The Flow of Genetic Information

- The information content of DNA is in the form of specific sequences of nucleotides
- The DNA inherited by an organism leads to specific traits by dictating the synthesis of proteins
- Proteins are the links between genotype and phenotype
- Gene expression, the process by which DNA directs protein synthesis, includes two stages: transcription and translation
- The ribosome is part of the cellular machinery for translation, polypeptide synthesis

Basic Principles of Transcription and Translation

- RNA is the intermediate between genes and the proteins for which they code
- Transcription is the synthesis of RNA under the direction of DNA
- Transcription produces messenger RNA (mRNA)
- Translation is the synthesis of a polypeptide, which occurs under the direction of mRNA
- Ribosomes are the sites of translation
• In prokaryotes, mRNA produced by transcription is immediately translated without more processing

• In a eukaryotic cell, the nuclear envelope separates transcription from translation

• Eukaryotic RNA transcripts are modified through RNA processing to yield finished mRNA

• A primary transcript is the initial RNA transcript from any gene

• The central dogma is the concept that cells are governed by a cellular chain of command: DNA → RNA → protein

• Eukaryotic RNA transcripts are modified through RNA processing to yield finished mRNA

---

a) Bacterial cell. In a bacterial cell which lacks a nucleus, mRNA produced by transcription is immediately translated without additional processing.

b) Eukaryotic cell. The nucleus provides a separate compartment for transcription. The original RNA transcript called pre-mRNA is processed in various ways before leaving the nucleus as mRNA.

Overview: the roles of transcription and translation in the flow of genetic information. In a cell inherited information flows from DNA to RNA to protein. The two main stages of information flow are transcription and translation.
(a) Bacterial cell

(b) Eukaryotic cell
(b) Eukaryotic cell

- **TRANSCRIPTION**
  - DNA
  - Pre-mRNA

- **RNA PROCESSING**
  - Pre-mRNA
  - mRNA

- **TRANSLATION**
  - Ribosome
  - Polypeptide

**Nuclear envelope**
The Genetic Code

- How are the instructions for assembling amino acids into proteins encoded into DNA?

- There are 20 amino acids, but there are only four nucleotide bases in DNA

- How many bases correspond to an amino acid?

Codons: Triplets of Bases

- The flow of information from gene to protein is based on a triplet code: a series of nonoverlapping, three-nucleotide words

- These triplets are the smallest units of uniform length that can code for all the amino acids

- Example: AGT at a particular position on a DNA strand results in the placement of the amino acid serine at the corresponding position of the polypeptide to be produced
During transcription, one of the two DNA strands called the **template strand** provides a template for ordering the sequence of nucleotides in an RNA transcript.

During translation, the mRNA base triplets, called **codons**, are read in the 5′ to 3′ direction.

Each codon specifies the amino acid to be placed at the corresponding position along a polypeptide.

Colons along an mRNA molecule are read by translation machinery in the 5′ to 3′ direction.

Each codon specifies the addition of one of 20 amino acids.

The triplet code. For each gene, one strand of DNA functions as a template for transcription. The base pairing rules for DNA synthesis also guide transcription, but uracil (U) takes the place of thymine (T) in RNA. During translation the mRNA is read as a sequence of base triplets called codons. Each codon specifies an amino acid to be added to the growing polypeptide chain. The mRNA is read in the 5′⇒ 3′ direction.
Cracking the Code

- All 64 codons were deciphered by the mid-1960s

- Of the 64 triplets, 61 code for amino acids; 3 triplets are “stop” signals to end translation

- The genetic code is redundant but not ambiguous; no codon specifies more than one amino acid

- Codons must be read in the correct reading frame (correct groupings) in order for the specified polypeptide to be produced

The dictionary of the genetic code. The three bases of an mRNA codon are designated here as the first, second and third bases reading in the 5′ ⇒ 3′ direction along the mRNA. The codon AUG not only stands for the amino acid methionine (Met) but also functions as a start signal for ribosomes to begin translating the mRNA at that point. Three of the 64 codons function as “stop” signals marking the end of a genetic message.
Evolution of the Genetic Code

- The genetic code is nearly universal, shared by the simplest bacteria to the most complex animals
- Genes can be transcribed and translated after being transplanted from one species to another
- Because diverse forms of life share a common genetic code, one species can be programmed to produce proteins characteristic of a second species by introducing DNA from the second species into the first

Transcription is the DNA-directed synthesis of RNA: a closer look

- Transcription, the first stage of gene expression, can be examined in more detail
- The three stages of transcription:
  - Initiation
  - Elongation
  - Termination
Molecular Components of Transcription

- RNA synthesis is catalyzed by **RNA polymerase**, which pries the DNA strands apart and hooks together the RNA nucleotides.
- RNA synthesis follows the same base-pairing rules as DNA, except uracil substitutes for thymine.
- The DNA sequence where RNA polymerase attaches is called the **promoter**; in bacteria, the sequence signaling the end of transcription is called the **terminator**.
- The stretch of DNA that is transcribed is called a **transcription unit**.
1) Initiation. After RNA polymerase binds to the promoter, the DNA strands unwind and the polymerase initiates RNA synthesis at the start point on the template strand.

2) Elongation. The polymerase moves downstream unwinding the DNA and elongating the RNA transcript 5'→3'. In the wake of transcription, the DNA strands re-form a double helix.

3) Termination. Eventually, the RNA transcript is released and the polymerase detaches from the DNA.

The stages of transcription: initiation, elongation, and termination. Thus, a general depiction of transcription applies to both bacteria and eukaryotes but the details of termination differ, as described in the text. Also in a bacterium, the transcript is immediately usable as mRNA in a eukaryote, the RNA transcript must first undergo processing.
RNA Polymerase Binding and Initiation of Transcription

- Promoters signal the initiation of RNA synthesis
- **Transcription factors** mediate the binding of RNA polymerase and the initiation of transcription
- The completed assembly of transcription factors and RNA polymerase II bound to a promoter is called a transcription initiation complex
- A promoter called a **TATA box** is crucial in forming the initiation complex in eukaryotes

1. A eukaryotic promoter commonly includes a TATA box a nucleotide sequence containing TATA about 25 nucleotides upstream from the transcription start point. By convention nucleotide sequences are given as they occur on the nontemplate strand.

2. Several transcription factors one recognizing the TATA box must bind to the DNA before RNA polymerase II can do so.

3. Additional transcription factors (purple) bind to the DNA along with RNA polymerase II forming the transcription initiation complex. The DNA double helix then unwinds and RNA synthesis begins at the start point on the template strand.

The initiation of transcription at a eukaryotic promoter. In eukaryotic cells proteins called transcription factors mediate the initiation of transcription by RNA polymerase II.
Elongation of the RNA Strand

- As RNA polymerase moves along the DNA, it untwists the double helix, 10 to 20 bases at a time.
- Transcription progresses at a rate of 40 nucleotides per second in eukaryotes.
- A gene can be transcribed simultaneously by several RNA polymerases.

Termination of Transcription

- The mechanisms of termination are different in bacteria and eukaryotes.
- In bacteria, the polymerase stops transcription at the end of the terminator.
- In eukaryotes, the polymerase continues transcription after the pre-mRNA is cleaved from the growing RNA chain; the polymerase eventually falls off the DNA.
Eukaryotic cells modify RNA after transcription

- Enzymes in the eukaryotic nucleus modify pre-mRNA before the genetic messages are dispatched to the cytoplasm.
- During RNA processing, both ends of the primary transcript are usually altered.
- Also, usually some interior parts of the molecule are cut out, and the other parts spliced together.

Alteration of mRNA Ends

- Each end of a pre-mRNA molecule is modified in a particular way:
  - The 5′ end receives a modified nucleotide **5′ cap**
  - The 3′ end gets a **poly-A tail**

- These modifications share several functions:
  - They seem to facilitate the export of mRNA
  - They protect mRNA from hydrolytic enzymes
  - They help ribosomes attach to the 5′ end
RNA processing addition of the 5’ cap and poly A tail. Enzymes modify the two ends of a eukaryotic pre mRNA molecule. The modified ends may promote the export of mRNA from the nucleus and they help protect the mRNA from degradation. When the mRNA reaches the cytoplasm the modified ends in conjunction with certain cytoplasmic proteins facilitate ribosome attachment. The 5’ cap and poly A tail are not translated into protein, nor are the regions called the 5’ untranslated regions (5’ UTR) and 3’ untranslated regions (3’ UTR).

Split Genes and RNA Splicing

- Most eukaryotic genes and their RNA transcripts have long noncoding stretches of nucleotides that lie between coding regions.
- These noncoding regions are called intervening sequences, or introns.
- The other regions are called exons because they are eventually expressed, usually translated into amino acid sequences.
- RNA splicing removes introns and joins exons, creating an mRNA molecule with a continuous coding sequence.
RNA processing: mRNA splicing. The RNA molecule shown here codes for β globin one of the polypeptides of hemoglobin. The numbers under the RNA refer to the codons. β globin is 146 amino acids long. The β globin gene and its pre mRNA transcript have three exons corresponding to sequences that will leave the nucleus as RNA. (The 5' UTR and 3' UTR are parts of exons because they are included in the mRNA however they do not code for protein). During RNA processing the introns are cut out and the exons are spliced together. In many genes the introns are much larger relative to the exons than they are in the β globin gene. The mRNA is not drawn to scale.

- In some cases, RNA splicing is carried out by spliceosomes

- **Spliceosomes** consist of a variety of proteins and several small nuclear ribonucleoproteins (snRNPs) that recognize the splice sites
The roles of snRNPs and spliceosomes in pre-mRNA splicing. The diagram shows only a portion of the pre-mRNA transcript, additional introns and exons lie downstream from the pictured ones here. 1) Small nuclear ribonucleoprotein (snRNPs) and other proteins form a molecular complex called a spliceosome on a pre-mRNA molecules containing exons and introns. 2) Within the spliceosome snRNA base pairs with nucleotides at specific sites along the intron. 3) The spliceosome cuts the pre mRNA releasing the intron and at the same time splices the exons together. The spliceosome then comes apart releasing mRNA which now contains only exons.

**Ribozymes**

- **Ribozymes** are catalytic RNA molecules that function as enzymes and can splice RNA

- The discovery of ribozymes rendered obsolete the belief that all biological catalysts were proteins

- Three properties of RNA enable it to function as an enzyme
  - It can form a three-dimensional structure because of its ability to base pair with itself
  - Some bases in RNA contain functional groups
  - RNA may hydrogen-bond with other nucleic acid molecules
The Functional and Evolutionary Importance of Introns

- Some genes can encode more than one kind of polypeptide, depending on which segments are treated as exons during RNA splicing
- Such variations are called alternative RNA splicing
- Because of alternative splicing, the number of different proteins an organism can produce is much greater than its number of genes

- Proteins often have a modular architecture consisting of discrete regions called domains
- In many cases, different exons code for the different domains in a protein
- Exon shuffling may result in the evolution of new proteins
Translation is the RNA-directed synthesis of a polypeptide: *a closer look*

- A cell translates an mRNA message into protein with the help of **transfer RNA (tRNA)**
- Molecules of tRNA are not identical:
  - Each carries a specific amino acid on one end
  - Each has an **anticodon** on the other end; the anticodon base-pairs with a complementary codon on mRNA
Translation: the basic concept. As a molecule of mRNA is moved through a ribosome codons are translated into amino acids one by one. The interpreters are tRNA molecules, each type with a specific anticodon at one end and a corresponding amino acid at the other end. A tRNA adds its amino acid cargo to a growing polypeptide chain when the anticodon hydrogen bonds to a complementary codon on the mRNA.

The Structure and Function of Transfer RNA

- A tRNA molecule consists of a single RNA strand that is only about 80 nucleotides long
- Flattened into one plane to reveal its base pairing, a tRNA molecule looks like a cloverleaf
- Because of hydrogen bonds, tRNA actually twists and folds into a three-dimensional molecule
- tRNA is roughly L-shaped
Accurate translation requires two steps:

- First: a correct match between a tRNA and an amino acid, done by the enzyme **aminoacyl-tRNA synthetase**
- Second: a correct match between the tRNA anticodon and an mRNA codon

Flexible pairing at the third base of a codon is called **wobble** and allows some tRNAs to bind to more than one codon.
An aminoacyl tRNA synthethase joining a specific amino acid to a tRNA. Linkage of the tRNA and amino acid is an endergonic process that occurs at the expense of ATP. The ATP loses two phosphate groups becoming AMP (adenosine monophosphate).

**Ribosomes**

- Ribosomes facilitate specific coupling of tRNA anticodons with mRNA codons in protein synthesis
- The two ribosomal subunits (large and small) are made of proteins and **ribosomal RNA (rRNA)**
A ribosome has three binding sites for tRNA:

- The **P site** holds the tRNA that carries the growing polypeptide chain
- The **A site** holds the tRNA that carries the next amino acid to be added to the chain
- The **E site** is the exit site, where discharged tRNAs leave the ribosome
Building a Polypeptide

- The three stages of translation:
  - Initiation
  - Elongation
  - Termination

- All three stages require protein “factors” that aid in the translation process

Ribosome Association and Initiation of Translation

- The initiation stage of translation brings together mRNA, a tRNA with the first amino acid, and the two ribosomal subunits

- First, a small ribosomal subunit binds with mRNA and a special initiator tRNA

- Then the small subunit moves along the mRNA until it reaches the start codon (AUG)

- Proteins called initiation factors bring in the large subunit that completes the translation initiation complex
The initiation of translation

1) A small ribosomal subunit binds to a molecule of mRNA. In a bacterial cell the mRNA binding site on this subunit recognizes a specific nucleotide sequence on the mRNA just upstream of the start codon. An initiator tRNA with the anticodon UAC base pairs with the start codon, AUG. This tRNA carries the amino acid methionine Met.

2) The arrival of a large ribosomal subunit completes the initiation complex. Proteins called initiation factors (not shown) are required to bring all the translation components together. GTP provides the energy for the assembly. The initiator tRNA is in the P site, the A site is available to the tRNA bearing the next amino acid.

Elongation of the Polypeptide Chain

- During the elongation stage, amino acids are added one by one to the preceding amino acid.

- Each addition involves proteins called elongation factors and occurs in three steps: codon recognition, peptide bond formation, and translocation.
Termination of Translation

- Termination occurs when a stop codon in the mRNA reaches the A site of the ribosome.
- The A site accepts a protein called a release factor.
- The release factor causes the addition of a water molecule instead of an amino acid.
- This reaction releases the polypeptide, and the translation assembly then comes apart.
The release factor hydrolyzes the bond between the tRNA in the P site and the last amino acid of the polypeptide chain. The polypeptide is thus freed from the ribosome.

When a ribosome reaches a stop codon on mRNA, the A site of the ribosome accepts a release factor, a protein shaped like a tRNA instead of an aminoacyl tRNA.

The termination of translation. Like elongation, termination requires GTP hydrolysis as well as additional factors which are not shown here.

**Polyribosomes**

- A number of ribosomes can translate a single mRNA simultaneously, forming a **polyribosome** (or **polysome**).
- Polyribosomes enable a cell to make many copies of a polypeptide very quickly.
Completing and Targeting the Functional Protein

- Often translation is not sufficient to make a functional protein
- Polypeptide chains are modified after translation
- Completed proteins are targeted to specific sites in the cell
**Protein Folding and Post-Translational Modifications**

- During and after synthesis, a polypeptide chain spontaneously coils and folds into its three-dimensional shape.
- Proteins may also require post-translational modifications before doing their job.
- Some polypeptides are activated by enzymes that cleave them.
- Other polypeptides come together to form the subunits of a protein.

**Targeting Polypeptides to Specific Locations**

- Two populations of ribosomes are evident in cells: free ribosomes (in the cytosol) and bound ribosomes (attached to the ER).
- Free ribosomes mostly synthesize proteins that function in the cytosol.
- Bound ribosomes make proteins of the endomembrane system and proteins that are secreted from the cell.
- Ribosomes are identical and can switch from free to bound.
The signal mechanism for targeting proteins to the ER. A polypeptide destined for the endomembrane system or for secretion from the cell begins with a signal peptide a series of amino acids that targets it for the ER. This figure shows the synthesis of a secretory protein and its simultaneous import into the ER. In the ER and then in the Golgi, the protein will be processed further. Finally a transport vesicle will convey it to the plasma membrane for release from the cell.
RNA plays multiple roles in the cell: *a review*

<table>
<thead>
<tr>
<th>Type of RNA</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messenger RNA (mRNA)</td>
<td>Carries information specifying amino acid sequences of proteins from DNA to ribosomes</td>
</tr>
<tr>
<td>Transfer RNA (tRNA)</td>
<td>Serves as adapter molecule in protein synthesis; translates mRNA codons into amino acids</td>
</tr>
<tr>
<td>Ribosomal RNA (rRNA)</td>
<td>Plays catalytic (ribozyme) roles and structural roles in ribosomes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of RNA</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary transcript</td>
<td>Serves as a precursor to mRNA, rRNA, or tRNA, before being processed by splicing or cleavage</td>
</tr>
<tr>
<td>Small nuclear RNA (snRNA)</td>
<td>Plays structural and catalytic roles in spliceosomes</td>
</tr>
<tr>
<td>SRP RNA</td>
<td>Is a component of the signal-recognition particle (SRP)</td>
</tr>
</tbody>
</table>
RNA’s diverse functions range from structural to informational to catalytic.

Properties that enable RNA to perform many different functions:

- Can hydrogen-bond to other nucleic acids
- Can assume a three-dimensional shape
- Has functional groups that allow it to act as a catalyst (ribozyme)
While gene expression differs among the domains of life, the concept of a gene is universal

- Archaea are prokaryotes, but share many features of gene expression with eukaryotes

Comparing Gene Expression in Bacteria, Archaea, and Eukarya

- Bacteria and eukarya differ in their RNA polymerases, termination of transcription and ribosomes; archaea tend to resemble eukarya in these respects

- Bacteria can simultaneously transcribe and translate the same gene

- In eukarya, transcription and translation are separated by the nuclear envelope. In addition extensive RNA processing occurs in the nucleus

- In archaea, transcription and translation are likely coupled
Coupled transcription and translation in bacteria. In bacterial cells, the translation of mRNA can begin as soon as the leading (5') end of the mRNA molecule peels away from the DNA template. The micrograph (TEM) shows a strand of *Escherichia coli* DNA being transcribed by RNA polymerase molecules. Attached to each RNA polymerase molecule is a growing strand of mRNA which is already being translated by ribosomes. The newly synthesized polypeptides are not visible in the micrograph but are shown in the diagram.

---

**What Is a Gene? Revisiting the Question**

- The idea of the gene itself is a unifying concept of life

- We have considered a gene as:
  - A discrete unit of inheritance
  - A region of specific nucleotide sequence in a chromosome
  - A DNA sequence that codes for a specific polypeptide chain

- In summary, a gene can be defined as a region of DNA that can be expressed to produce a final functional product, either a polypeptide or an RNA molecule
A summary of transcription and translation in a eukaryotic cell

1) Transcription- RNA is transcribed from a DNA template.

2) RNA processing- In eukaryotes the RNA transcript (pre mRNA) is spliced and modified to produce mRNA, which moves from the nucleus to the cytoplasm.

3) The mRNA leaves the nucleus and attaches to a ribosome.

4) Amino acid activation- Each amino acid attaches to its proper tRNA with the help of a specific enzyme and ATP.

5) Translation- A succession of tRNAs add their amino acids to the polypeptide chain as the mRNA is moved through the ribosome one codon at a time. (When completed the polypeptide is released from the ribosome.

Conducting the Genetic Orchestra

- Prokaryotes and eukaryotes alter gene expression in response to their changing environment.

- 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.
Individual bacteria respond to environmental change by regulating their gene expression

- A bacterium can tune its metabolism to the changing environment and food sources
- This metabolic control occurs on two levels:
  - Adjusting activity of metabolic enzymes
  - Regulating genes that encode metabolic enzymes

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

- An operon is the entire stretch of DNA that includes the operator, the promoter, and the genes that they control.
- In bacteria, genes are often clustered into operons, composed of:
  - An operator, an “on-off” switch
  - A promoter
  - Genes for metabolic enzymes
- 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.

- 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.
Eukaryotic gene expression can be regulated at any stage

- All organisms must regulate which genes are expressed at any given time
- In multicellular organisms 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, the expression of different genes by cells with the same genome
- In each type of differentiated cell, a unique subset of genes is expressed
- Many key stages of gene expression can be regulated in eukaryotic cells
- Errors in gene expression can lead to diseases including cancer
- Gene expression is regulated at many stages
- All our cells start off with the same set of genes
- A small percentage of these genes are expressed in all our cells—housekeeping genes like for example for glycolysis
- Through development and our lives different cells selectively express different genes
- For example RBCs transcribe hemoglobin genes whereas the eye does not transcribe this gene and instead it expresses crystalline DNA

Stages in gene expression that can be regulated in eukaryotic cells. In this diagram the colored boxes indicate the processes most often regulated; each color indicates the type of molecule that is affected (blue=DNA, orange=RNA, purple=protein).

The nuclear envelope separating transcription from translation in eukaryotic cells offers an opportunity for post-transcriptional control in the form of RNA processing that is absent in prokaryotes. In addition eukaryotes have a greater variety of control mechanisms operating before transcription and after translation. The expression of any given gene, however does not necessarily involve every stage shown; for example not every polypeptide is cleaved.
Controls before transcription

- Promoters
- Enhancers
- Methylation and acetylation
- Rearrangement- multiplication- for example polytene chromosomes contain several copies of genes allowing more RNA and subsequently more proteins to get produced
Regulation of Chromatin Structure

- Genes within highly packed heterochromatin are usually not expressed
- Chemical modifications to histones and DNA of chromatin influence both chromatin structure and gene expression

Histone Modifications

- In **histone acetylation**, acetyl groups are attached to positively charged lysines in histone tails
- This process loosens chromatin structure, thereby promoting the initiation of transcription
- The addition of methyl groups (methylation) can condense chromatin; the addition of phosphate groups (phosphorylation) next to a methylated amino acid can loosen chromatin
- The histone code hypothesis proposes that specific combinations of modifications help determine chromatin configuration and influence transcription
A simple model of histone tails and the effect of histone acetylation. In addition to acetylation, histones can undergo several other types of modifications that also help determine the chromatin configuration in a region.

**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.
Chemical Modifications

- Methylation of DNA can inactivate genes
- Acetylation of histones allows DNA unpacking and transcription

The Histone Code

Methylation of histone or of DNA usually turns a gene off.
Acetylation of histone usually turns a gene on.
**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**

- 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 control elements, segments of noncoding DNA that help regulate transcription by binding certain proteins.

- Control elements and the proteins they bind are critical to the precise regulation of gene expression in different cell types.
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, groups 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.
- Bound activators cause mediator proteins to interact with proteins at the promoter.
1) Activator proteins bind to distal control elements grouped as an enhancer in the DNA. This enhancer has three binding sites.

2) A DNA bending protein brings the bound activators closer to the promoter. General transcription factors, mediator proteins, and RNA polymerase are nearby.

3) The activators bind to certain mediator proteins and general transcription factors, helping them form an active transcription initiation complex on the promoter.

A model for the action of enhancers and transcription activators: Bending of the DNA by a protein enables enhancers to influence a promoter hundreds or even thousands of nucleotides away. Specific transcription factors called activators bind to the enhancer DNA sequences and then to a group of mediator proteins, which in turn bind to general transcription factors assembling the transcription initiation complex. These protein-protein interactions facilitate the correct positioning of the complex on the promoter and the initiation of RNA synthesis. Only one enhancer (with three orange control elements) is shown here, but a gene may have several enhancers that act at different times or on different cell types.

**Coordinately Controlled Genes in Eukaryotes**

- Unlike the genes of a prokaryotic operon, each of the coordinately controlled 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

Control of RNA processing

- 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
- Alternative splicing- For example different muscle cells express slightly different forms of the troponin gene by utilizing alternative splicing. This generates proteins with somewhat unique functions
- Nuclear envelope- UTRs contain zipcodes which allow the RNA to exit the nucleus with the aid of proteins which recognize it and selectively bind to it. In addition these unique zipcodes specify to which cytoplasmic location these RNA must move. This is crucial during embryonic development.
Alternative RNA splicing of the troponin T gene. The primary transcript of this gene can be spliced in more than one way, generating different mRNA molecules. Notice that one mRNA molecule has ended up with exon 3 (green) and the other with exon 4 (purple). These two mRNAs are translated into different but related muscle proteins.

**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
- The mRNA life span is determined in part by sequences in the leader and trailer regions
**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.

- Alternatively, translation of all mRNAs in a cell may be regulated simultaneously.

- For example, translation initiation factors are simultaneously activated in an egg following fertilization.

**Protein Processing and Degradation**

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

- **Proteasomes** are giant protein complexes that bind protein molecules and degrade them.
Degradation of a protein by a proteasome. A proteasome, an enormous protein complex shaped like a trash can, chops up unneeded proteins in the cell. In most cases the proteins attacked by a proteasome have been tagged with short chains of ubiquitin a small protein. Steps 1 and 3 require ATP. Eukaryotic proteasomes are as massive as ribosomal subunits and are distributed throughout the cell. Their shape somewhat resembles that of chaperone proteins, which protect protein structure rather than destroy it.

Control after translation (post-translational)

- Example: phosphorylation and other protein modifications which occur after the protein has been synthesized can change their activity
Noncoding RNAs play multiple roles in controlling gene expression

- Only a small fraction of DNA codes for proteins, rRNA, and tRNA
- A significant amount of the genome may be transcribed into noncoding RNAs
- 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
The phenomenon of inhibition of gene expression by RNA molecules is called **RNA interference (RNAi)**.

This is caused by single-stranded small interfering RNAs (siRNAs) and can lead to degradation of an mRNA or block its translation.

siRNAs and miRNAs are similar but form from different RNA precursors.
Chromatin Remodeling and Silencing of Transcription by Small RNAs

- siRNAs play a role in heterochromatin formation and can block large regions of the chromosome
- Small RNAs may also block transcription of specific genes