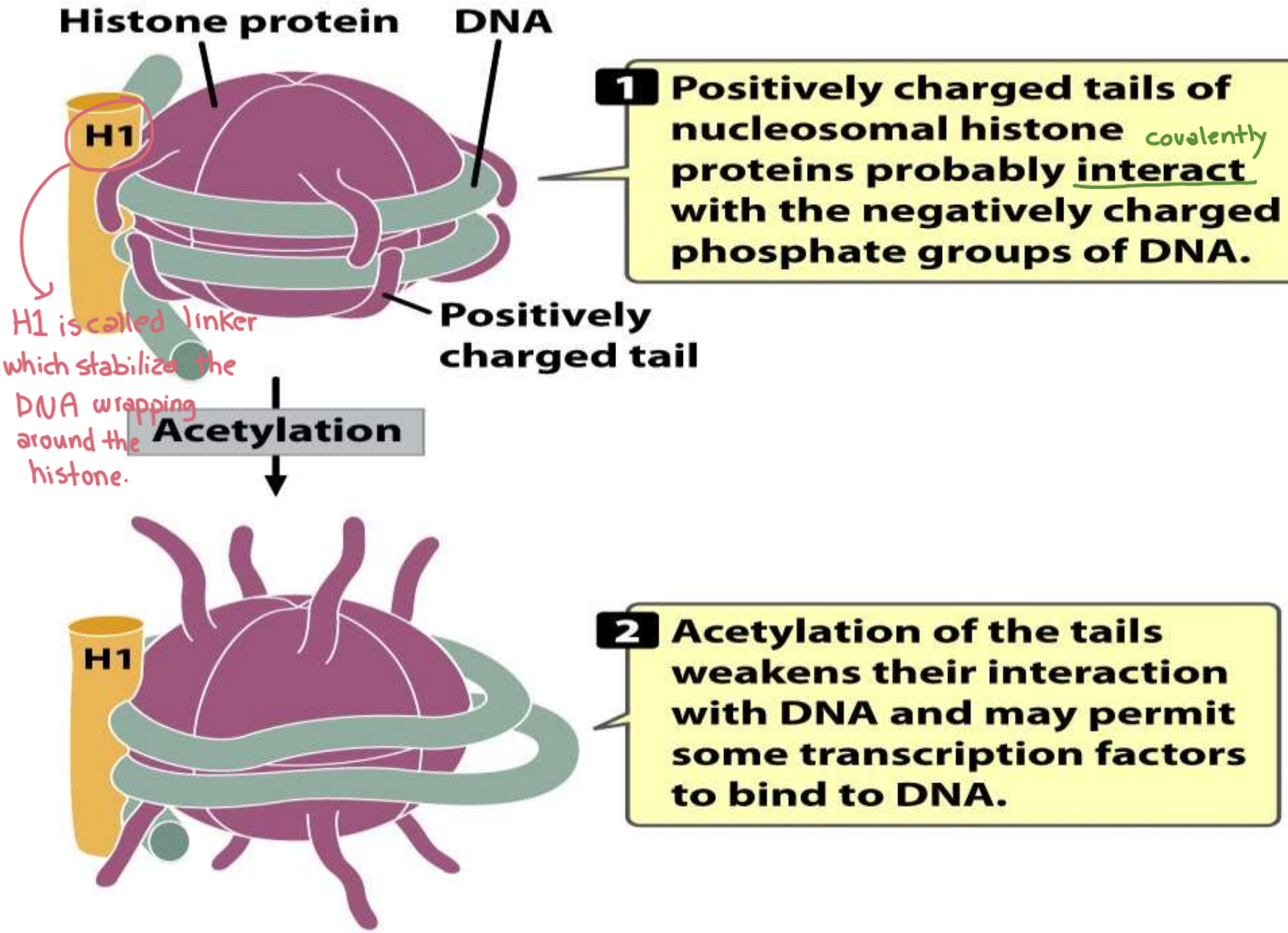


Figure 17-2
Genetics: A Conceptual Approach, Third Edition

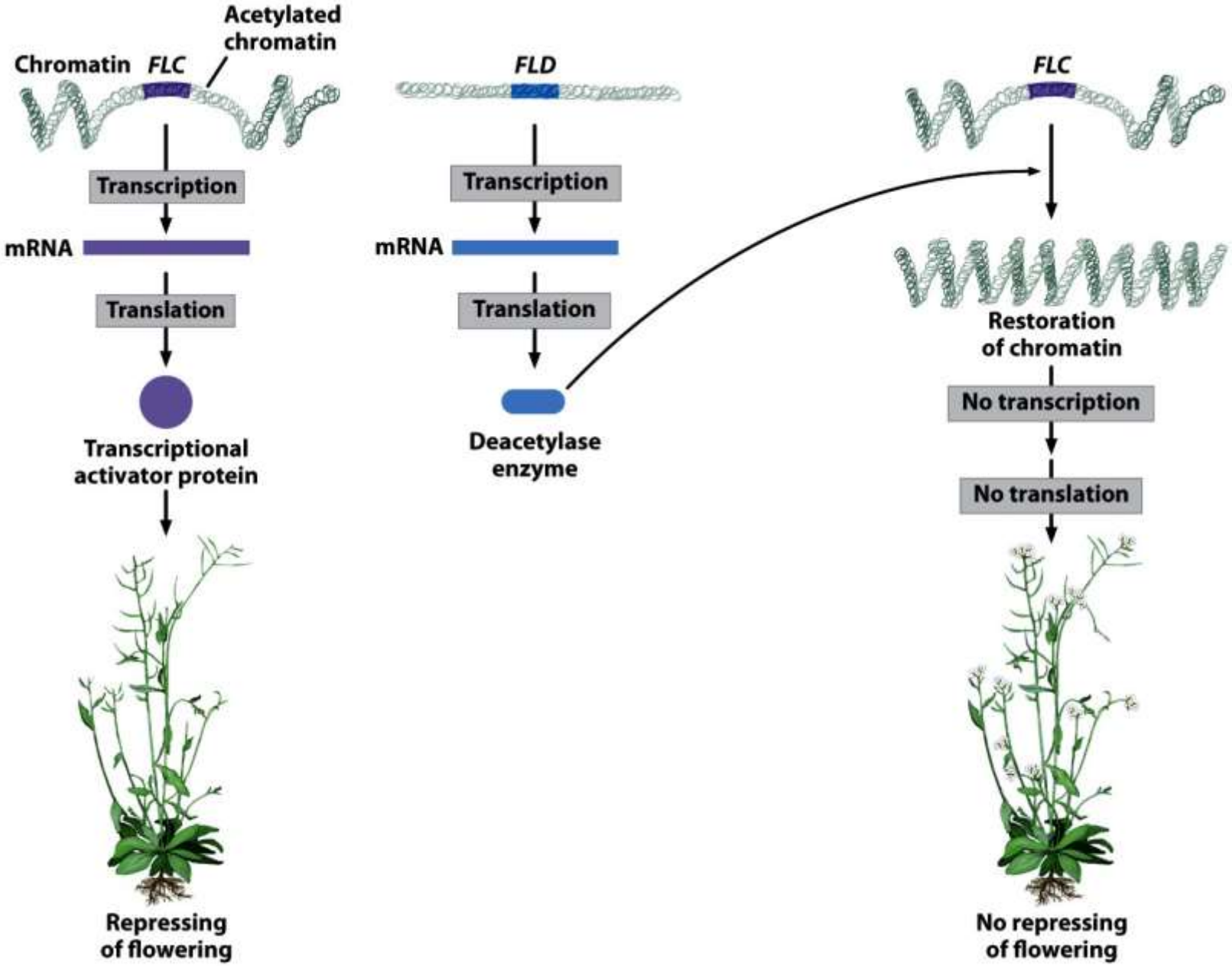


Figure 17-3
Genetics: A Conceptual Approach, Third Edition

Methylation

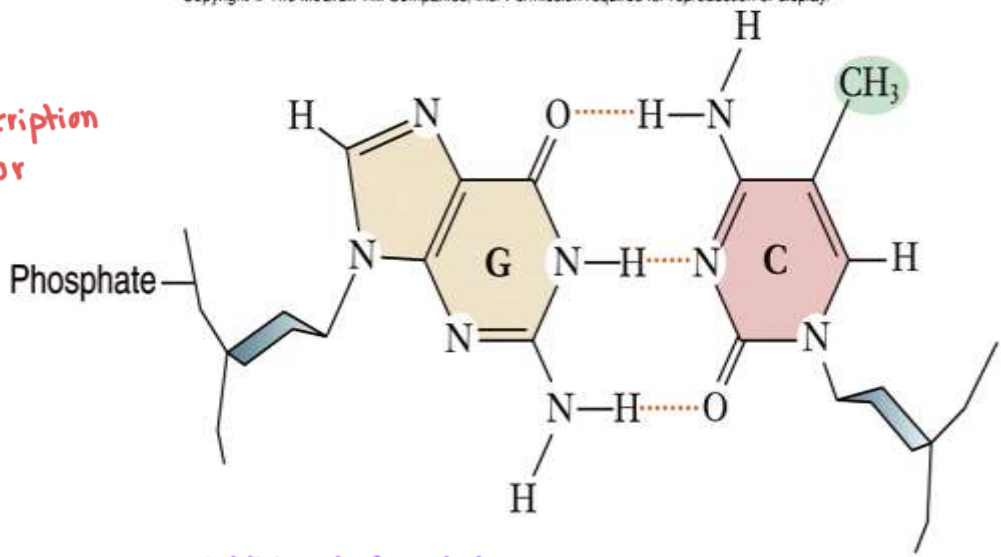
It is used for inhibiting the rate of transcription; it converts large parts of chromatin to be inactive

- Addition of methyl group to DNA or histone proteins is associated with the control of gene expression.
- Clusters of methylated cytosine nucleotides bind to a protein that prevents activators from binding to DNA.
- Methylated histone proteins are associated with inactive regions of chromatin.

So, acetate group bind to basic amino acid in histone and it binds for initiating transcription but methyl group binds to cytosine in DNA or to histone and it binds for inhibiting transcription.

Chromatin remodeling

- Chromatin-remodeling complexes: bind directly to DNA sites and reposition nucleosomes
- DNA methylation of cytosine bases adjacent to guanine nucleotides (CpG)–CpG islands



Additional of methyl group:
1) on DNA 》 methylated cytosine bind protein 》 prevents activators binding
2) on histone 》 inactive regions of chromatin

Posttranscriptional Regulation

- Control of gene expression usually involves the control of transcription initiation.
- But gene expression can be controlled after transcription, with mechanisms such as:

A- RNA interference (This process occur by DICER and RISC) B- alternative splicing

C- RNA editing D- mRNA degradation

A- **RNA interference** involves the use of small RNA molecules

- The enzyme **Dicer** chops double stranded RNA into small pieces of RNA Micro RNA and siRNA are involved
- **Micro-RNAs** bind to complementary RNA to prevent translation
- **Small interfering RNAs** degrade particular mRNAs before translation

B- **Alternative splicing**

- Introns are spliced out of pre-mRNAs to produce the mature mRNA that is translated.
- The spliceosome recognizes different splice sites in different tissue types.

It is optional splicing which means it is not necessary that all exons that produced after removing introns to be joined together. So, it will produce different types of proteins related to exons joining.

- The mature mRNAs in each tissue possess different exons, resulting in different polypeptide products from the same gene.

C- <u>RNA editing</u> creates mature mRNA that are not truly encoded by the genome.

- Editing in the coding region to produce isoforms of the same protein and pre- and post-coding region affecting pre-mRNA splicing, mRNA transport to cytosol or RNA translation effectiveness.

for this reason, the number of proteins in our cells is more than the number of genes.
Nearly double the number, in addition the the definition genome , another one was added which is proteome (complete set of proteins expressed by an organism)

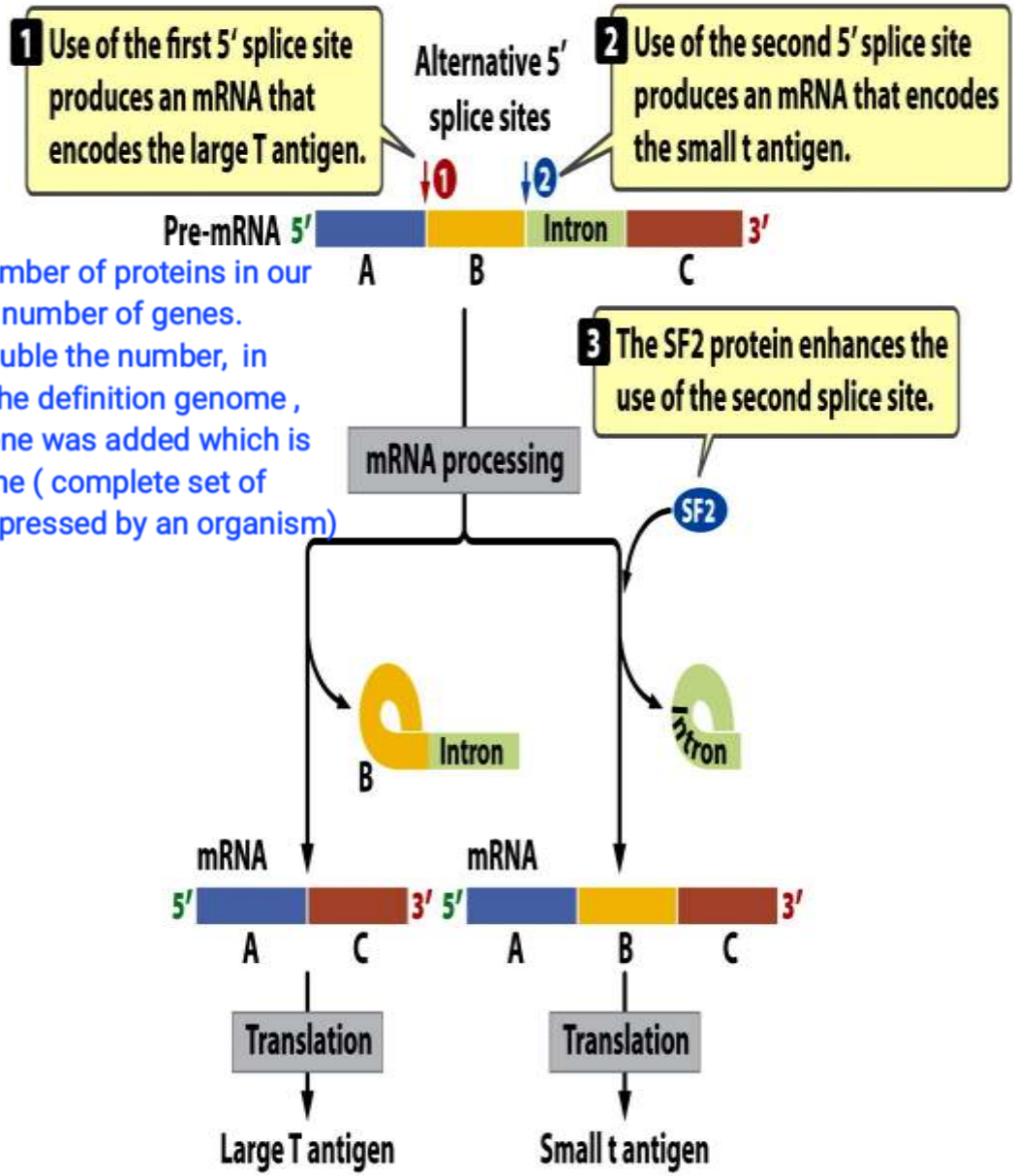


Figure 17-9
Genetics: A Conceptual Approach, Third Edition

there is apoprotein B and it has 2 types:

1) apo B48 which exists in compartments of chylomicron (lipoprotein)

2) apo B100 which exists in very low density and intermediate density lipoproteins and low density lipoprotein(VLDL, LDL and IDL).

In this case, the same gene produces 2 different proteins as a result of RNA editing not alternative splicing.

For example:

• Apolipoprotein B exists in 2 isoforms.

One isoform is produced by editing the mRNA to create a stop codon instead of glutamine codon (Apo B48 in intestine) which is produced by intestinal mucosal cells and (Apo B 100 in liver) which will complete transcription.

This RNA editing is tissue-specific. Which means it does not occur in all cells.

Mature mRNA molecules have various half-lives depending on the gene and thelocation (tissue) of expression and in turn it is affecting the amount of the polypeptide produced.

RNA editing is very important to produce different types of proteins from the gene not by alternative splicing.

- For example:
  - Apolipoprotein B exists in 2 isoforms
  - One isoform is produced by editing the mRNA to create a stop codon instead of glutamine codon (Apo B48 in intestine) and (Apo B 100 in liver)
  - This RNA editing is tissue-specific
- Mature mRNA molecules have various half-lives depending on the gene and the location (tissue) of expression and in turn it is affecting the amount of the polypeptide produced.

in the presence of a protein called localization element.

D- **Control of RNA Transport and intracellular Localization**

- The control of transporting nuclear mRNAs to cytosol and their
- localization to specific cellular compartment which is affected by the post-transcriptional processing of hn-RNA
- The intracellular localization is under effect of elements (localization elements) specified by cis-acting elements (mostly found in 3'UTR)
- Localization elements are recognized by trans-acting factors (RNA-binding proteins)

(another name for localization element). This protein is needed after processing of RNA to be transported to cytosol using it.

E- **RNA processing and degradation can regulate some genes**

- **The degradation of RNA**

- 5′-cap removal
- Shortening of the poly (A) tail
- Degradation of 5′ UTR, coding sequence, and 3′ UTR

5' cap is 7- methyl GTP which has 2 functions: 1) to be recognized by specific protein synthesizing machinary for transition (ribosomes) and 2) to protect mRNA from 5' exonucleases; so, when 5' cap is removed, the 5' end will beexposed to attacks by 5' exonucleases.

located in the 3' end of mRNA for protection against 3' exonucleases. So, when shortening occurs, the 3' end will be exposed to attacks by 3' exonucleases.

3 mechanisms

Mechanisms of Gene regulation by RNA interference RNA cleavage:

1) RISC containing an siRNA, pair with mRNA molecules and cleavage to the mRNA
   - Inhibition of translation
2) Transcriptional silencing: altering chromatin structure
3) Silencer-independent degradation of mRNA

RISC is incorporated with micro RNA and siRNA for degradation of mRNA when it is no longer needed; as in cases of massive increments in rates of transcription which leads to decreasing the rates of translation as the amount of mRNA is limited.

RISC : RNA Induced Silencer Complex

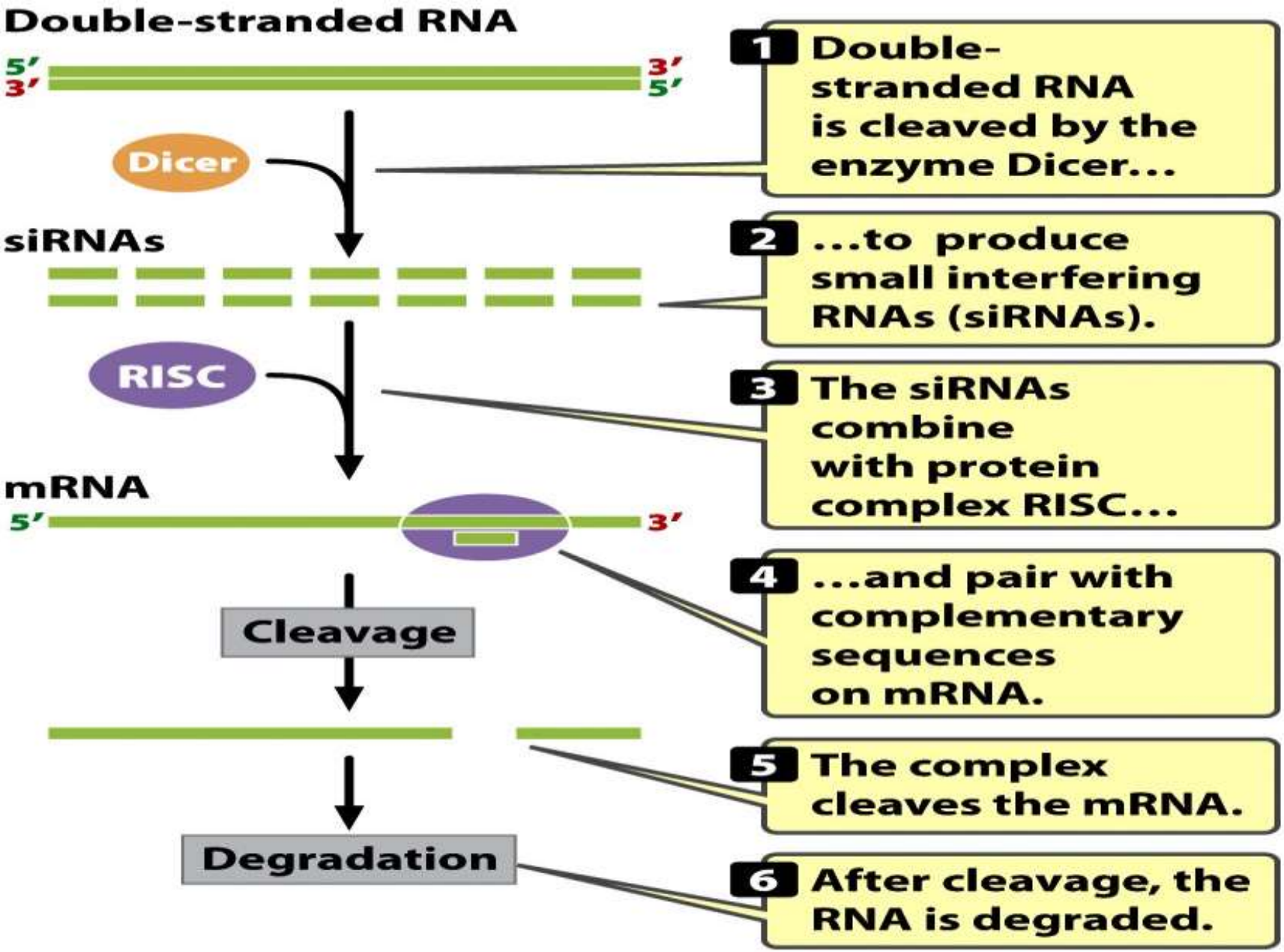


Figure 17-12a
Genetics: A Conceptual Approach, Third Edition

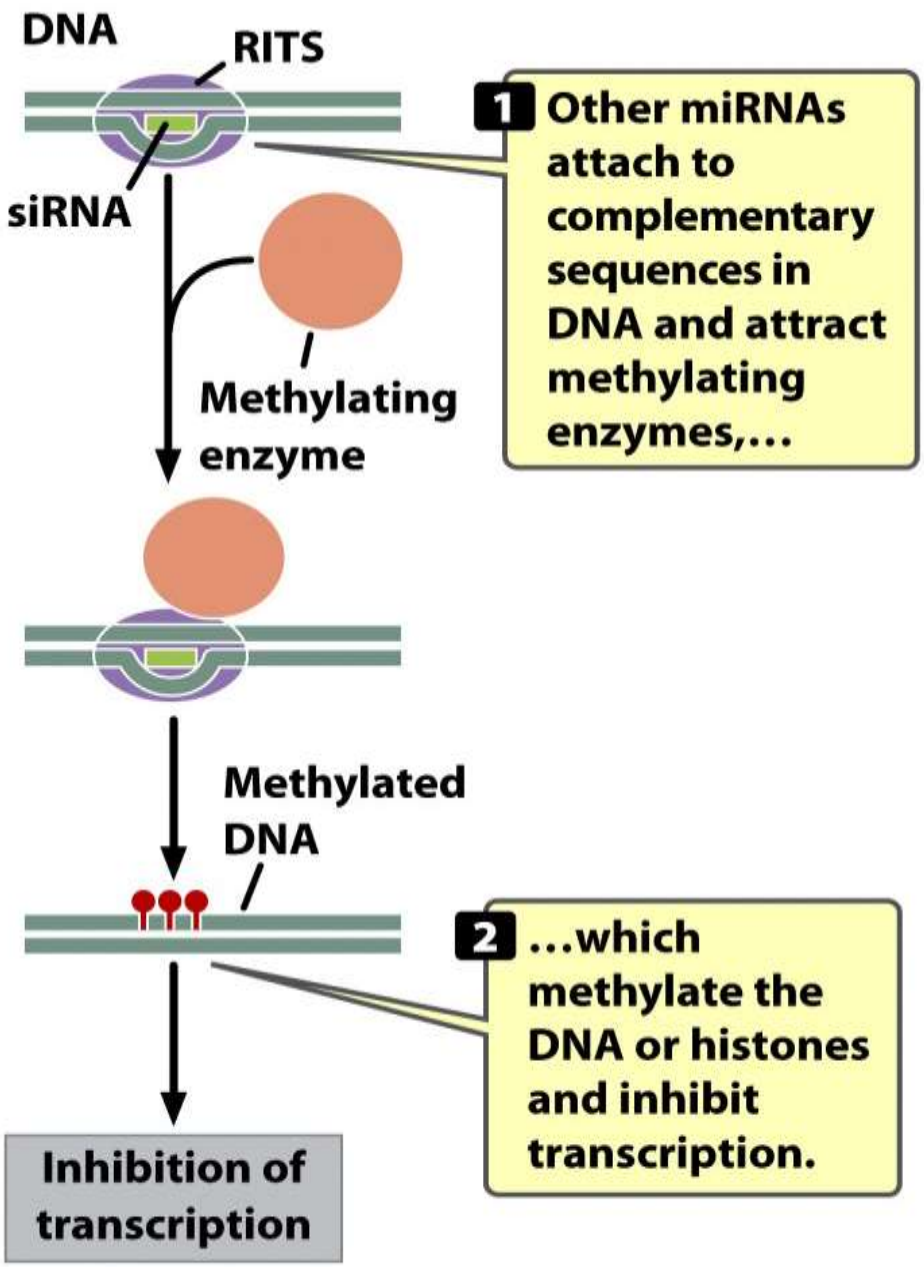


Figure 17-12c
Genetics: A Conceptual Approach, Third Edition

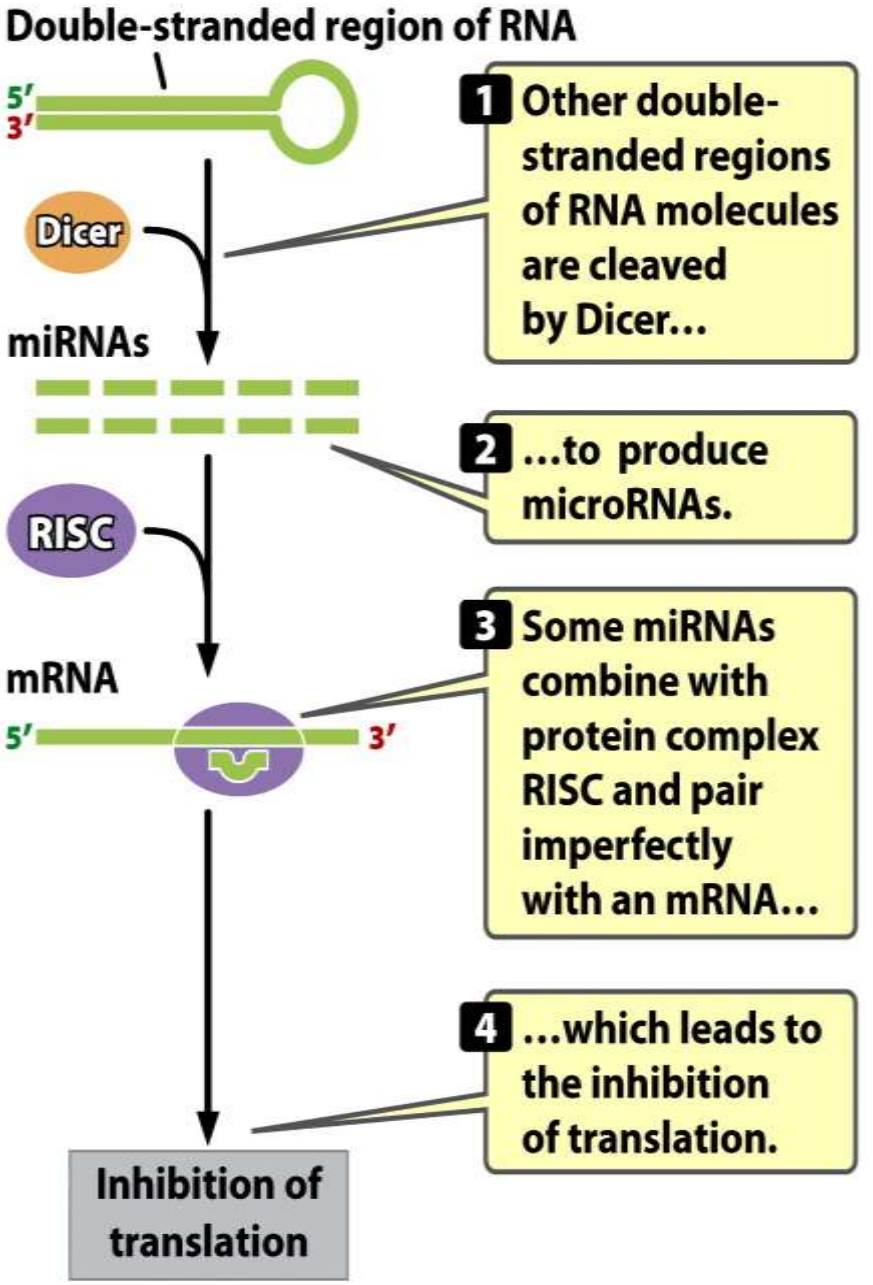


Figure 17-12b
Genetics: A Conceptual Approach, Third Edition

The Fate of Proteins after translation

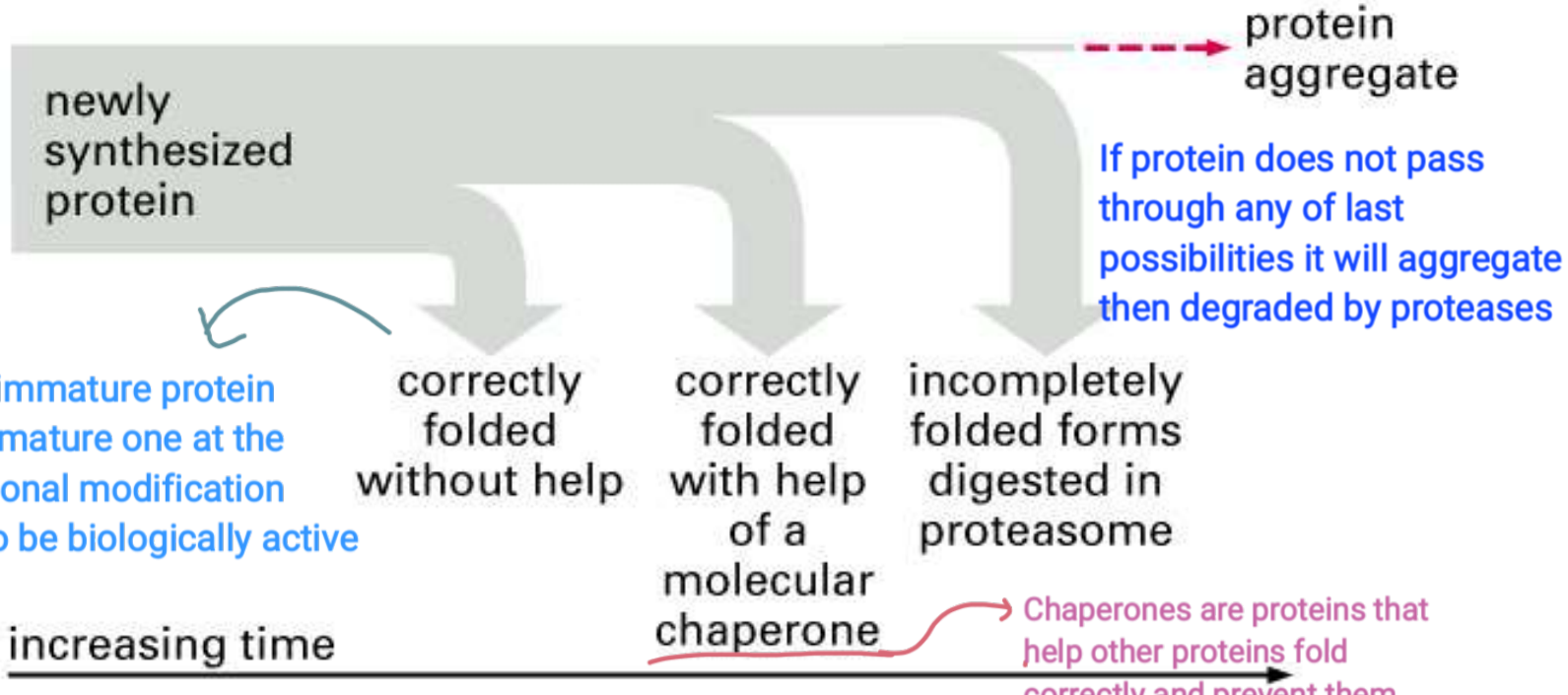


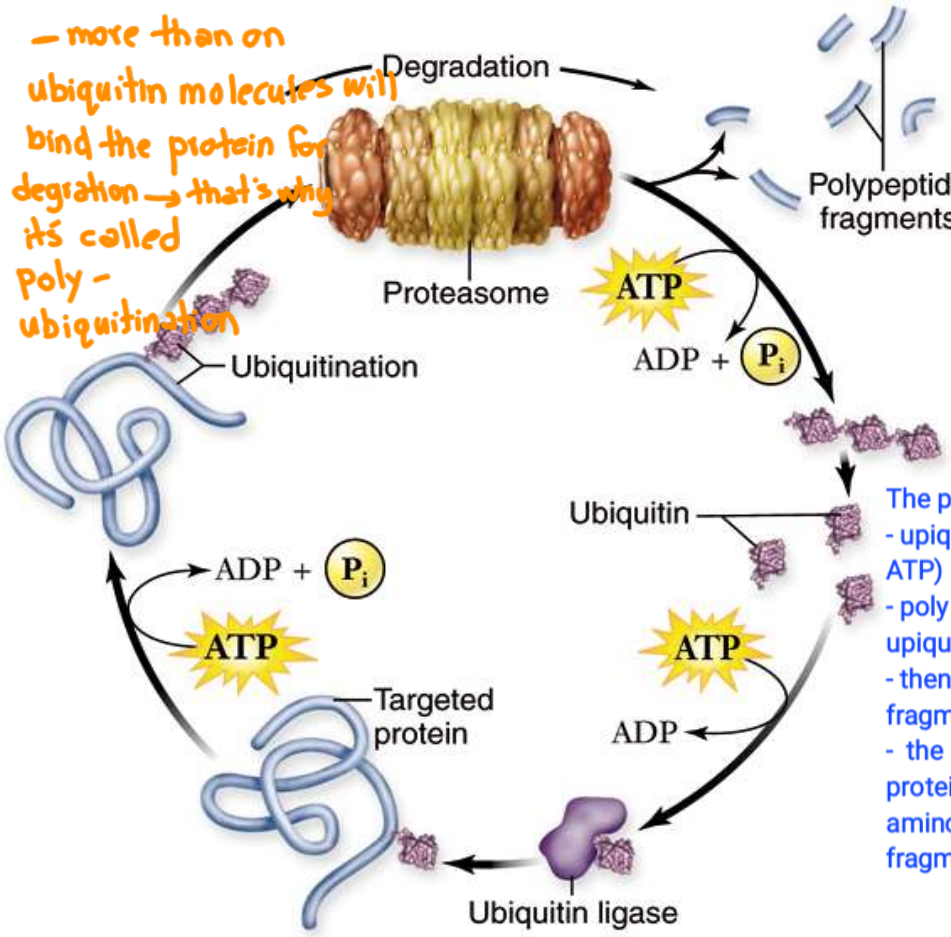
Figure 6–85. Molecular Biology of the Cell, 4th Edition.

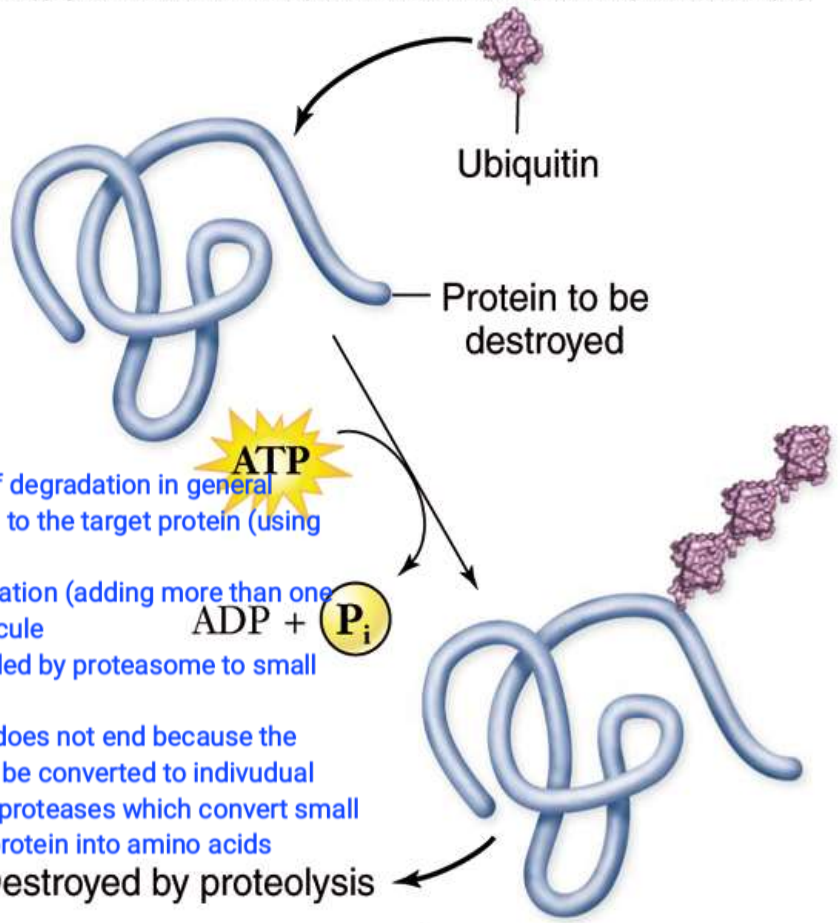
Protein Degradation

Degradation occurs to the protein which is not correctly folded or if there is excess production of a specific protein or a specific cell is going to control the production of specific gene.

- Proteins are produced and degraded continually in the cell.
- Proteins to be degraded are tagged with **ubiquitin**.
- Degradation of proteins marked with ubiquitin occurs at the **proteasome**.

ubiquitin is a protein in nature that is a small molecule (it has a molecular weight of about 7KDa). When it binds to protein it makes a signal to degrade this protein by proteosome. There are a group of enzymes that will bind ubiquitin with the particular protein





The process of degradation in general
- upiquitin bind to the target protein (using ATP)
- poly upiquitination (adding more than one upiquitin molecule
- then it degraded by proteasome to small fragment
- the process does not end because the protein should be converted to indivudual amino acid by proteases which convert small fragments of protein into amino acids

E1: ubiquitin activating enzyme; E2/3: ubiquitin ligase

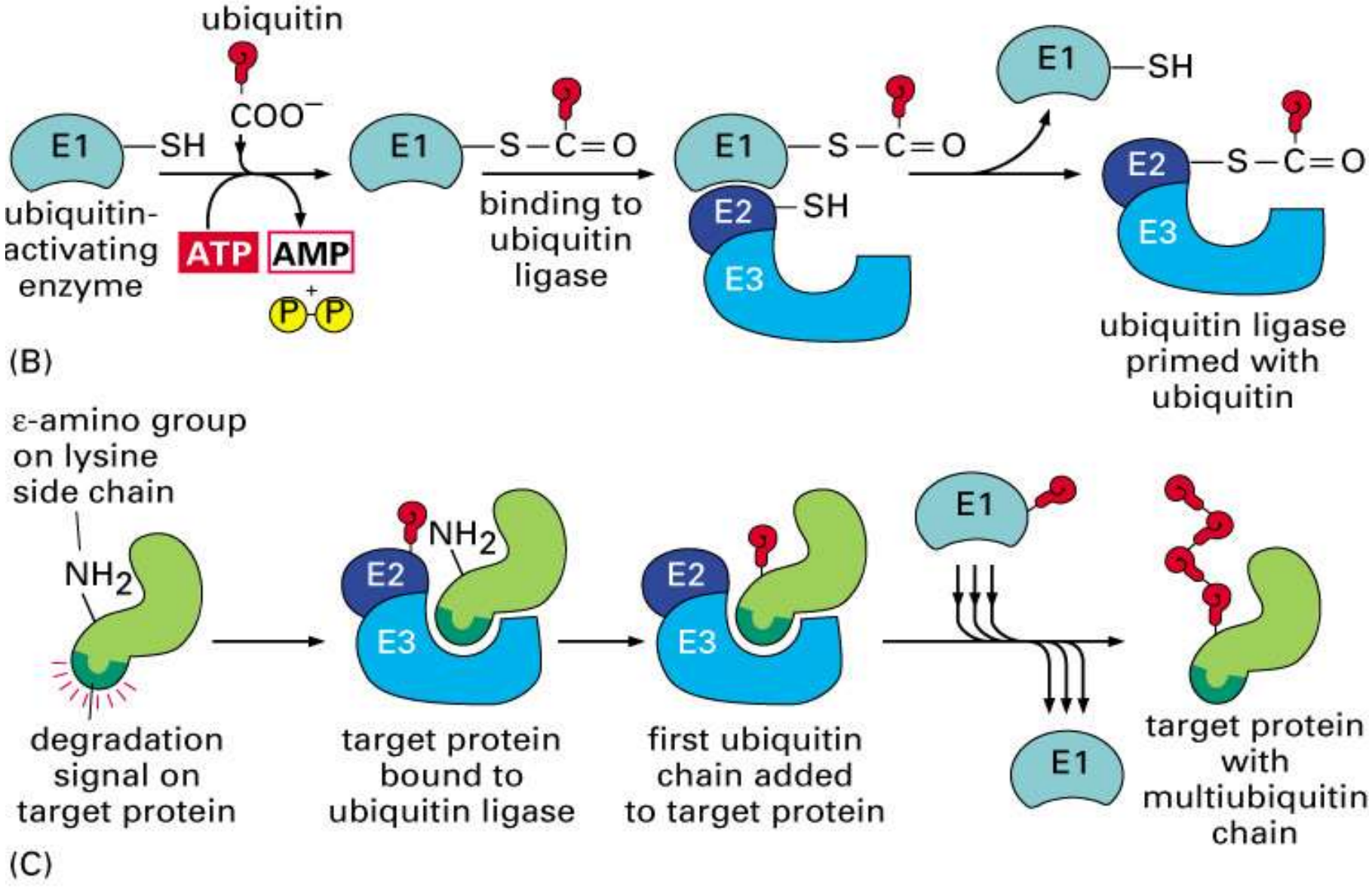


Figure 6–87 part 2 of 2. Molecular Biology of the Cell, 4th Edition.

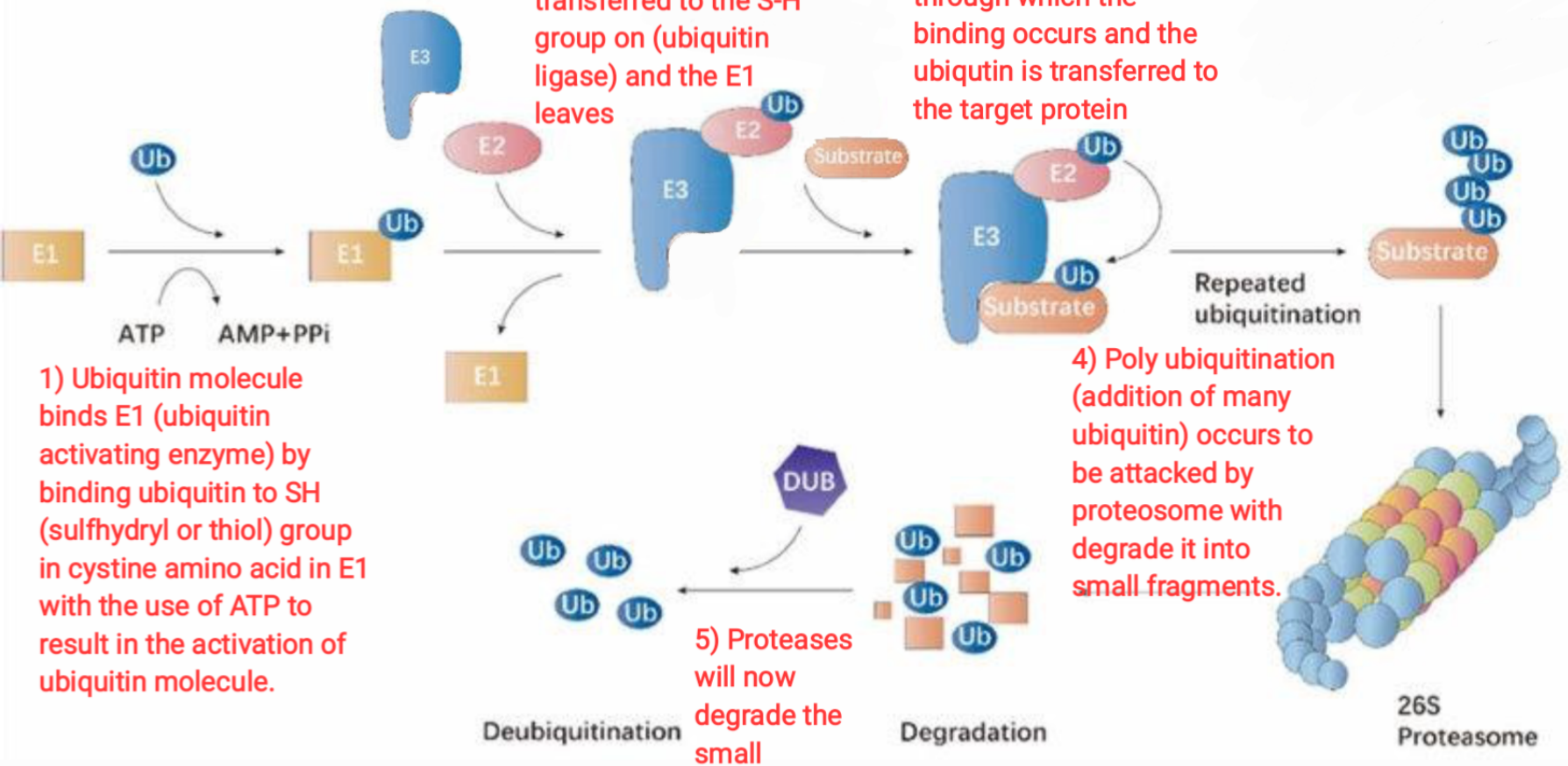
1) Ubiquitin molecule binds E1 (ubiquitin activating enzyme) by binding ubiquitin to SH (sulfhydryl or thiol) group in cystine amino acid in E1 with the use of ATP to result in the activation of ubiquitin molecule.

2) After activation, the complex will bind to E2 and E3 (ubiquitin ligase).
And the ubiquitin is transferred to the S-H group on (ubiquitin ligase) and the E1 leaves

3) Target protein will bind to the (ubiquitin ligase). Note that target protein has degradation signal which occurs through amino group on lysine side chain, through which the binding occurs and the ubiqutin is transferred to the target protein

4) Poly ubiquitination (addition of many ubiquitin) occurs to be attacked by proteosome with degrade it into small fragments.

5) Proteases will now degrade the small fragments into individual amino acids



Activation of a upiquitin ligase occur by using ATP tnen ligases bind to ligand which followed by binding the target protein

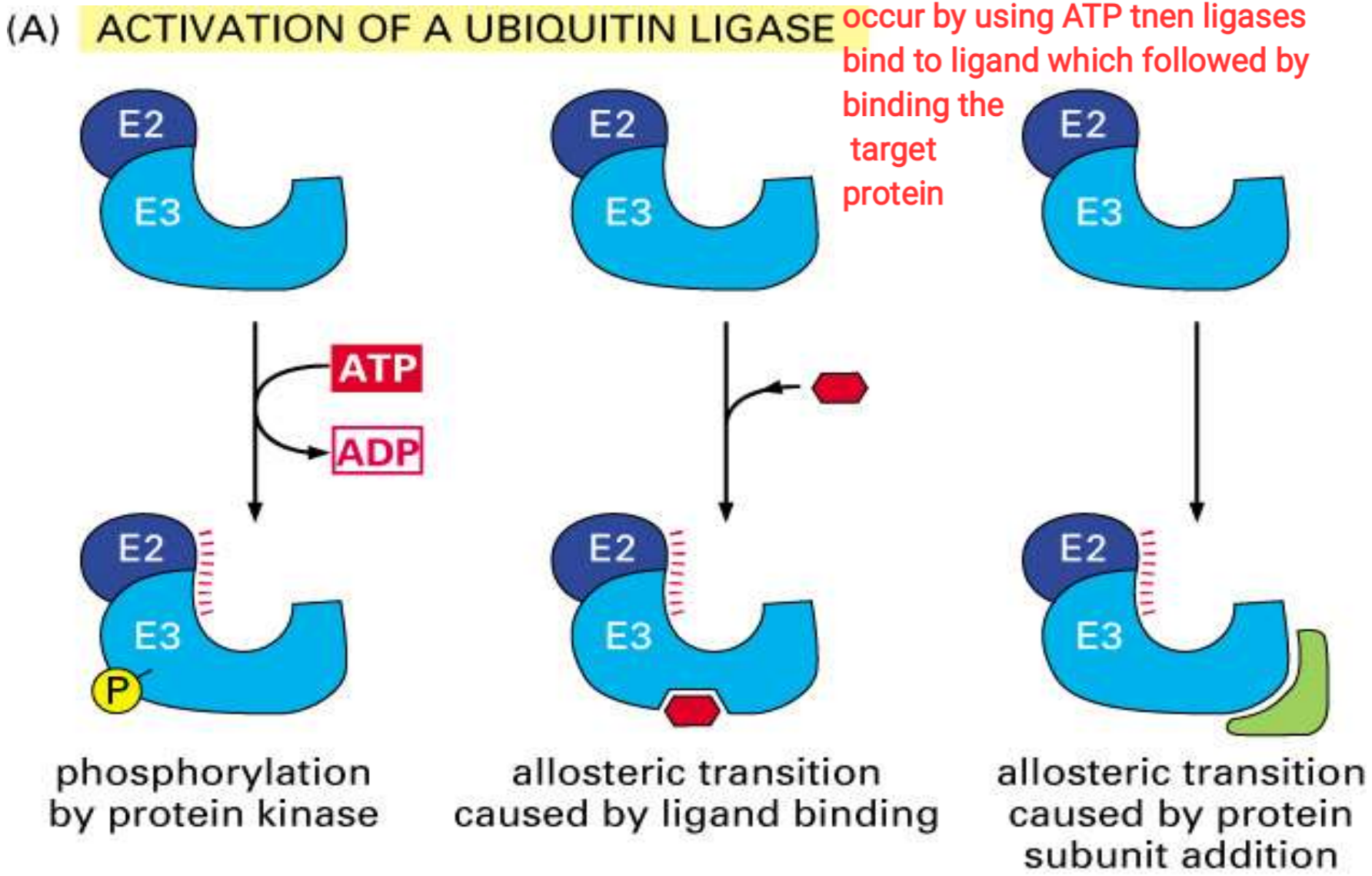
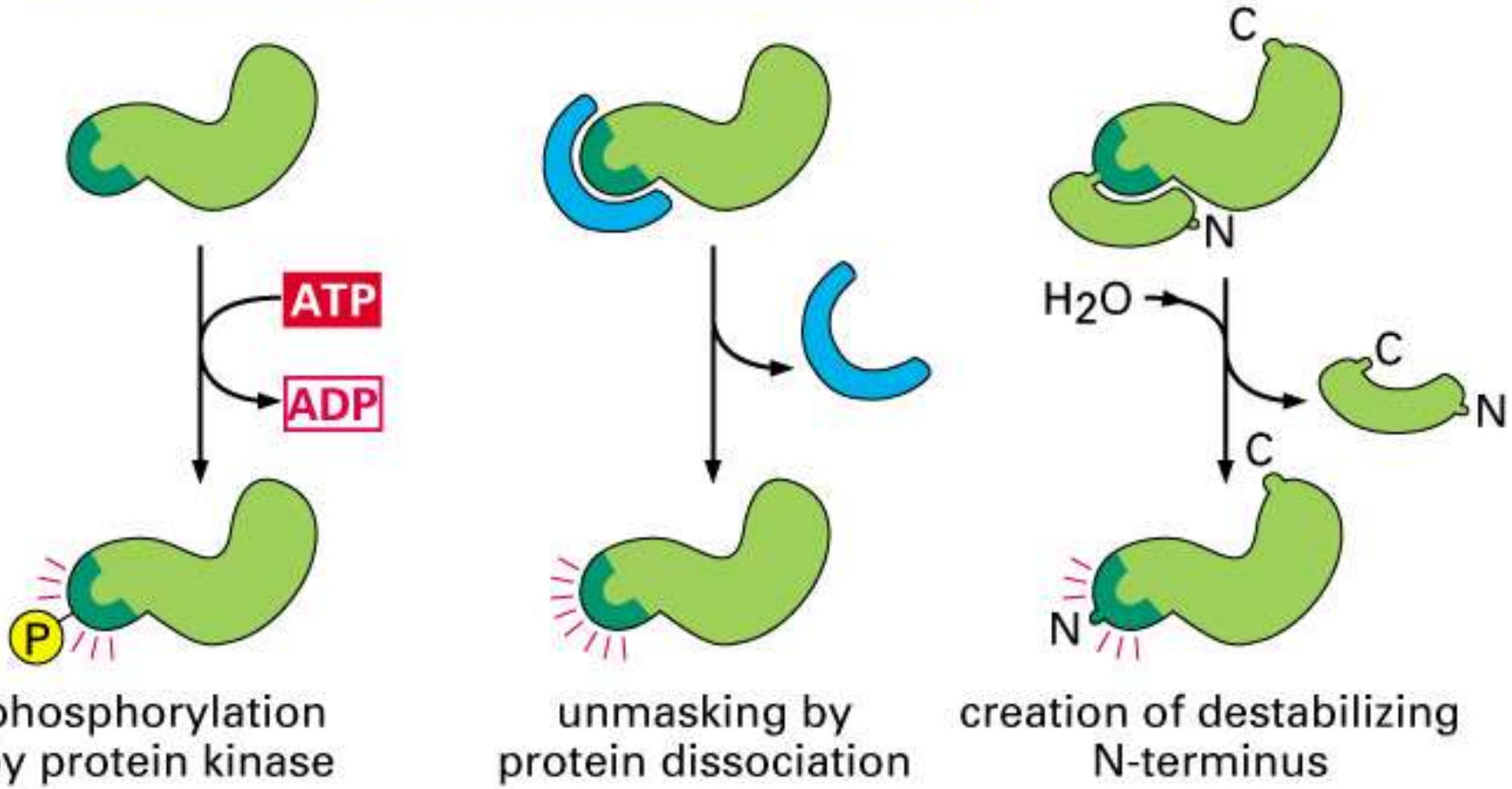


Figure 6–88 part 1 of 2. Molecular Biology of the Cell, 4th Edition.

(B) ACTIVATION OF A DEGRADATION SIGNAL



phosphorylation by protein kinase

unmasking by protein dissociation

creation of destabilizing N-terminus

Degradation signal activated by using ATP so it should be a kinase enzyme to transfer phosphate group from ATP to the target protein then followed by binding the upiquitin molecule which result in creation of destabilizing N terminus

So the present of degradation signal and poly upiquitination will give the chance for proteolytic enzyme belonging to proteasome system to degrade the target protein

Prokaryotic vs. Eukaryotic

- Bacterial genetics are different.
- Prokaryote genes are grouped in operons.
- Prokaryotes have one type of RNA polymerase for all types of RNA
- mRNA is not modified
- The existence of introns in prokaryotes is extremely rare.
- To initiate transcription in bacteria, sigma factors bind tRNA polymerases. RNA polymerases/ sigma factors complex can then bind to promoter prior to the gene coding region.
- In prokaryotes, the newly synthesized mRNA is polycistronic (polygenic) (code for more than one polypeptide chain).
- In prokaryotes, transcription of a gene and translation of the resulting mRNA occur simultaneously, so many polysomes are found associated with an active gene.

* Prokaryotic Cell:

Its genes are grouped in operons

It has one type of RNA polymerase which consists of 4 subunits (a2BB'), it is helped (guided) by sigma factor which can recognize the initiation site by the presence of promoter boxes of prokaryotic genes.

Transcription and translation occur simultaneously.

mRNA in prokaryotes is polycistronic

* Eukaryotic Cell:

It has 3 types of RNA polymerase: 1) produces 3 types of rRNA, 2) produces sRNA andmRNA and 3) produces one type of RNA, tRNA and snRNA (so, RNA polymerase 3 isthe least specialized one because it produces 3 different types of RNA).

Its genes are monocistronic

Transcription and translation are separated because it has nuclear membrane.

mRNA should be processed before translation and it is transferred by localization element to the cytosol.