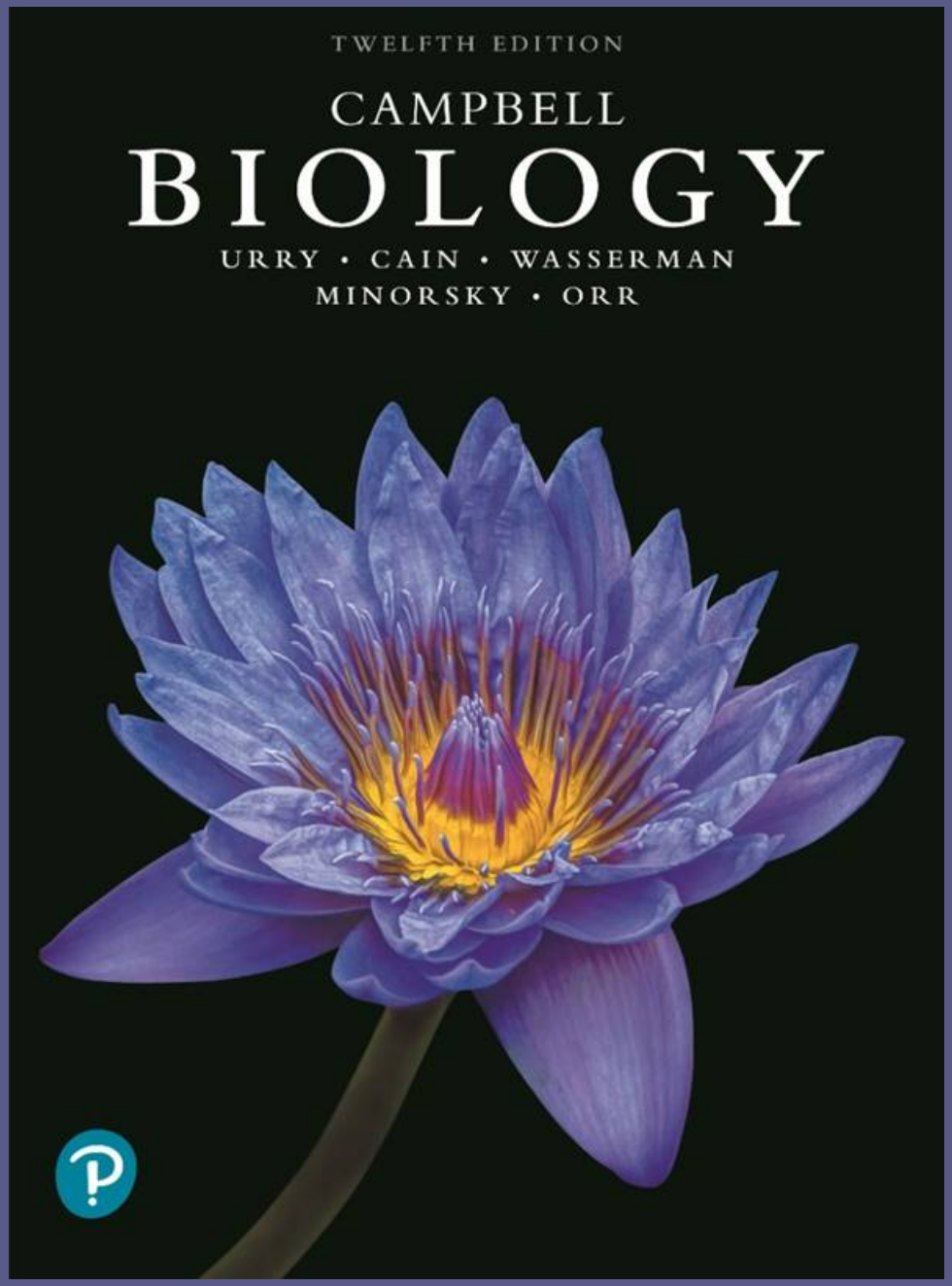


Chapter 17

Gene Expression: From Gene to Protein

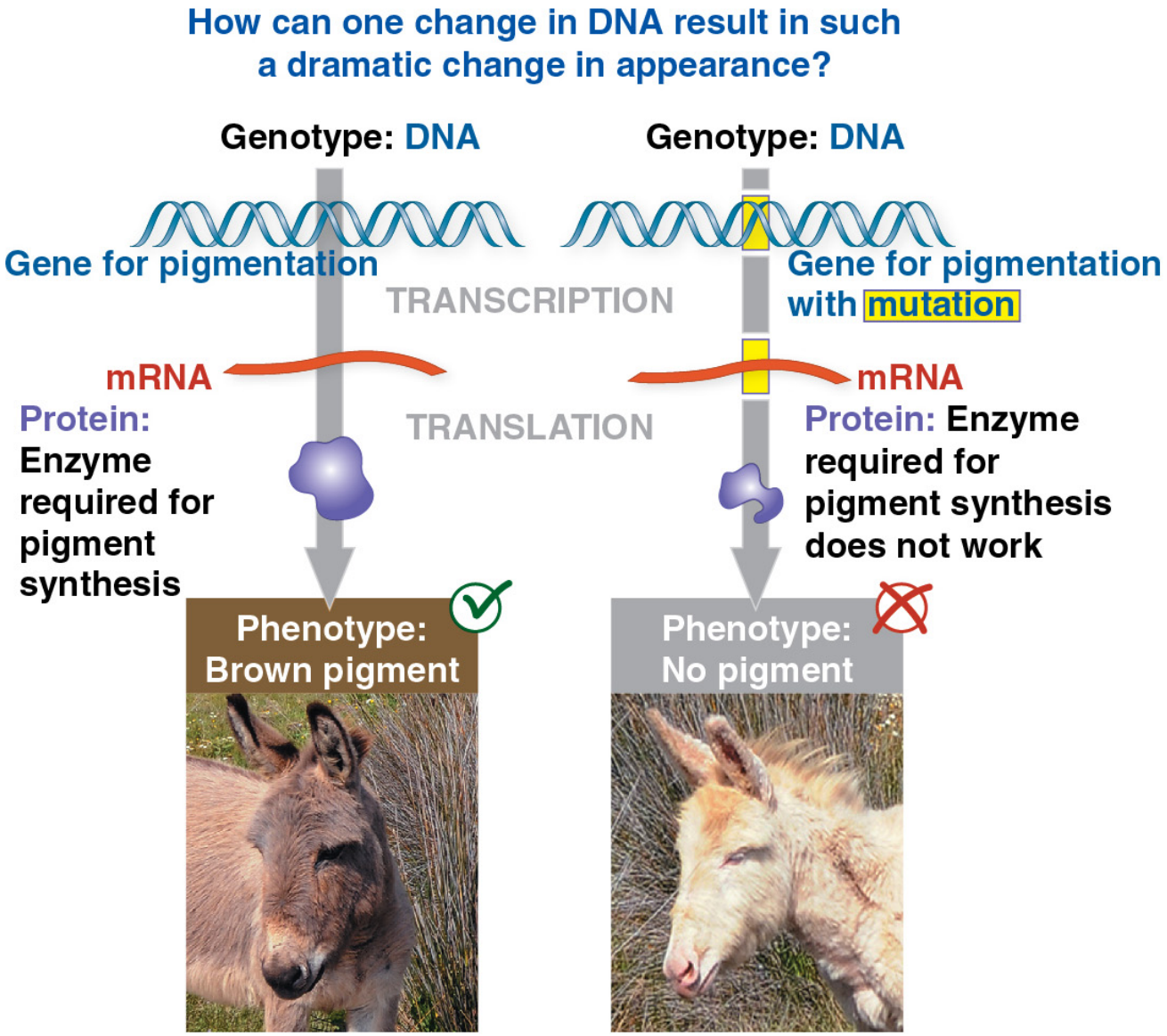
Lecture Presentations by
Nicole Tunbridge and
Kathleen Fitzpatrick

Figure 17.1a

Figure 17.1b

Gene expression is the process by which DNA directs the synthesis of proteins.

Proteins are the link between genotype and phenotype.



CONCEPT 17.1: Genes specify proteins via transcription and translation

- 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
- How was the relationship between proteins and DNA discovered?

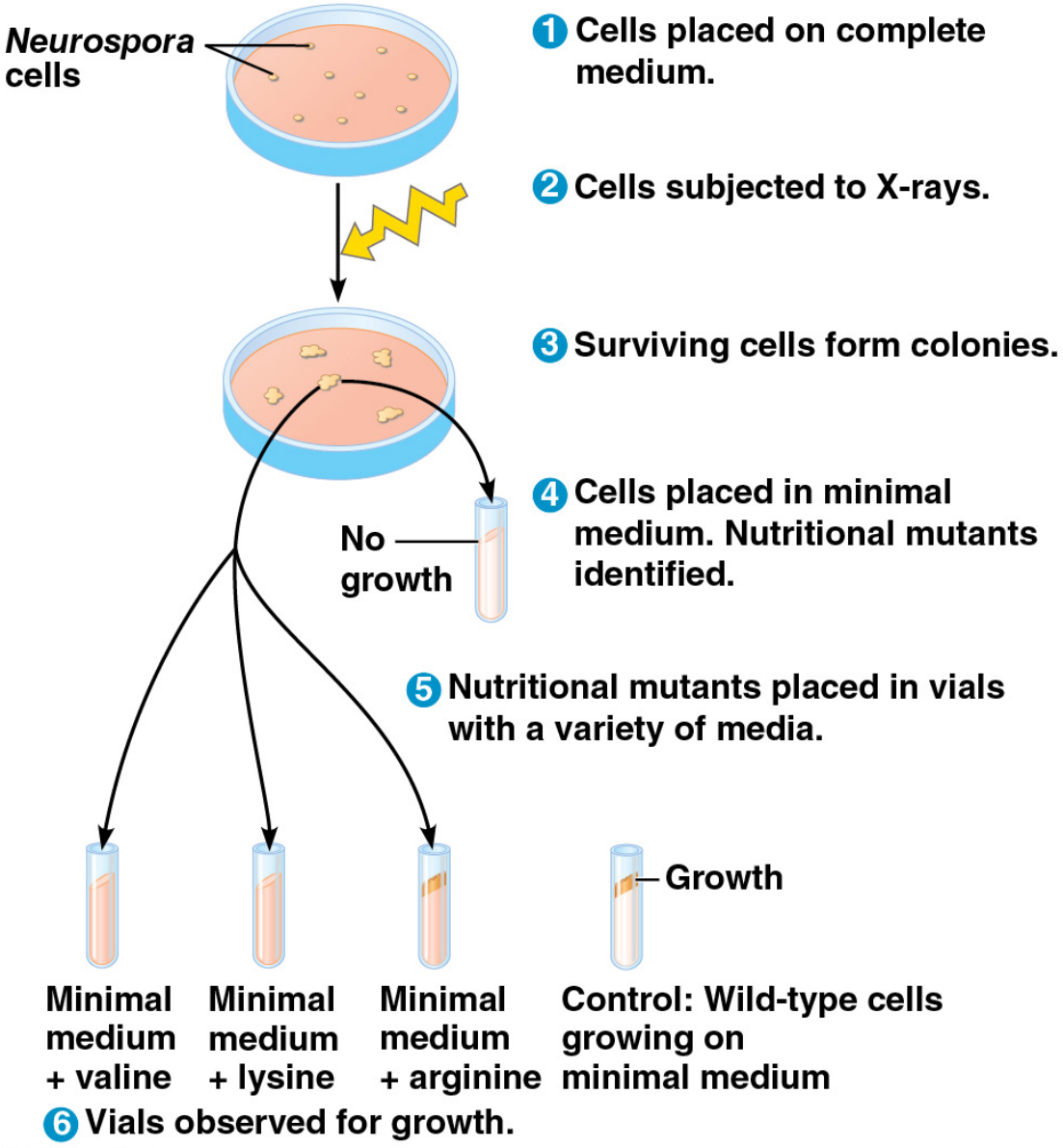
Evidence from the Study of Metabolic Defects

- In 1902, British physician Archibald Garrod first suggested that genes dictate phenotypes through enzymes that catalyze specific chemical reactions
- He thought symptoms of an inherited disease reflect an inability to synthesize a certain enzyme
- Cells synthesize and degrade molecules in a series of steps, a metabolic pathway

Nutritional Mutants in Neurospora: Scientific Inquiry

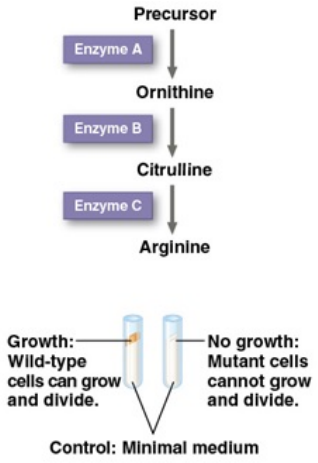
- George Beadle and Edward Tatum exposed bread mold to X-rays, creating mutants that were unable to survive on minimal media
- Their colleagues Adrian Srb and Norman Horowitz identified three classes of arginine-deficient mutants
- Each lacked a different enzyme necessary for synthesizing arginine

Figure 17.2


Neurospora cells
1 Cells placed on complete medium.
2 Cells subjected to X-rays.
3 Surviving cells form colonies.
4 Cells placed in minimal medium. Nutritional mutants identified.
No growth
5 Nutritional mutants placed in vials with a variety of media.
Growth
Minimal medium + valine
Minimal medium + lysine
Minimal medium + arginine
Control: Wild-type cells growing on minimal medium
6 Vials observed for growth.

- The results of the experiments provided support for the one gene–one enzyme hypothesis
- The hypothesis states that the function of a gene is to dictate production of a specific enzyme

Figure 17.3



Results Table — Classes of *Neurospora crassa*

Condition	Wild type	Class I mutants	Class II mutants	Class III mutants
Minimal medium (MM) (control)				
MM + ornithine				
MM + citrulline				
MM + arginine (control)				
Summary of results	Can grow with or without any supplements	Can grow on ornithine, citrulline, or arginine	Can grow only on citrulline or arginine	Require arginine to grow

Gene (codes for enzyme)	Wild type	Class I mutants (mutation in gene *A*)	Class II mutants (mutation in gene *B*)	Class III mutants (mutation in gene *C*)
Gene *A* → Enzyme A	Precursor	Precursor	Precursor	Precursor
Gene *B* → Enzyme B	Ornithine	Ornithine	Ornithine	Ornithine
Gene *C* → Enzyme C	Citrulline	Citrulline	Citrulline	Citrulline
	Arginine	Arginine	Arginine	Arginine

Data from A. M. Srb and N. H. Horowitz, The ornithine cycle in *Neurospora* and its genetic control, *Journal of Biological Chemistry* 154:129–139 (1944).

The Products of Gene Expression: A Developing Story

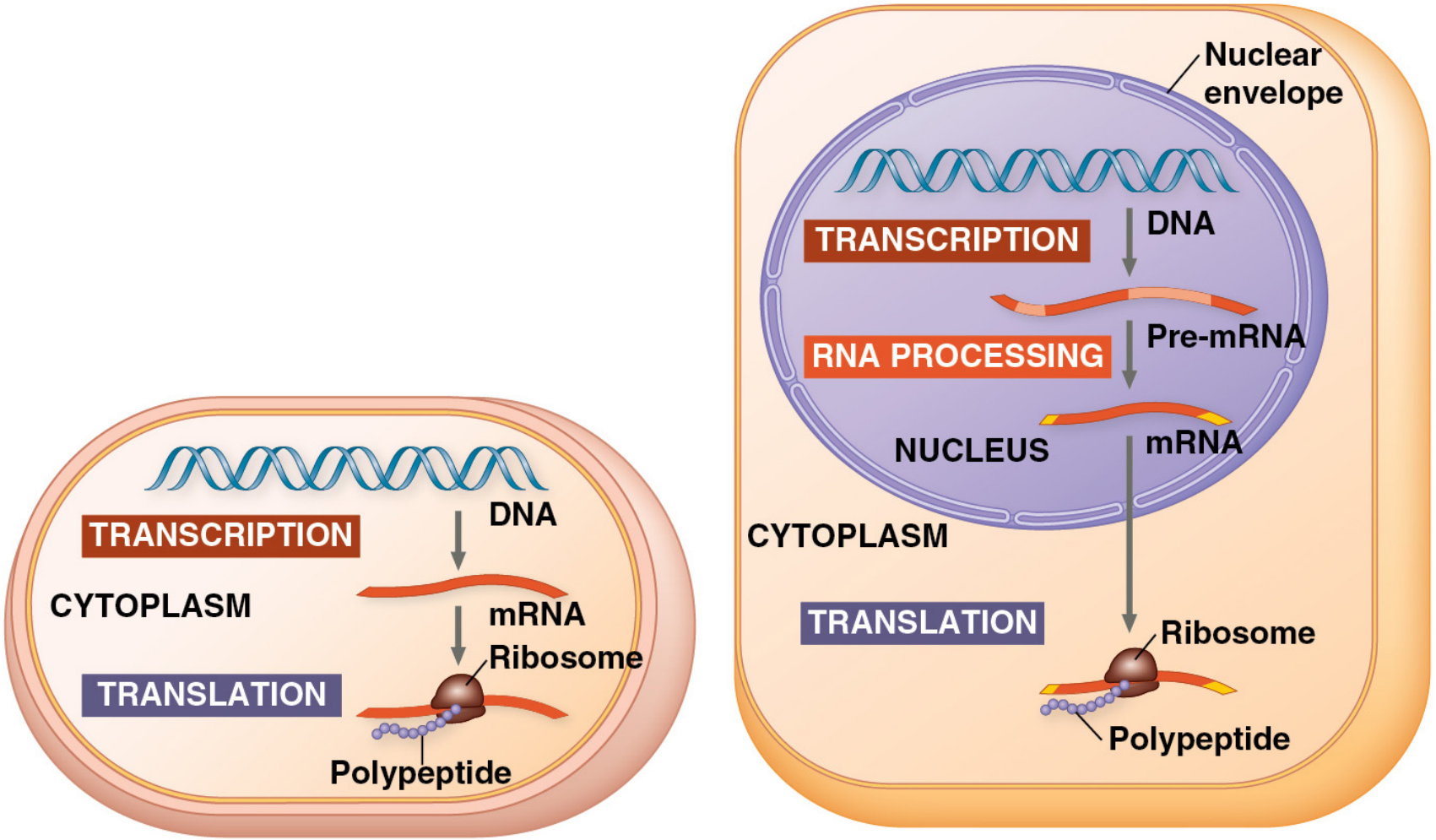
- Not all proteins are enzymes, so researchers later revised the hypothesis: one gene–one protein
- Many proteins are composed of several polypeptides, each of which has its own gene
- Therefore, Beadle and Tatum's hypothesis is now restated as the one gene–one polypeptide hypothesis
- It is common to refer to gene products as proteins rather than more precisely as polypeptides

Basic Principles of Transcription and Translation

- RNA is the bridge between genes and protein synthesis
- **Transcription** is the synthesis of RNA using information in DNA
- Transcription produces **messenger RNA (mRNA)**
- **Translation** is the synthesis of a polypeptide, using information in the mRNA
- **Ribosomes** are the sites of translation

- In prokaryotes, translation of mRNA can begin before transcription has finished
- In a eukaryotic cell, the nuclear envelope separates transcription from translation
- Eukaryotic RNA transcripts are modified through RNA processing to yield the finished mRNA

Figure 17.4


TRANSCRIPTION
DNA
CYTOPLASM
mRNA
Ribosome
TRANSLATION
Polypeptide
(a) Bacterial cell
Nuclear envelope
TRANSCRIPTION
DNA
RNA PROCESSING
Pre-mRNA
NUCLEUS
mRNA
CYTOPLASM
TRANSLATION
Ribosome
Polypeptide
(b) Eukaryotic cell

- A **primary transcript** is the initial RNA transcript from any gene prior to processing
- The central dogma is the concept that cells are governed by a cellular chain of command: DNA → RNA → protein

Figure 17.UN01

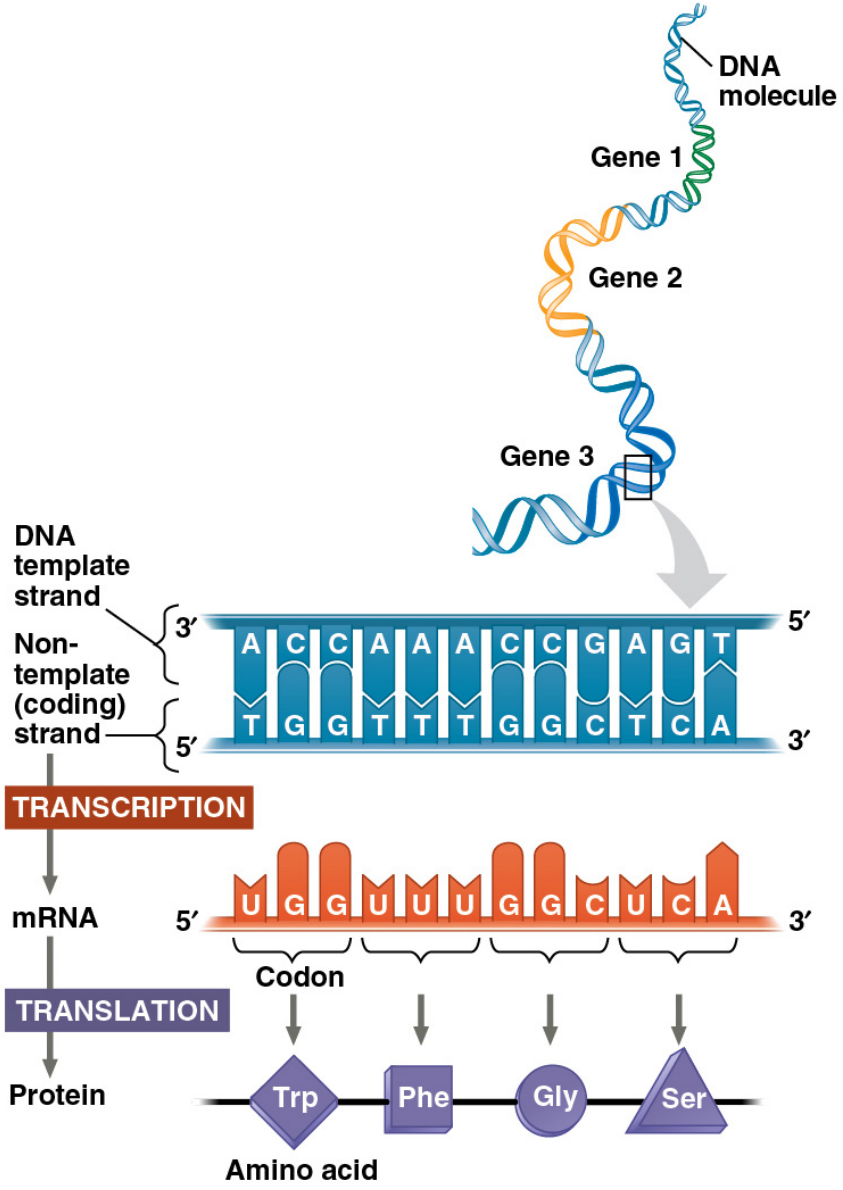
The Genetic Code

- 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 nucleotides correspond to an amino acid?

Codons: Triplets of Nucleotides

- The flow of information from gene to protein is based on a **triplet code**: a series of nonoverlapping, three-nucleotide words
- The words of a gene are transcribed into complementary nonoverlapping three-nucleotide words of mRNA
- These words are then translated into a chain of amino acids, forming a polypeptide

Figure 17.5


DNA molecule
Gene 1
Gene 2
Gene 3
DNA template strand
Non-template (coding) strand
3′
5′
A C C A A A C C G A G T
T G G T T T G G C T C A
5′
3′
TRANSCRIPTION
mRNA
5′
U G G U U U G G C U C A
3′
Codon
TRANSLATION
Protein
Trp
Phe
Gly
Ser
Amino acid

- One of the two DNA strands, the **template strand**, provides a template for ordering the sequence of complementary nucleotides in an RNA transcript
- The template strand is always the same strand for a given gene
- However, further along the chromosome, the opposite strand may be the template strand for a different gene

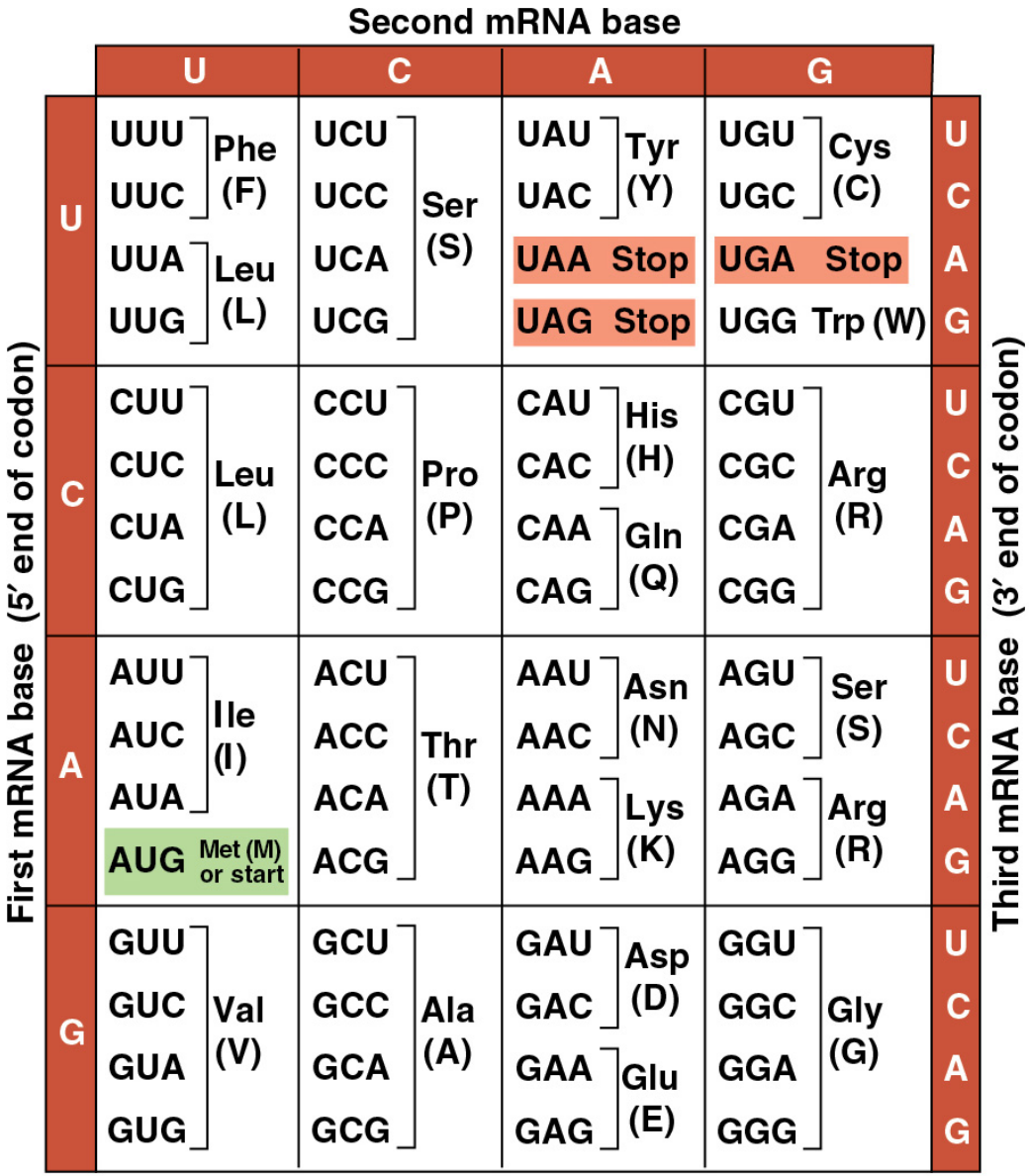
- Specific DNA sequences associated with the gene direct which strand is used as the template
- The mRNA molecule produced is complementary to the template strand
- During translation, the mRNA base triplets, called **codons**, are read in the 5′ → 3′ direction

- The nontemplate strand is called the **coding strand** because the nucleotides of this strand are identical to the codons, except that T is present in the DNA in place of U in the RNA
- Each codon specifies the amino acid (one of 20) to be placed at the corresponding position along a polypeptide

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 (more than one codon may specify a particular amino acid) 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

Figure 17.6


Second mRNA base
U
C
A
G
First mRNA base (5′ end of codon)
Third mRNA base (3′ end of codon)
U
C
A
G
UUU Phe (F)
UUC
UUA Leu (L)
UUG
UCU Ser (S)
UCC
UCA
UCG
UAU Tyr (Y)
UAC
UAA Stop
UAG Stop
UGU Cys (C)
UGC
UGA Stop
UGG Trp (W)
CUU Leu (L)
CUC
CUA
CUG
CCU Pro (P)
CCC
CCA
CCG
CAU His (H)
CAC
CAA Gln (Q)
CAG
CGU Arg (R)
CGC
CGA
CGG
AUU Ile (I)
AUC
AUA
AUG Met (M) or start
ACU Thr (T)
ACC
ACA
ACG
AAU Asn (N)
AAC
AAA Lys (K)
AAG
AGU Ser (S)
AGC
AGA Arg (R)
AGG
GUU Val (V)
GUC
GUA
GUG
GCU Ala (A)
GCC
GCA
GCG
GAU Asp (D)
GAC
GAA Glu (E)
GAG
GGU Gly (G)
GGC
GGA
GGG

Evolution of the Genetic Code

- The genetic code is nearly universal, shared by the simplest bacteria and the most complex animals
- Genes can be transcribed and translated after being transplanted from one species to another
- A language shared by all living things must have been operating very early in the history of life

Figure 17.7

(a) Tobacco plant expressing a firefly gene

(b) Mosquito larva expressing a jellyfish gene

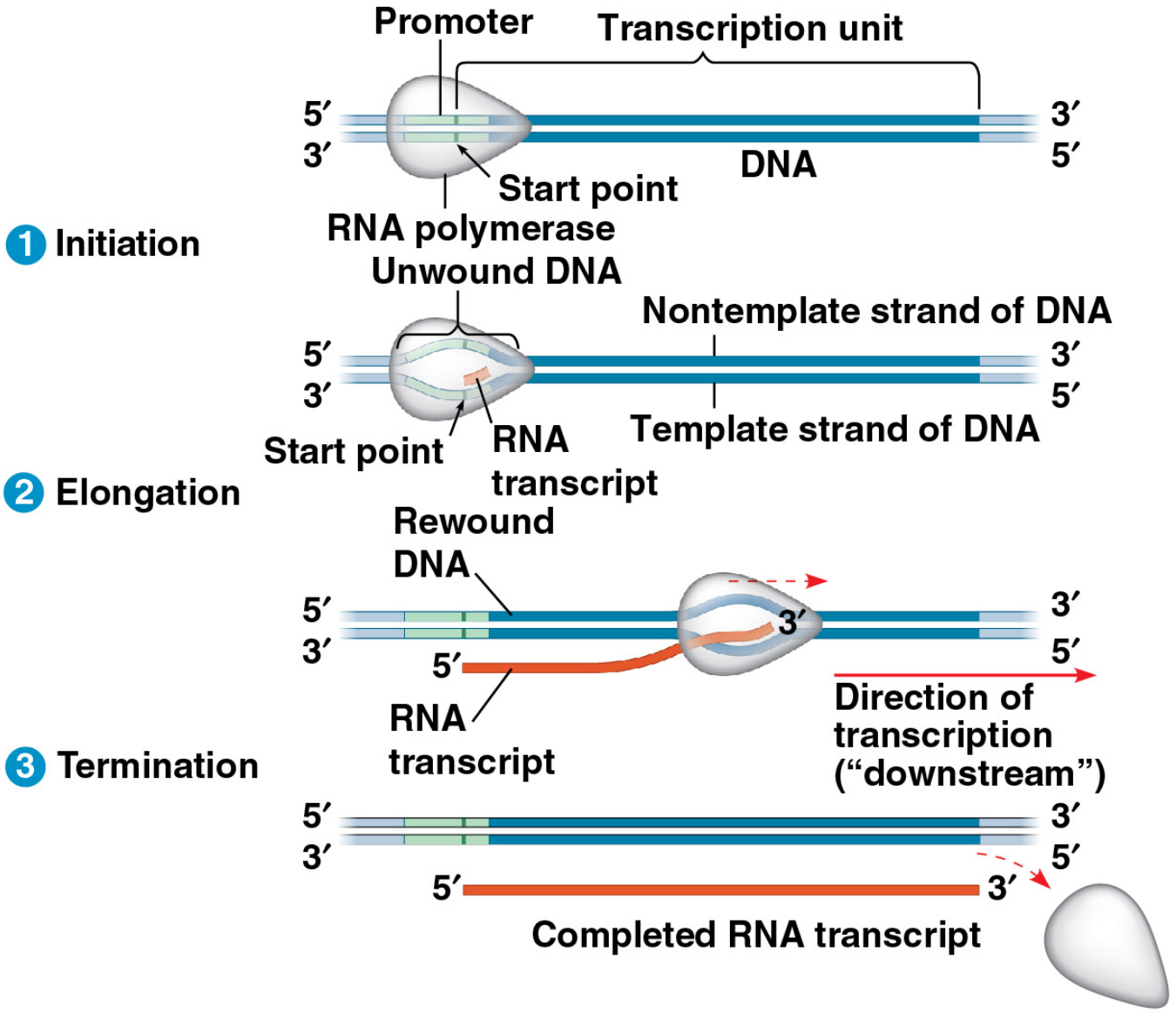
CONCEPT 17.2: Transcription is the DNA-directed synthesis of RNA: *a closer look*

- Transcription is the first stage of gene expression

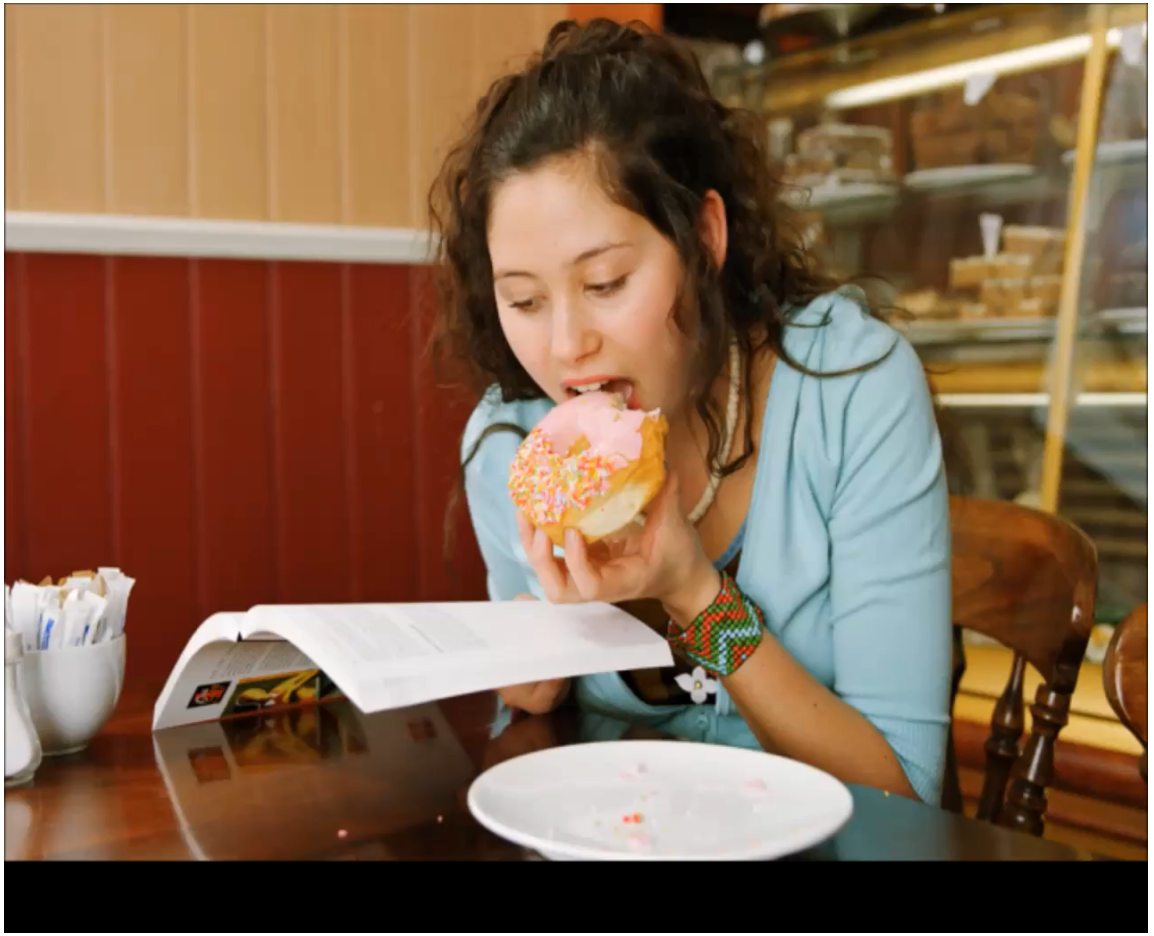
Molecular Components of Transcription

- RNA synthesis is catalyzed by **RNA polymerase**, which pries the DNA strands apart and joins together the RNA nucleotides
- The RNA is complementary to the DNA template strand
- RNA polymerase does not need any primer
- RNA synthesis follows the same base-pairing rules as DNA, except that uracil substitutes for thymine

Figure 17.8


Promoter
Transcription unit
5′
3′
3′
5′
DNA
Start point
RNA polymerase
1 Initiation
Unwound DNA
Nontemplate strand of DNA
5′
3′
3′
5′
Start point
RNA transcript
Template strand of DNA
2 Elongation
Rewound DNA
5′
3′
3′
5′
3′
5′
RNA transcript
Direction of transcription ("downstream")
3 Termination
5′
3′
3′
5′
5′
3′
Completed RNA transcript

BioFlix® Animation: Transcription



- 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**

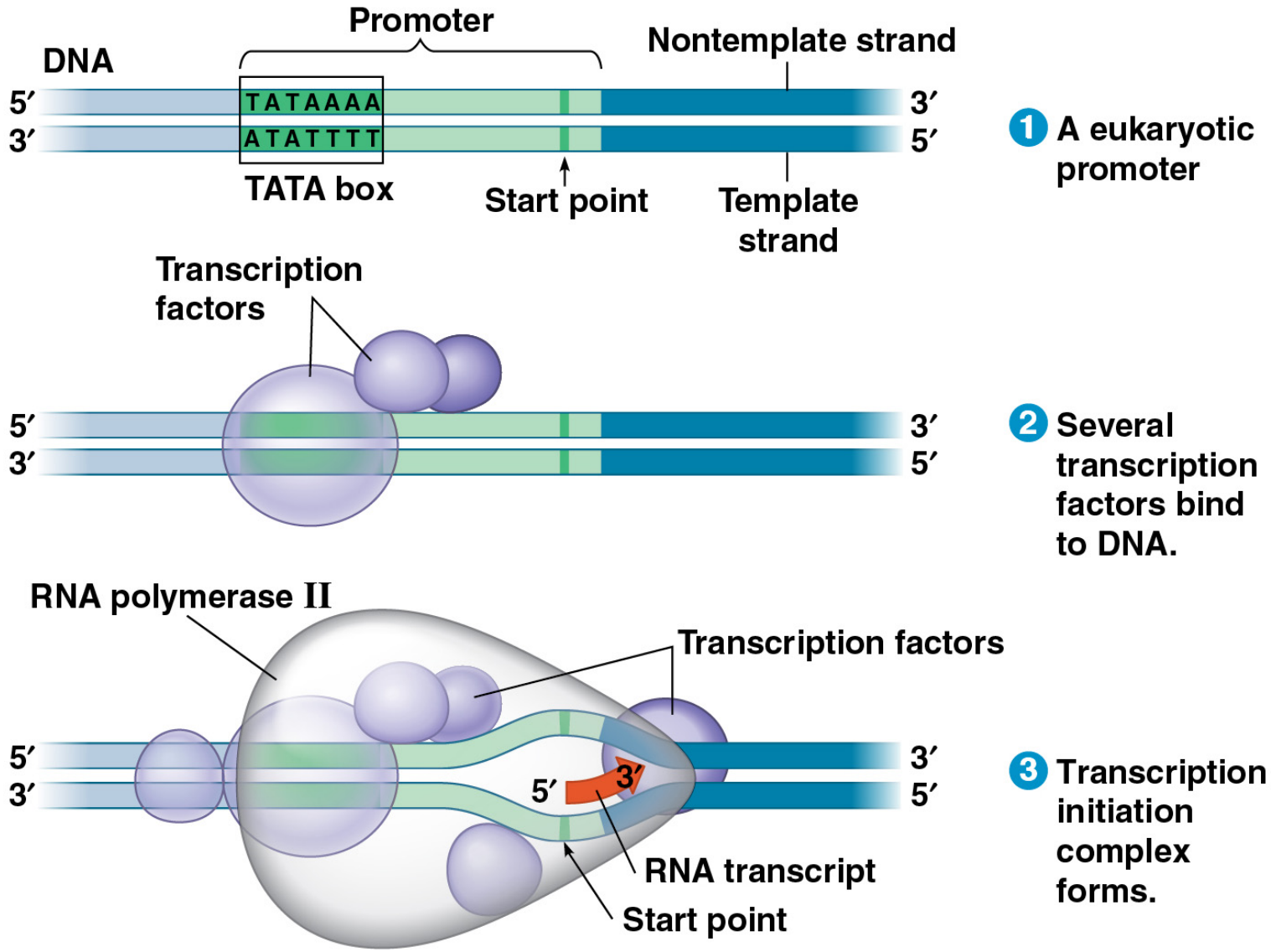
Synthesis of an RNA Transcript

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

RNA Polymerase Binding and Initiation of Transcription

- Promoters signal the transcription **start point** and usually extend several dozen nucleotide pairs upstream of the start point
- **Transcription factors** help guide 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

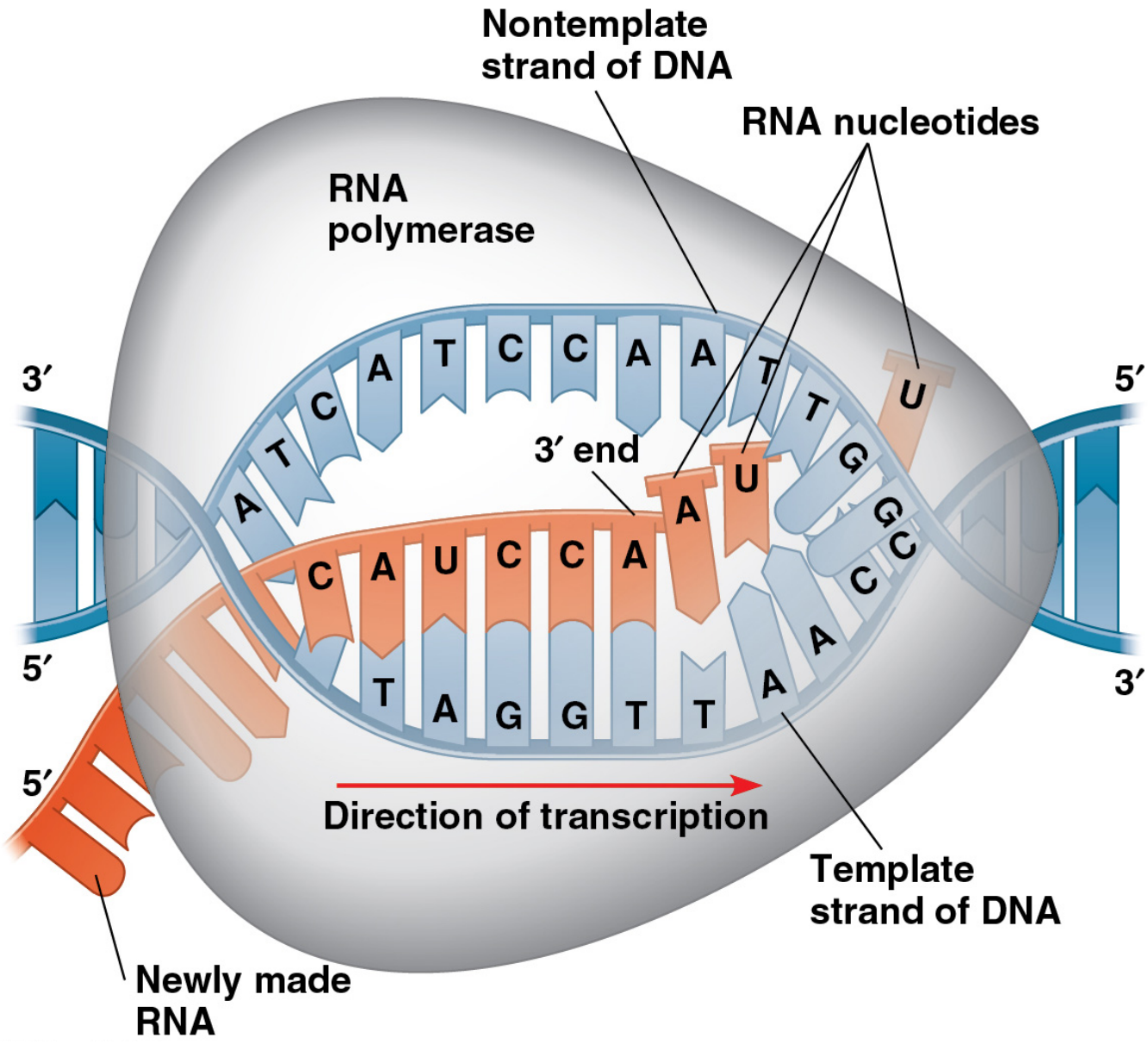
Figure 17.9


Promoter
DNA
Nontemplate strand
5′
TATAAAA
3′
3′
ATATTTT
5′
TATA box
Start point
Template strand
1 A eukaryotic promoter
Transcription factors
5′
3′
3′
5′
2 Several transcription factors bind to DNA.
RNA polymerase II
Transcription factors
5′
3′
3′
5′
5′
3′
RNA transcript
Start point
3 Transcription initiation complex forms.

Elongation of the RNA Strand

- As RNA polymerase moves along the DNA, it untwists the double helix, 10–20 nucleotides at a time
- Nucleotides are added to the 3′ end of the growing RNA molecule
- Transcription progresses at a rate of 40 nucleotides per second in eukaryotes
- A gene can be transcribed simultaneously by several RNA polymerases

Figure 17.10


Nontemplate strand of DNA
RNA nucleotides
RNA polymerase
3′
5′
3′ end
5′
3′
5′
Direction of transcription
Template strand of DNA
Newly made RNA

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 and the mRNA can be translated without further modification
- In eukaryotes, RNA polymerase II transcribes the polyadenylation signal sequence; the RNA transcript is released 10–35 nucleotides past this polyadenylation sequence

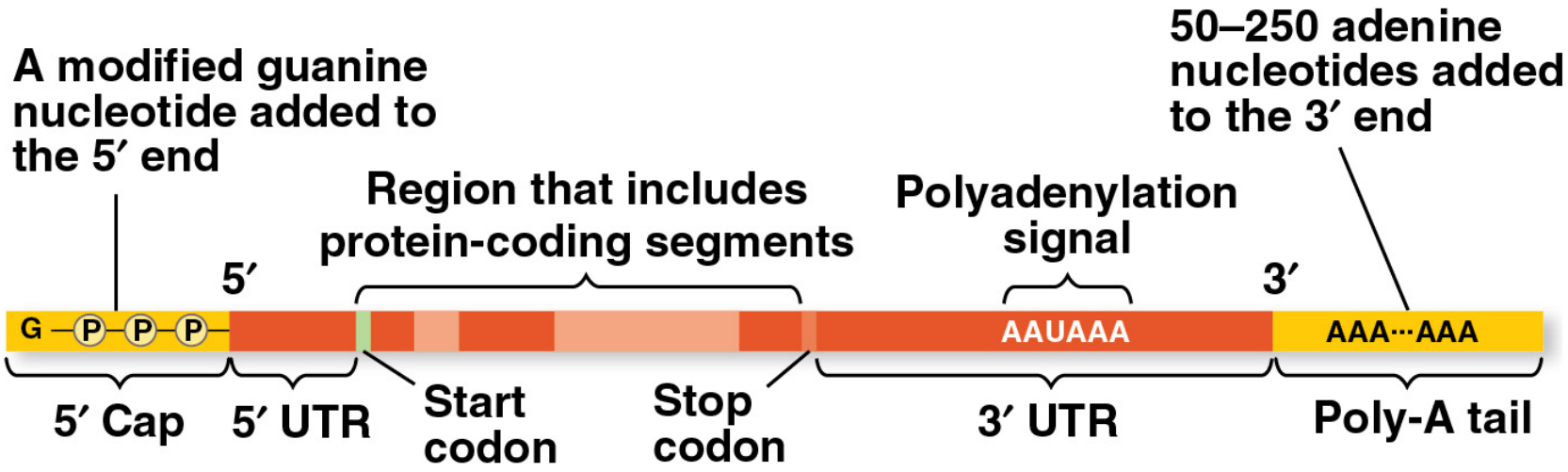
CONCEPT 17.3: Eukaryotic cells modify RNA after transcription

- Enzymes in the eukaryotic nucleus modify pre-mRNA (**RNA processing**) before the genetic messages are dispatched to the cytoplasm
- During RNA processing, both ends of the primary transcript are altered
- Also, in most cases, certain interior sections of the molecule are cut out and the remaining 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 to the cytoplasm
  - They protect mRNA from hydrolytic enzymes
  - They help ribosomes attach to the 5′ end

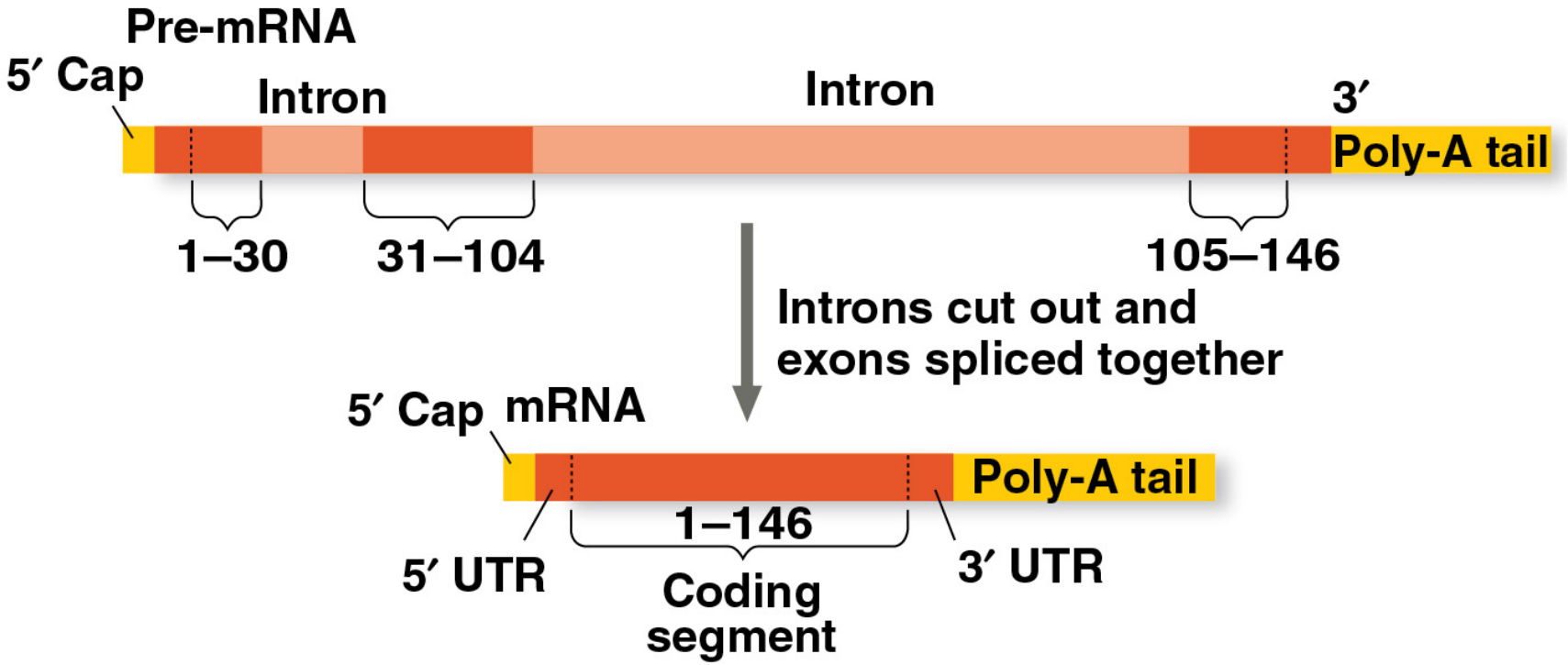
Figure 17.11


A modified guanine nucleotide added to the 5′ end
50–250 adenine nucleotides added to the 3′ end
Region that includes protein-coding segments
Polyadenylation signal
5′
3′
G
P
P
P
AAUAAA
AAA···AAA
5′ Cap
5′ UTR
Start codon
Stop codon
3′ UTR
Poly-A tail

Split Genes and RNA Splicing

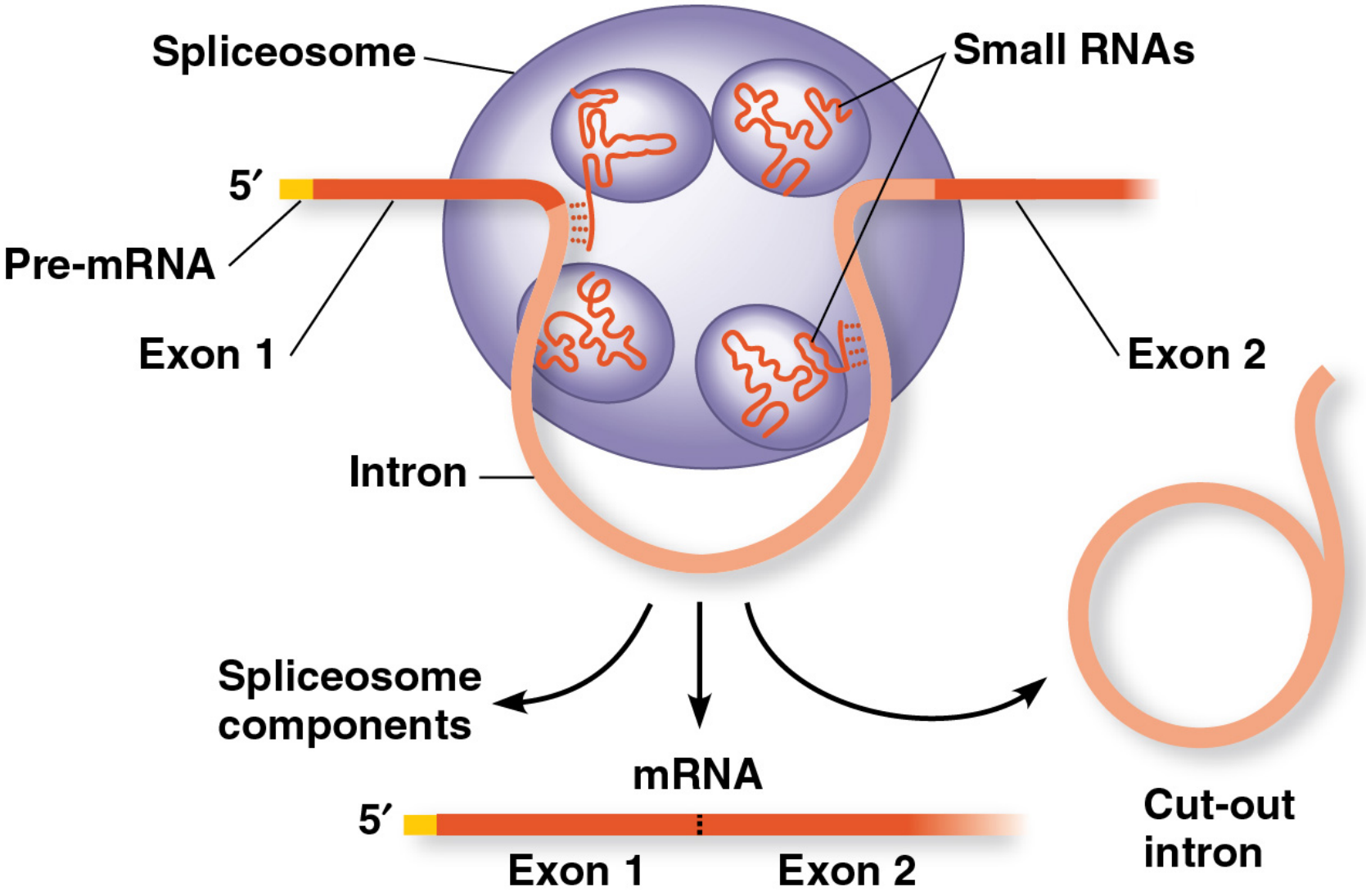
- Most eukaryotic genes and their RNA transcripts have long noncoding stretches of nucleotides that lie between coding regions
- These are removed through **RNA splicing**
- The noncoding segments in a gene are called intervening sequences, or **introns**
- The other regions are called **exons** because they are eventually expressed, usually translated into amino acid sequences

Figure 17.12


Pre-mRNA
5′ Cap
Intron
Intron
3′
Poly-A tail
1–30
31–104
105–146
Introns cut out and
exons spliced together
5′ Cap
mRNA
Poly-A tail
1–146
5′ UTR
3′ UTR
Coding
segment

- The removal of introns is accomplished by spliceosomes
- **Spliceosomes** consist of a variety of proteins and several small RNAs that recognize the splice sites
- The RNAs of the spliceosome also catalyze the splicing reaction

Figure 17.13


Spliceosome
Small RNAs
5′
Pre-mRNA
Exon 1
Exon 2
Intron
Spliceosome components
mRNA
5′
Exon 1
Exon 2
Cut-out intron

BioFlix® Animation: Overview of Transcription

Ribozymes

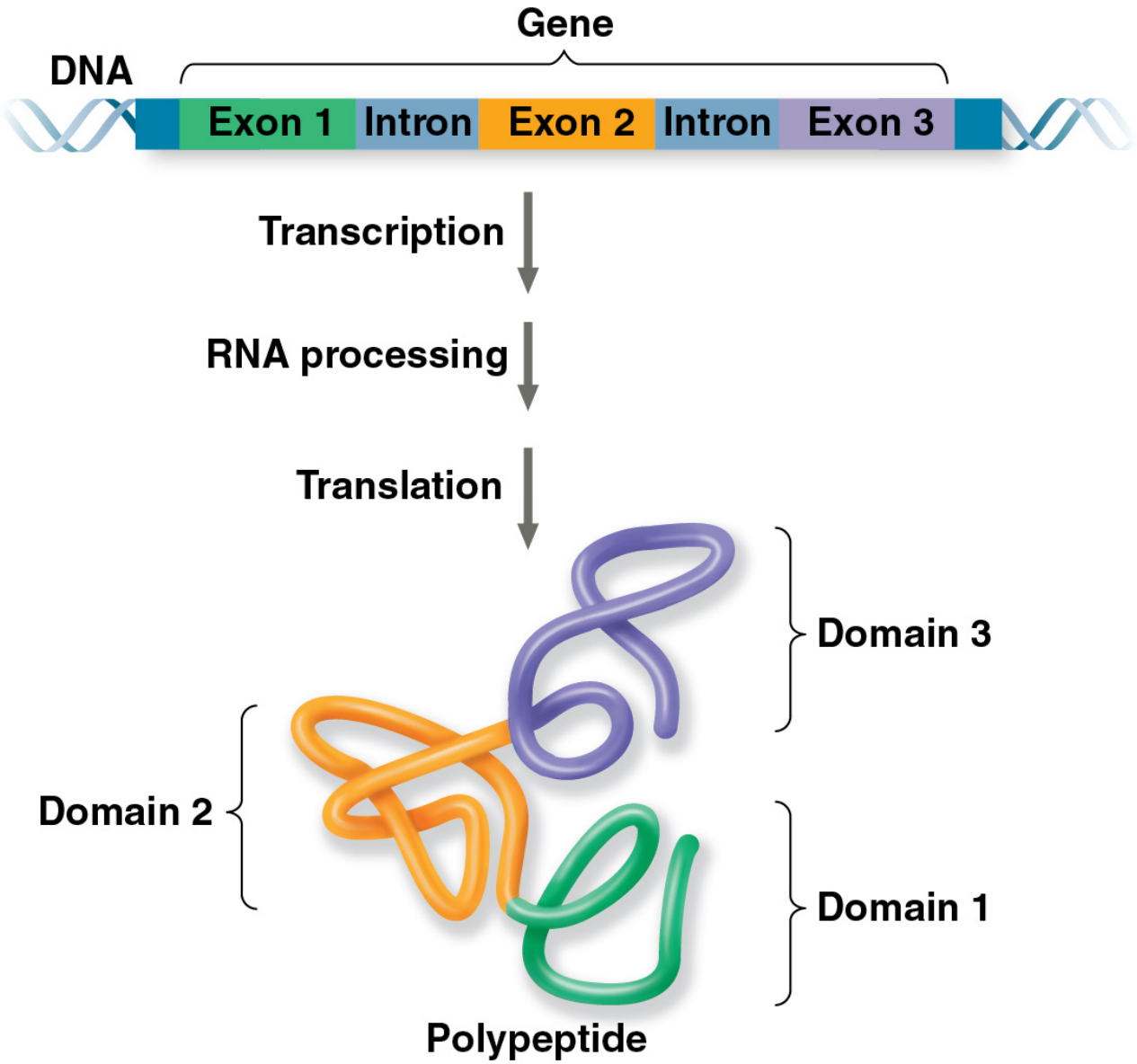
- **Ribozymes** are catalytic RNA molecules that function as enzymes and can splice RNA
- Three properties of RNA enable it to function as an enzyme
  - It can form a three-dimensional structure because of its ability to base-pair with itself
  - Some bases in RNA contain functional groups that may participate in catalysis
  - RNA may hydrogen-bond with other nucleic acid molecules

The Functional and Evolutionary Importance of Introns

- Some introns contain sequences that regulate gene expression and many affect gene products
- Some genes can encode more than one kind of polypeptide, depending on which segments are treated as exons during splicing
- This is called **alternative RNA splicing**
- Consequently, 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 by mixing and matching exons between different genes

Figure 17.14


Gene
DNA
Exon 1
Intron
Exon 2
Intron
Exon 3
Transcription
RNA processing
Translation
Domain 3
Domain 2
Domain 1
Polypeptide

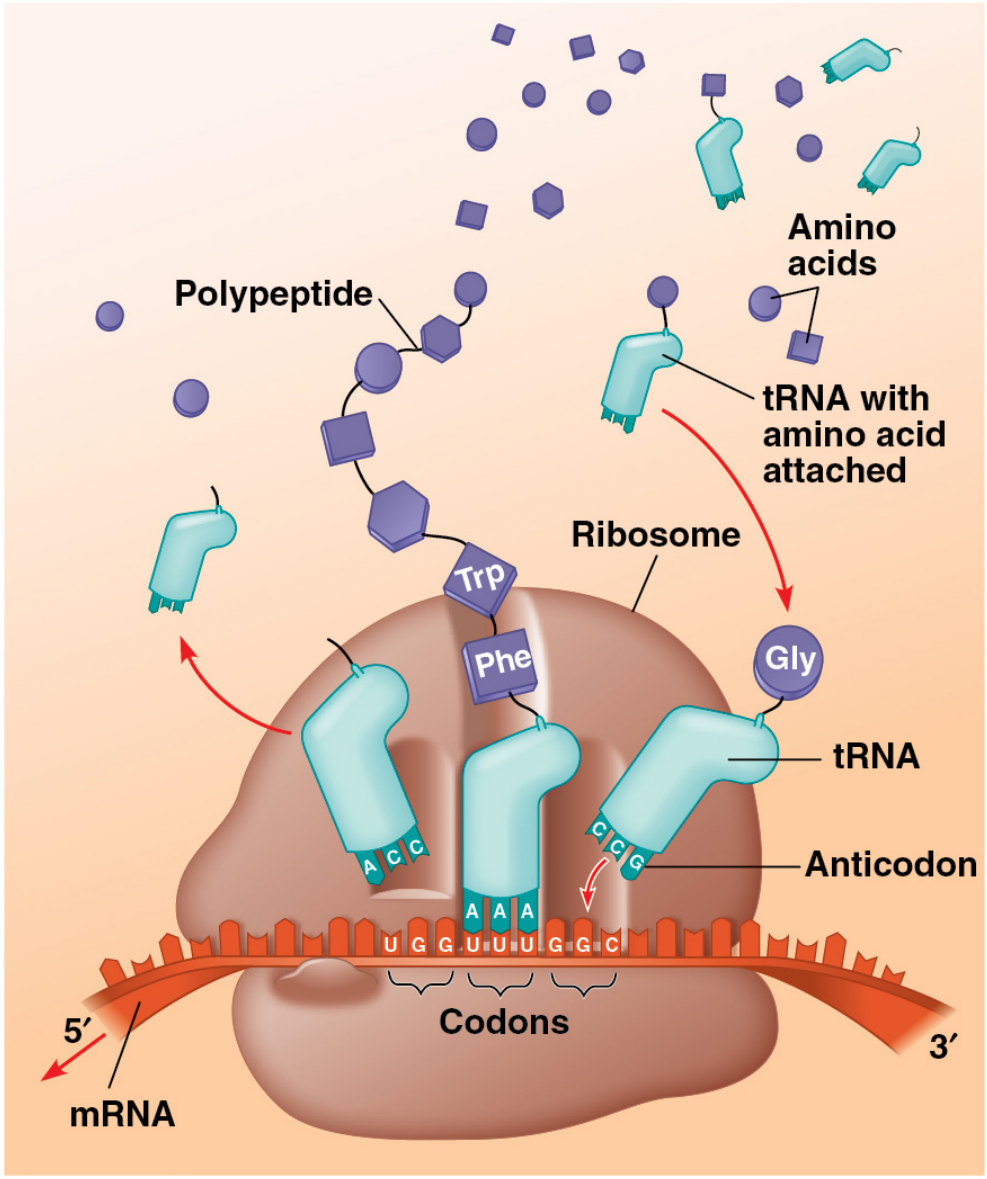
CONCEPT 17.4: Translation is the RNA-directed synthesis of a polypeptide: *a closer look*

- Genetic information flows from mRNA to protein through the process of translation

Molecular Components of Translation

- A cell translates an mRNA message into protein with the help of **transfer RNA (tRNA)**
- tRNAs transfer amino acids to the growing polypeptide in a ribosome
- Translation is a complex process in terms of its biochemistry and mechanics

Figure 17.15


Amino acids
Polypeptide
tRNA with amino acid attached
Ribosome
Trp
Phe
Gly
tRNA
Anticodon
A C C
C C G
A A A
U G G U U U G G C
Codons
5′
3′
mRNA

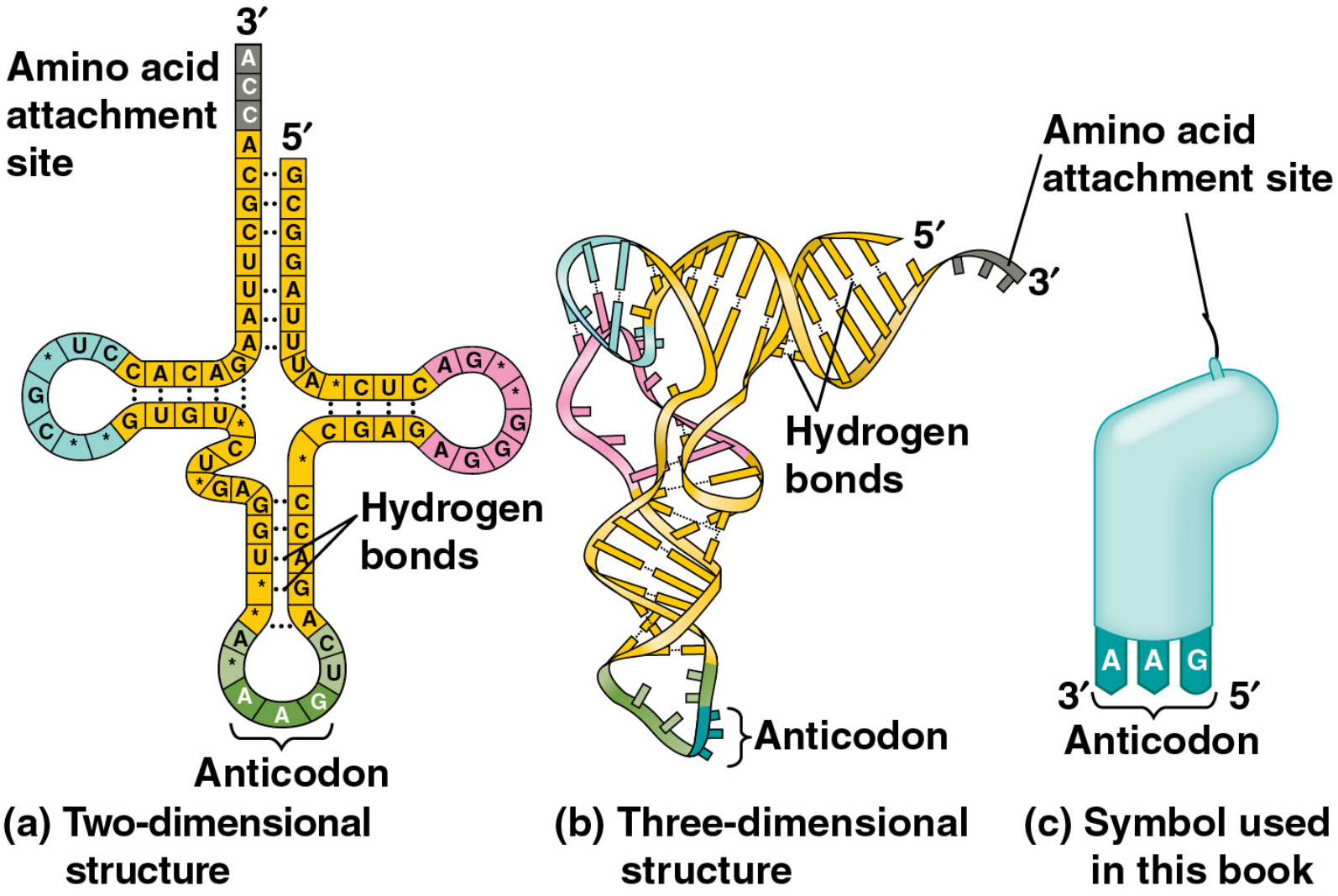
The Structure and Function of Transfer RNA

- Each tRNA molecule enables translation of a given mRNA codon into a certain amino acid
  - 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

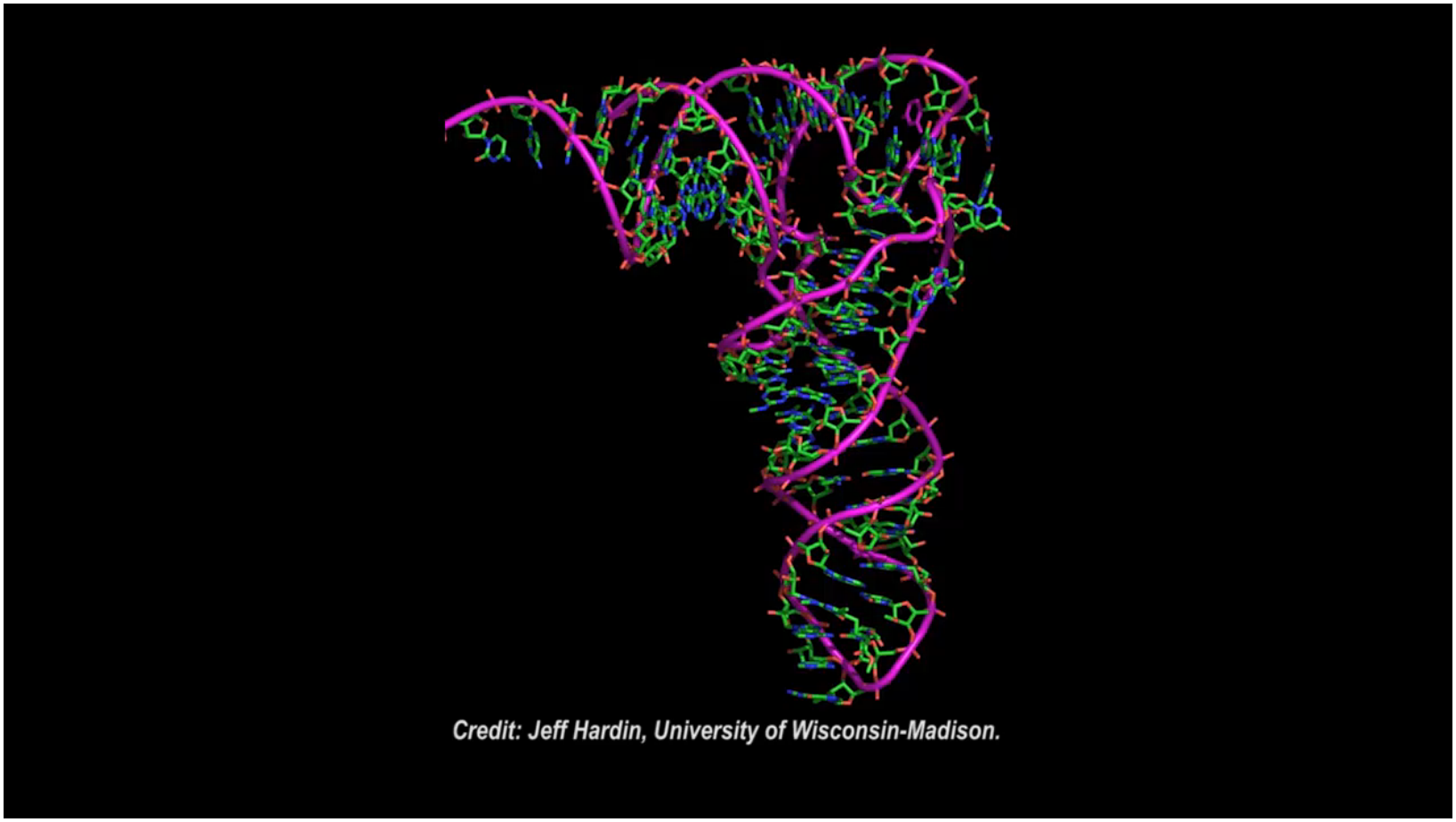
- A tRNA molecule consists of a single RNA strand that is only about 80 nucleotides long
- Flattened into one plane to reveal its base pairing, a tRNA molecule looks like a cloverleaf

- Because of hydrogen bonds, tRNA actually twists and folds into a three-dimensional molecule
- tRNA is roughly L-shaped with the 5′ and 3′ ends both located near one end of the structure
- The protruding 3′ end acts as an attachment site for an amino acid

Figure 17.16

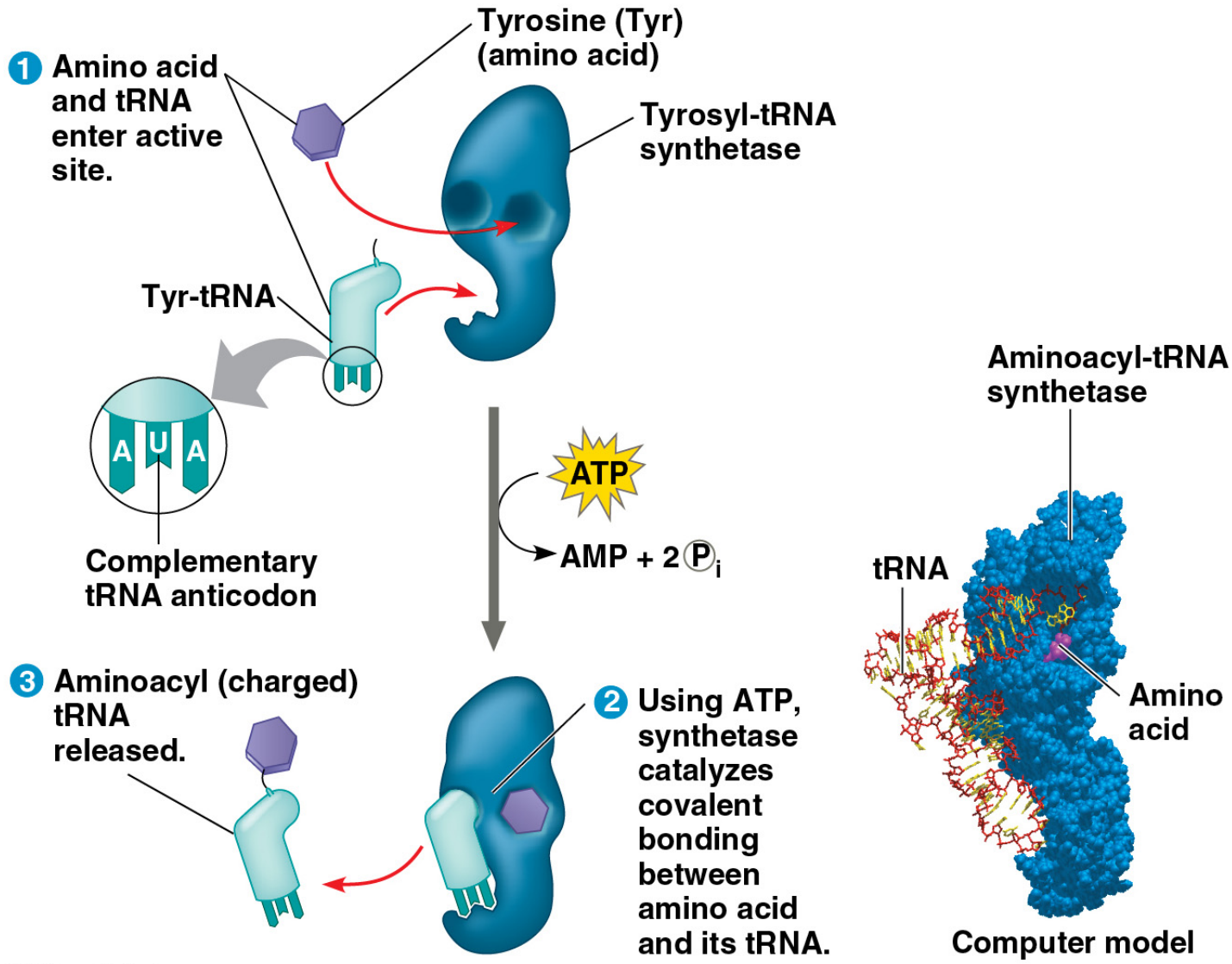

3′
Amino acid attachment site
5′
Hydrogen bonds
Anticodon
(a) Two-dimensional structure
5′
3′
Amino acid attachment site
Hydrogen bonds
Anticodon
(b) Three-dimensional structure
A A G
3′
5′
Anticodon
(c) Symbol used in this book

Video: Stick and Ribbon Rendering of a tRNA



- Accurate translation requires two instances of molecular recognition
  - 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

Figure 17.17

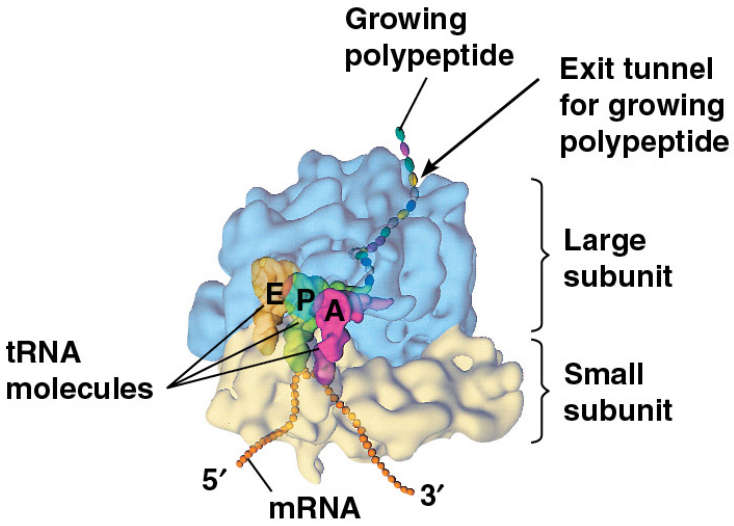

1 Amino acid and tRNA enter active site.
Tyrosine (Tyr) (amino acid)
Tyrosyl-tRNA synthetase
Tyr-tRNA
A U A
Complementary tRNA anticodon
ATP
AMP + 2 Pi
2 Using ATP, synthetase catalyzes covalent bonding between amino acid and its tRNA.
3 Aminoacyl (charged) tRNA released.
Aminoacyl-tRNA synthetase
tRNA
Amino acid
Computer model

The Structure and Function of Ribosomes

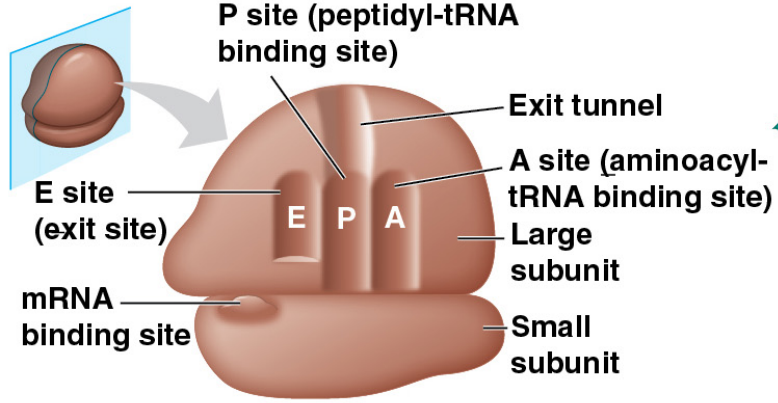
- Ribosomes facilitate specific coupling of tRNA anticodons with mRNA codons in protein synthesis
- Eukaryotic ribosomes are somewhat larger than bacterial ribosomes and differ in their molecular composition
- Some antibiotic drugs specifically inactivate bacterial ribosomes without harming eukaryotic ribosomes
- The two ribosomal subunits (large and small) are made of proteins and **ribosomal RNAs (rRNAs)**

- 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

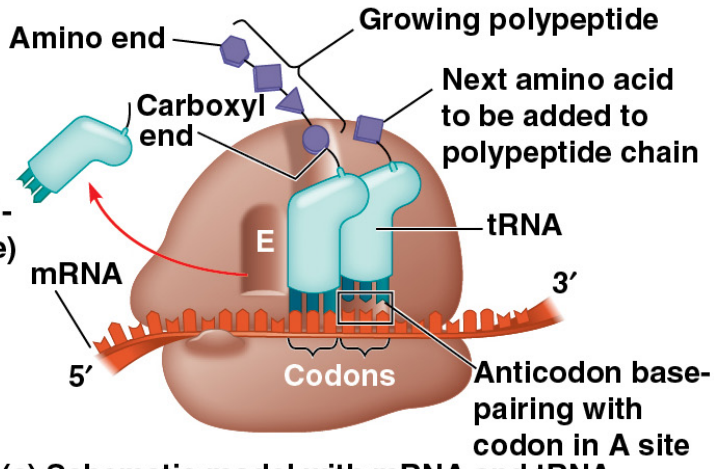
Figure 17.18



(a) Computer model of functioning ribosome



(b) Schematic model showing binding sites



(c) Schematic model with mRNA and tRNA

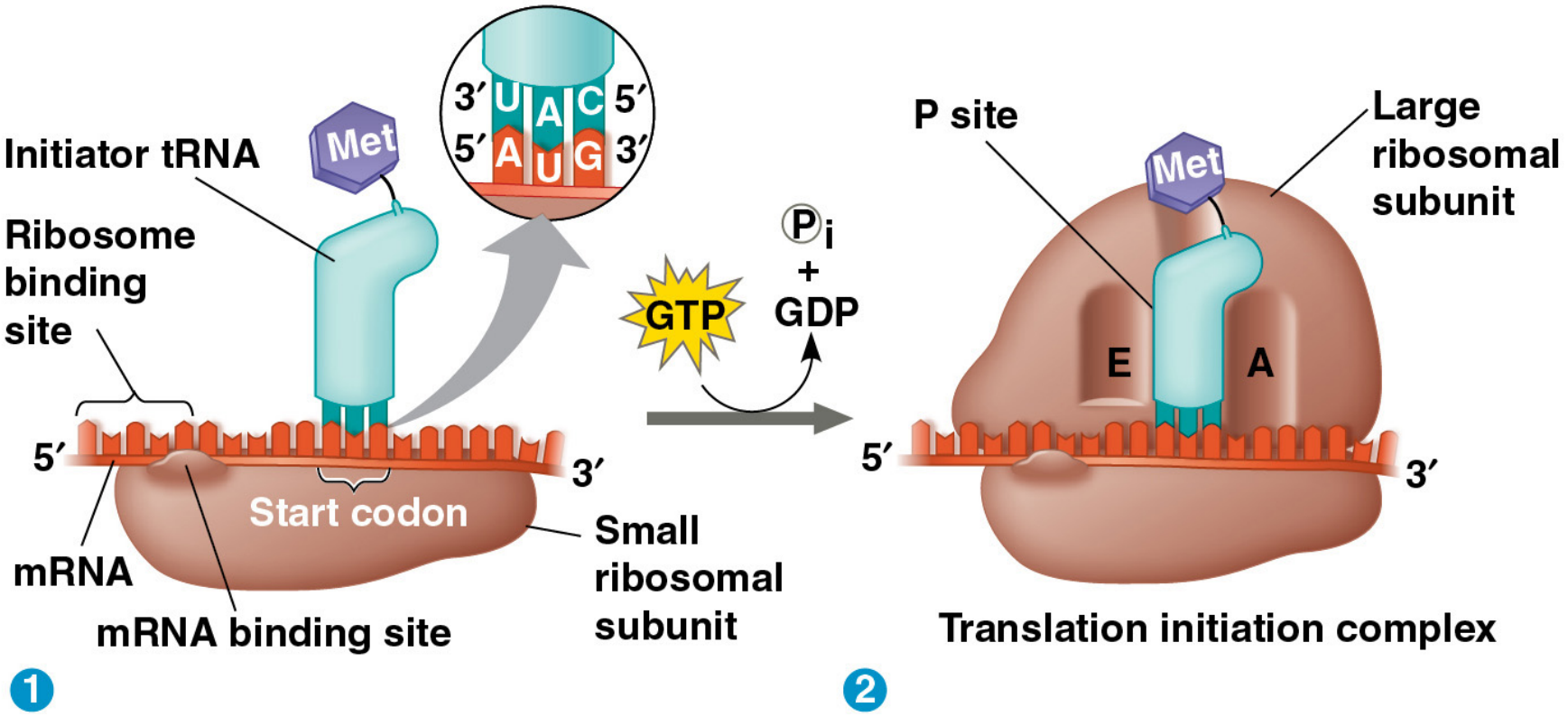
Building a Polypeptide

- The three stages of translation:
  - Initiation
  - Elongation
  - Termination
- All three stages require protein “factors” that aid in the translation process
- Energy is required for some steps, too

Ribosome Association and Initiation of Translation

- The initiation of translation starts when the small ribosomal subunit binds with mRNA and a special initiator tRNA
- The initiator tRNA carries the amino acid methionine
- 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

Figure 17.19

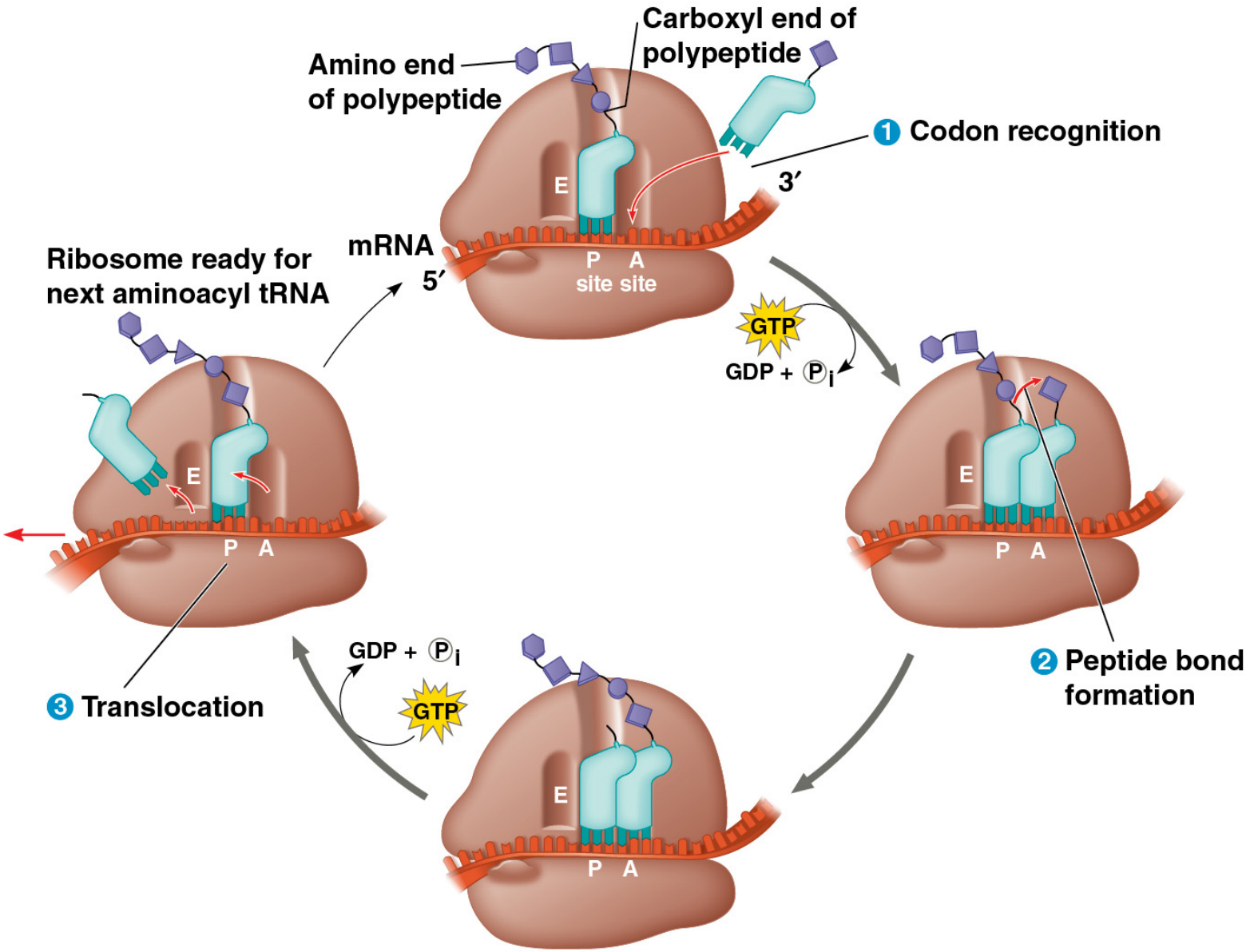

3′ U A C 5′
5′ A U G 3′
Initiator tRNA
Met
Ribosome binding site
5′
3′
Start codon
mRNA
mRNA binding site
Small ribosomal subunit
GTP
Pi + GDP
P site
Large ribosomal subunit
Met
E
A
5′
3′
Translation initiation complex
1
2

Elongation of the Polypeptide Chain

- During elongation, amino acids are added one by one to the C-terminus of the growing chain
- Each addition involves proteins called elongation factors
- Elongation occurs in three steps: codon recognition, peptide bond formation, and translocation
- Energy expenditure occurs in the first and third steps

- Translation proceeds along the mRNA in a 5′ → 3′ direction
- The ribosome and mRNA move relative to each other, codon by codon
- The elongation cycles takes less than a tenth of a second in bacteria
- Empty tRNAs released from the E site return to the cytoplasm, where they will be reloaded with the appropriate amino acid

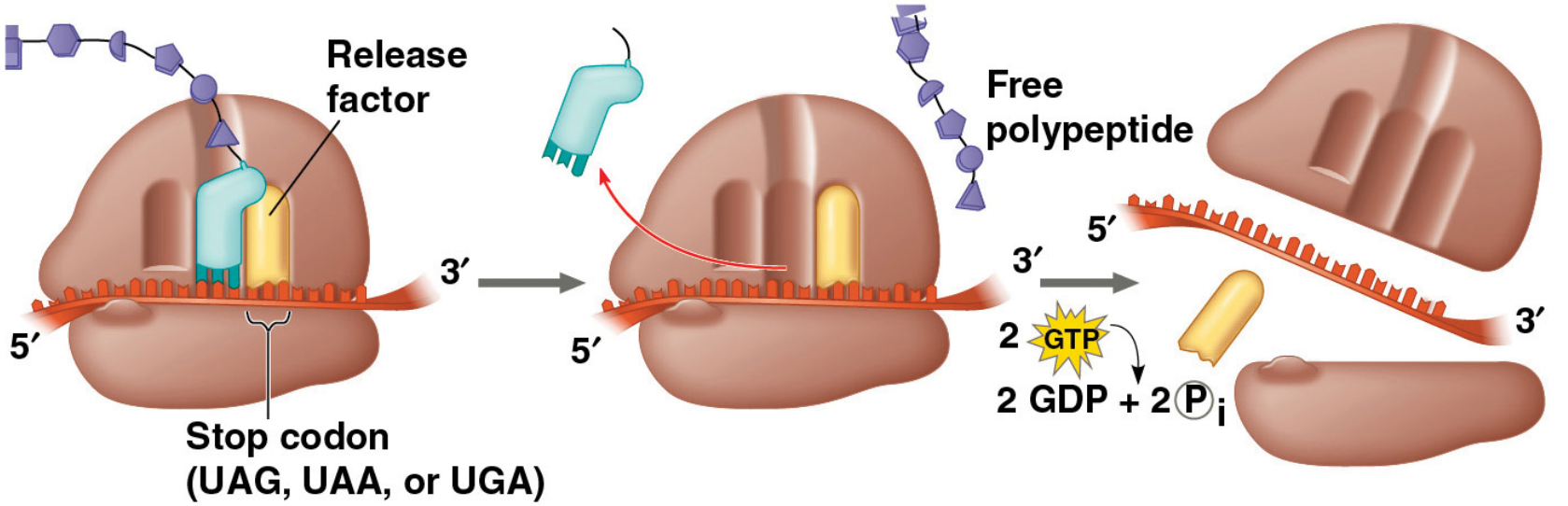
Figure 17.20


Amino end of polypeptide
Carboxyl end of polypeptide
1 Codon recognition
3′
E
P site
A site
mRNA
5′
Ribosome ready for next aminoacyl tRNA
GTP
GDP + Pi
2 Peptide bond formation
3 Translocation
GDP + Pi
GTP

Termination of Translation

- Elongation continues until a stop codon in the mRNA reaches the A site
- 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 comes apart

Figure 17.21


Release factor
3′
5′
Stop codon (UAG, UAA, or UGA)
5′
Free polypeptide
3′
5′
3′
2 GTP
2 GDP + 2 Pi
1 Ribosome reaches a stop codon on mRNA. A site accepts release factor.
2 Release factor promotes hydrolysis, freeing polypeptide.
3 Ribosomal subunits and other components dissociate.

Completing and Targeting the Functional Protein

- Often translation is not sufficient to make a functional protein
- Polypeptide chains are modified after translation or targeted to specific sites in the cell

Protein Folding and Post-Translational Modifications

- During synthesis, a polypeptide chain begins to coil and fold spontaneously into a specific shape: a three-dimensional molecule with secondary and tertiary structure
- A gene determines the primary structure, and the primary structure in turn determines shape
- Post-translational modifications may be required before the protein can begin doing its particular job in the cell

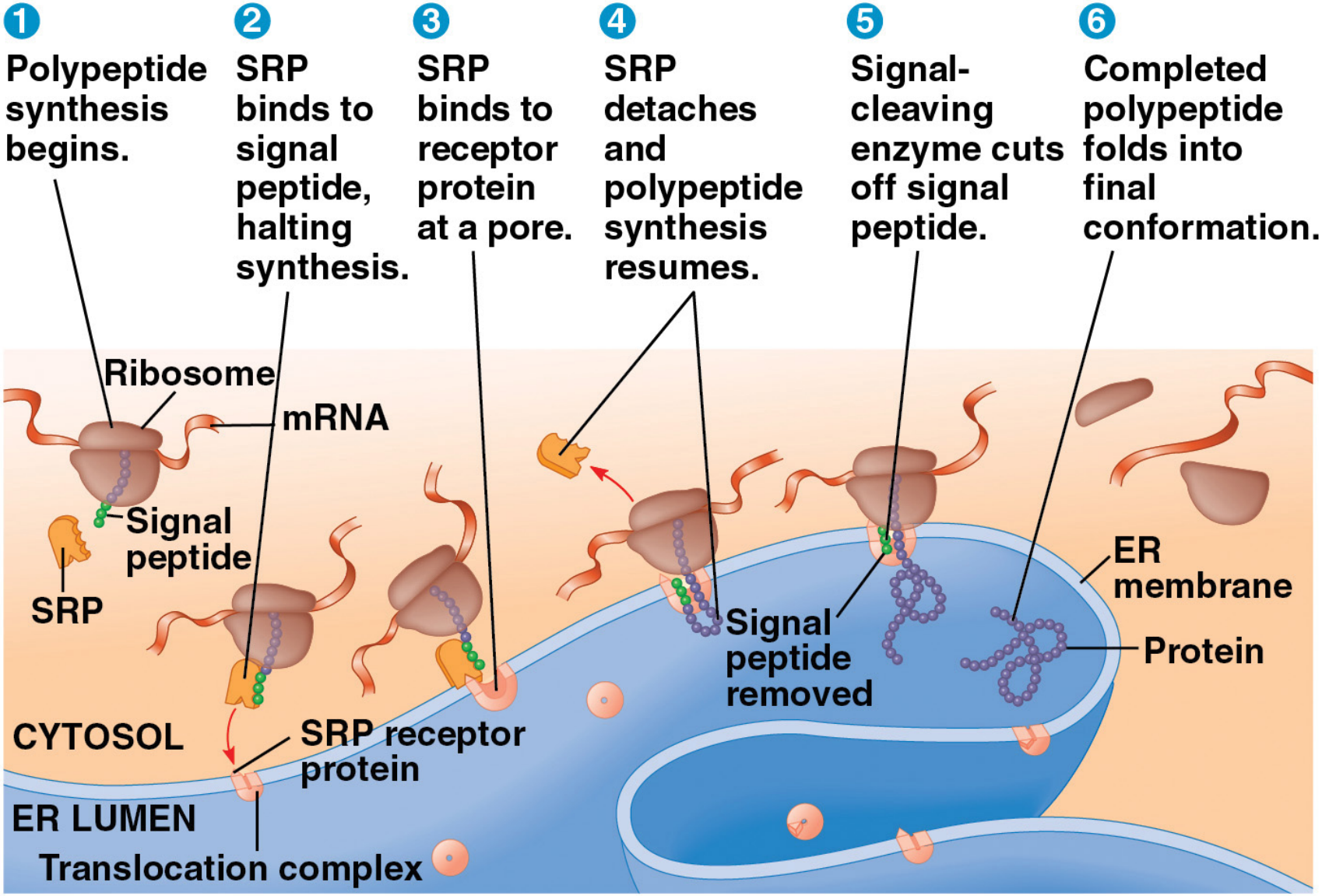
Targeting Polypeptides to Specific Locations

- 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

- Polypeptide synthesis always begins in the cytosol
- Synthesis finishes in the cytosol unless the polypeptide signals the ribosome to attach to the ER
- Polypeptides destined for the ER or for secretion are marked by a **signal peptide**
- The signal peptide is a sequence of about 20 amino acids at or near the leading end of the polypeptide

- A **signal-recognition particle (SRP)** binds to the signal peptide
- The SRP escorts the ribosome to a receptor protein built into the ER membrane
- The signal peptide is removed by an enzyme
- Other kinds of signal peptides target polypeptides to other organelles

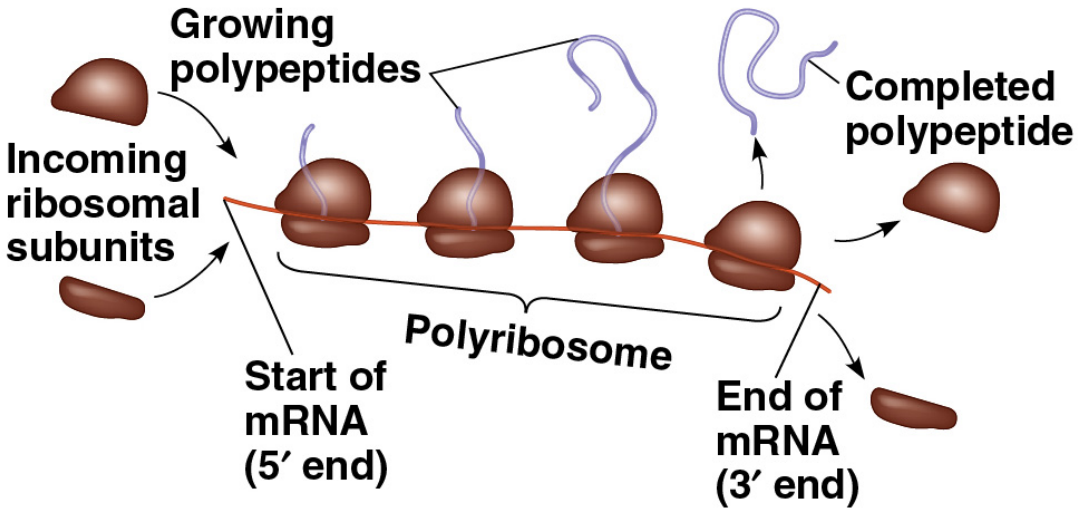
Figure 17.22


1
Polypeptide synthesis begins.
2
SRP binds to signal peptide, halting synthesis.
3
SRP binds to receptor protein at a pore.
4
SRP detaches and polypeptide synthesis resumes.
5
Signal-cleaving enzyme cuts off signal peptide.
6
Completed polypeptide folds into final conformation.
Ribosome
mRNA
Signal peptide
SRP
CYTOSOL
SRP receptor protein
ER LUMEN
Translocation complex
Signal peptide removed
ER membrane
Protein

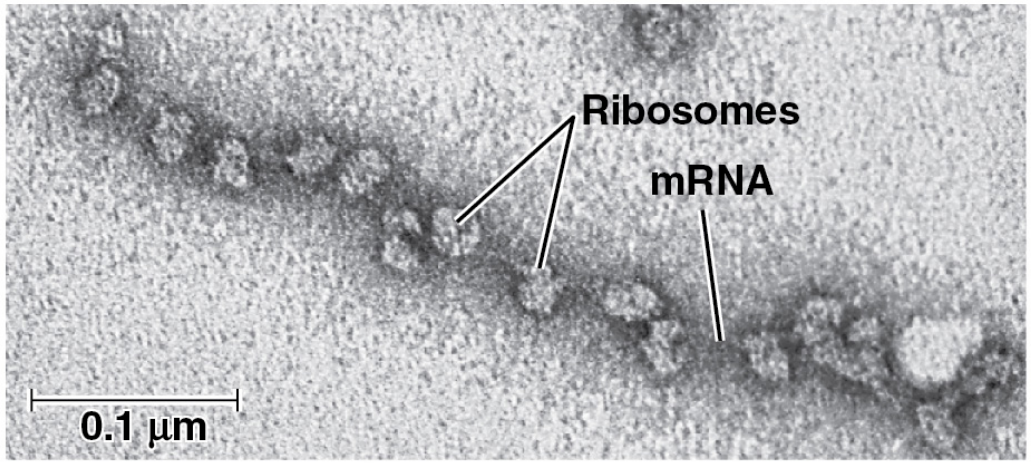
Making Multiple Polypeptides in Bacteria and Eukaryotes

- Multiple 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

Figure 17.23



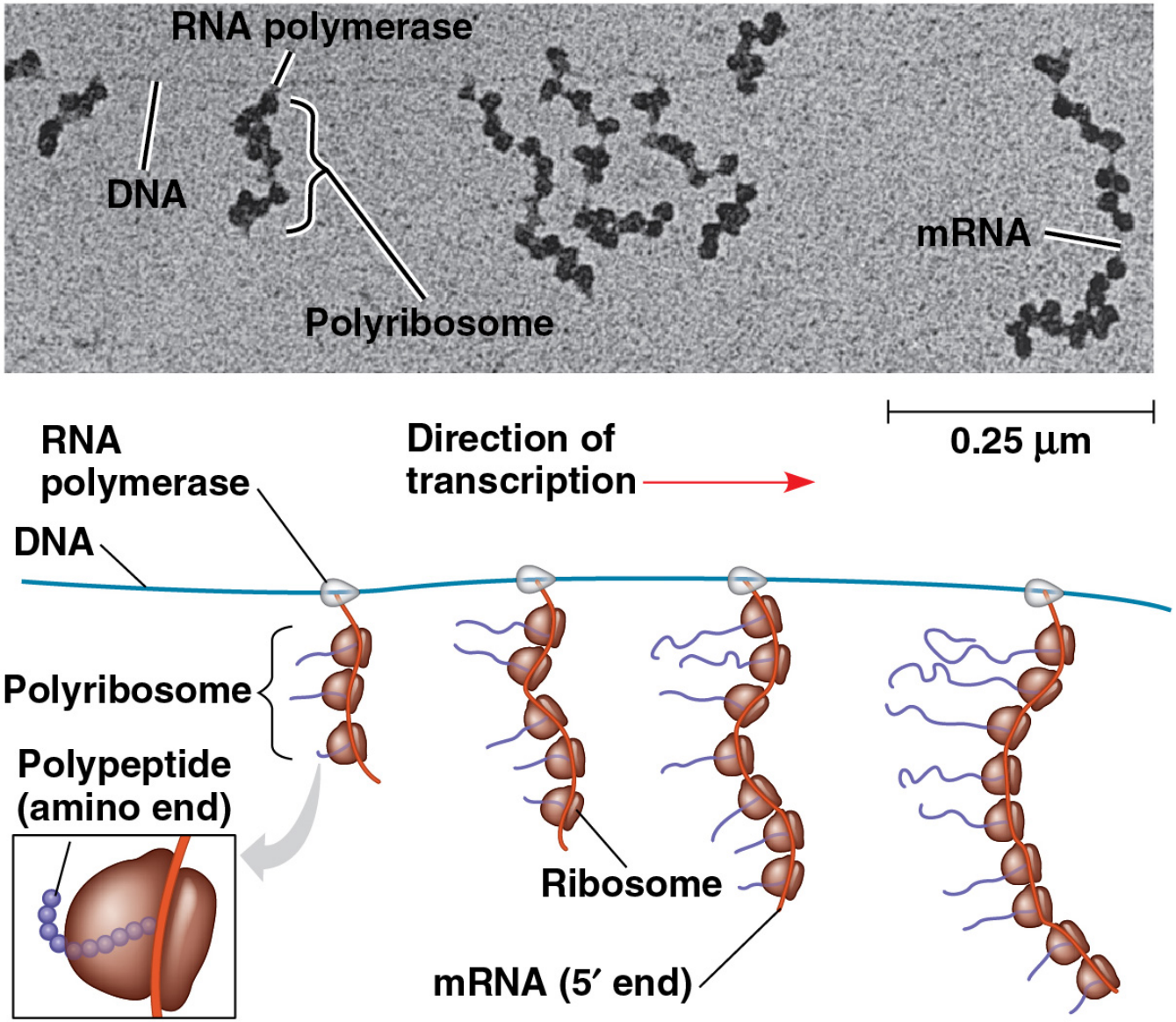
(a) Several ribosomes simultaneously translating one mRNA molecule



(b) A large polyribosome in a bacterial cell (TEM)

- A bacterial cell ensures a streamlined process by coupling transcription and translation
- In this case the newly made protein can quickly diffuse to its site of function

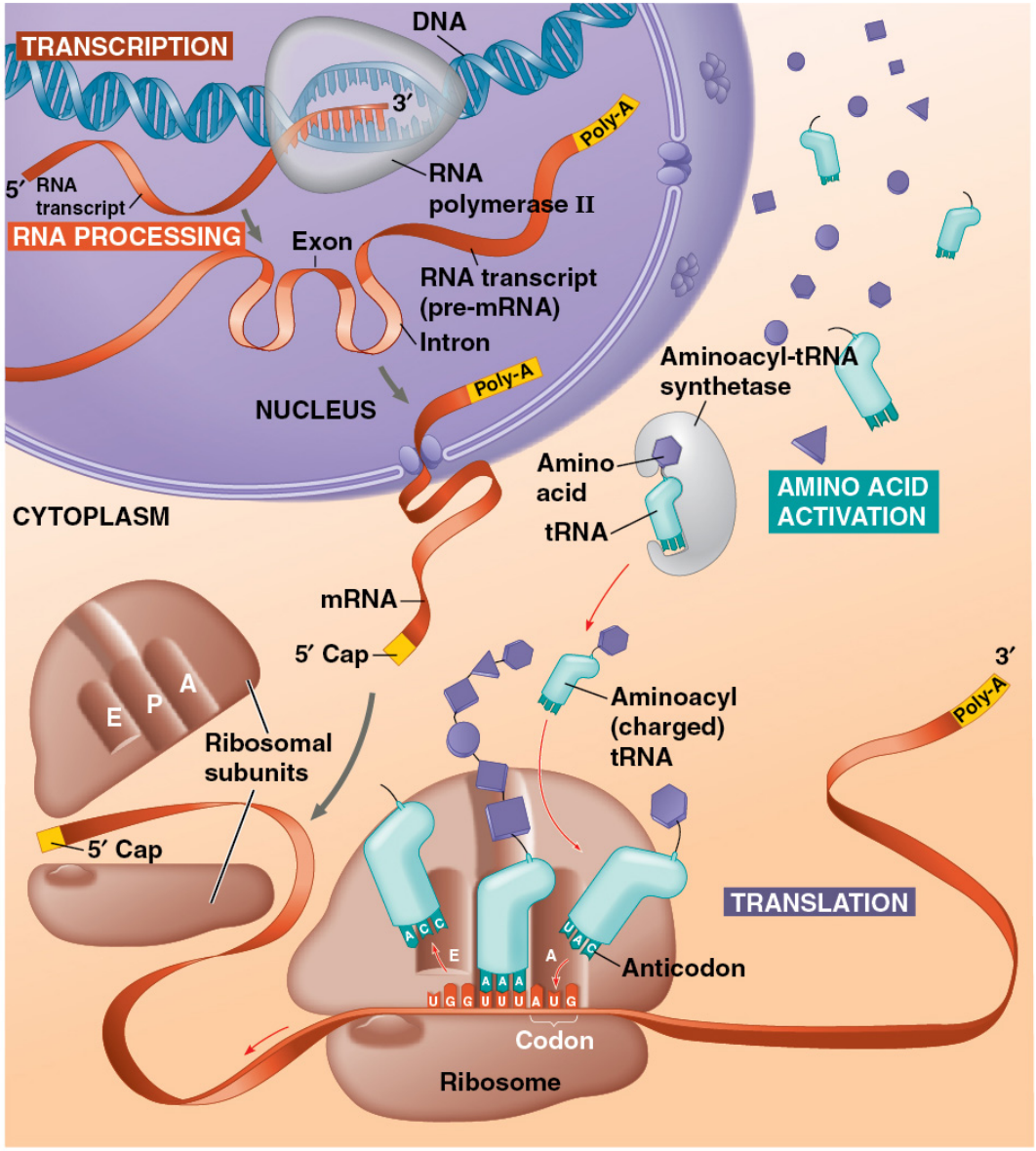
Figure 17.24


RNA polymerase
DNA
Polyribosome
mRNA
0.25 µm
RNA polymerase
Direction of transcription
DNA
Polyribosome
Polypeptide (amino end)
Ribosome
mRNA (5′ end)

Animation: Overview of Protein Synthesis in Bacteria

- In eukaryotes, the nuclear envelope separates the processes of transcription and translation
- RNA undergoes processing before leaving the nucleus

Figure 17.25


TRANSCRIPTION
DNA
3′
Poly-A
5′ RNA transcript
RNA polymerase II
RNA PROCESSING
Exon
RNA transcript (pre-mRNA)
Intron
NUCLEUS
Poly-A
Aminoacyl-tRNA synthetase
Amino acid
tRNA
AMINO ACID ACTIVATION
CYTOPLASM
mRNA
5′ Cap
E
P
A
Ribosomal subunits
5′ Cap
Aminoacyl (charged) tRNA
3′
Poly-A
TRANSLATION
Anticodon
Codon
Ribosome

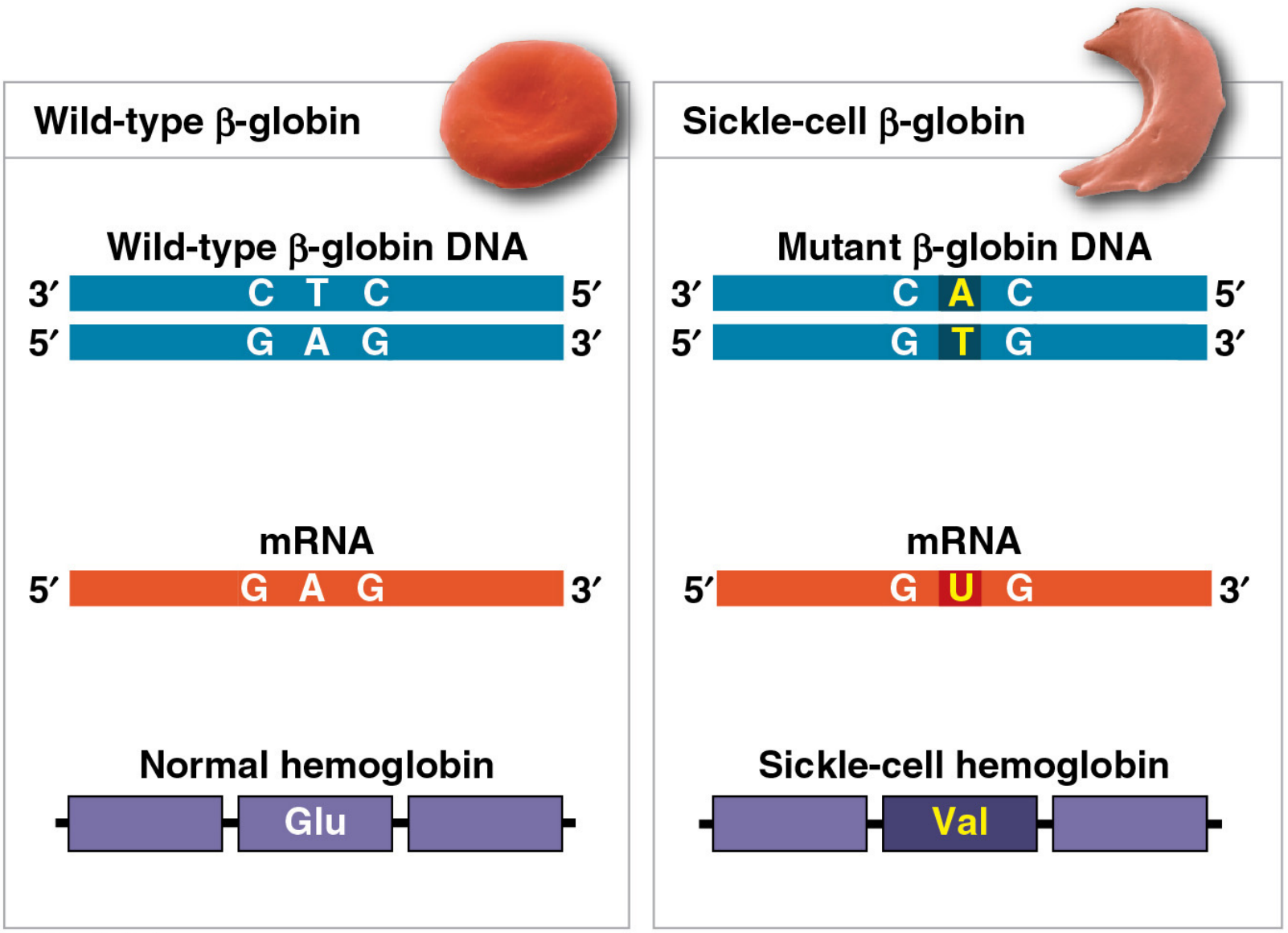
Animation: Overview of Protein Synthesis in Eukaryotes

CONCEPT 17.5: Mutations of one or a few nucleotides can affect protein structure and function

- **Mutations** are changes in the genetic information of a cell
- **Point mutations** are changes in just one nucleotide pair of a gene
- The change of a single nucleotide in a DNA template strand can lead to the production of an abnormal protein

- If a mutation has an adverse effect on the phenotype of the organism, the condition is referred to as a genetic disorder or hereditary disease

Figure 17.26


Wild-type β-globin
Wild-type β-globin DNA
3′ C T C 5′
5′ G A G 3′
mRNA
5′ G A G 3′
Normal hemoglobin
Glu
Sickle-cell β-globin
Mutant β-globin DNA
3′ C A C 5′
5′ G T G 3′
mRNA
5′ G U G 3′
Sickle-cell hemoglobin
Val

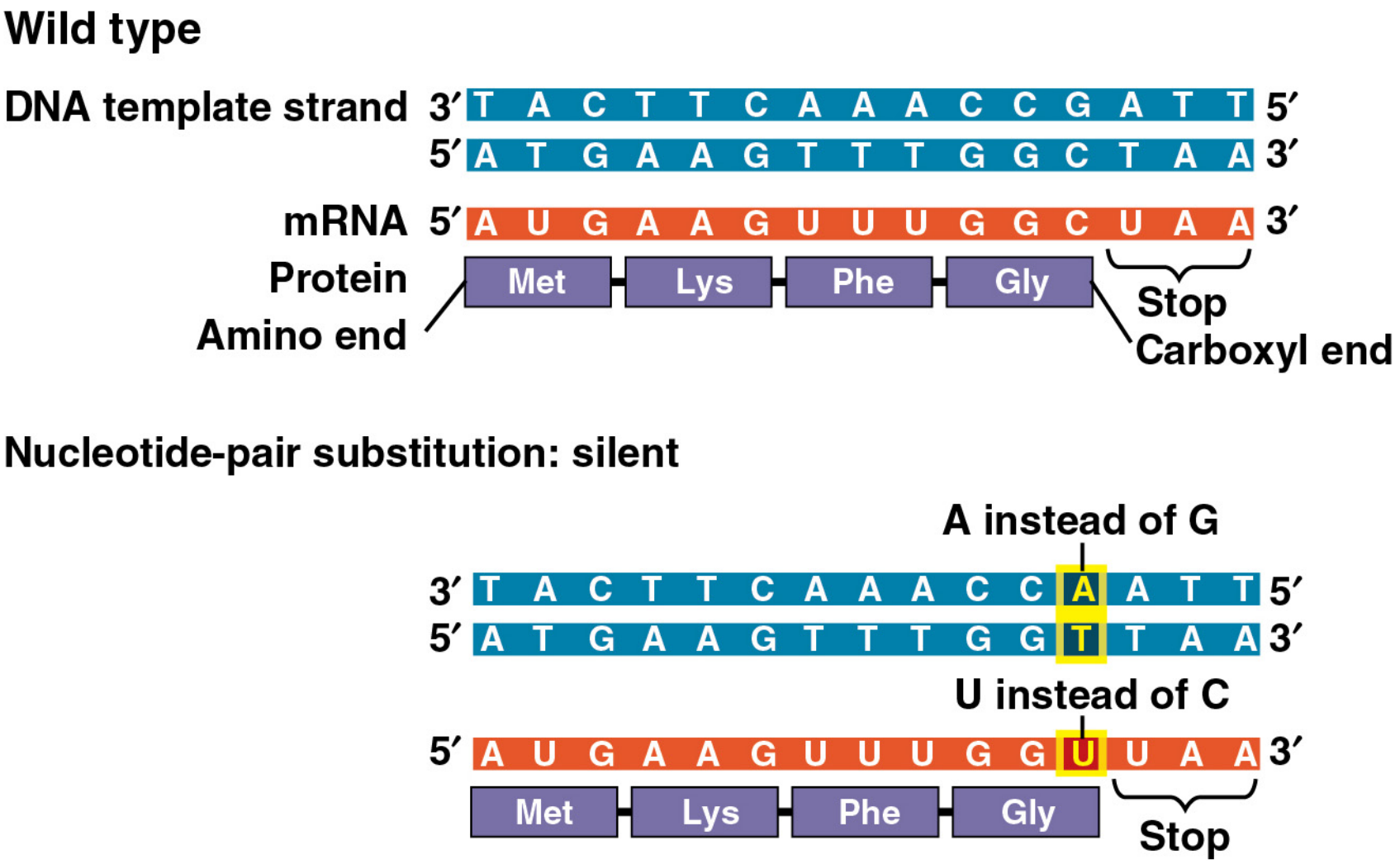
Types of Small-Scale Mutations

- Point mutations within a gene can be divided into two general categories:
  - Single nucleotide-pair substitutions
  - Nucleotide-pair insertions or deletions

Substitutions

- A **nucleotide-pair substitution** replaces one nucleotide and its partner with another pair of nucleotides
- **Silent mutations** have no effect on the amino acid produced by a codon because of redundancy in the genetic code
- **Missense mutations** still code for an amino acid, but not the correct amino acid
- **Nonsense mutations** change an amino acid codon into a stop codon; most lead to a nonfunctional protein

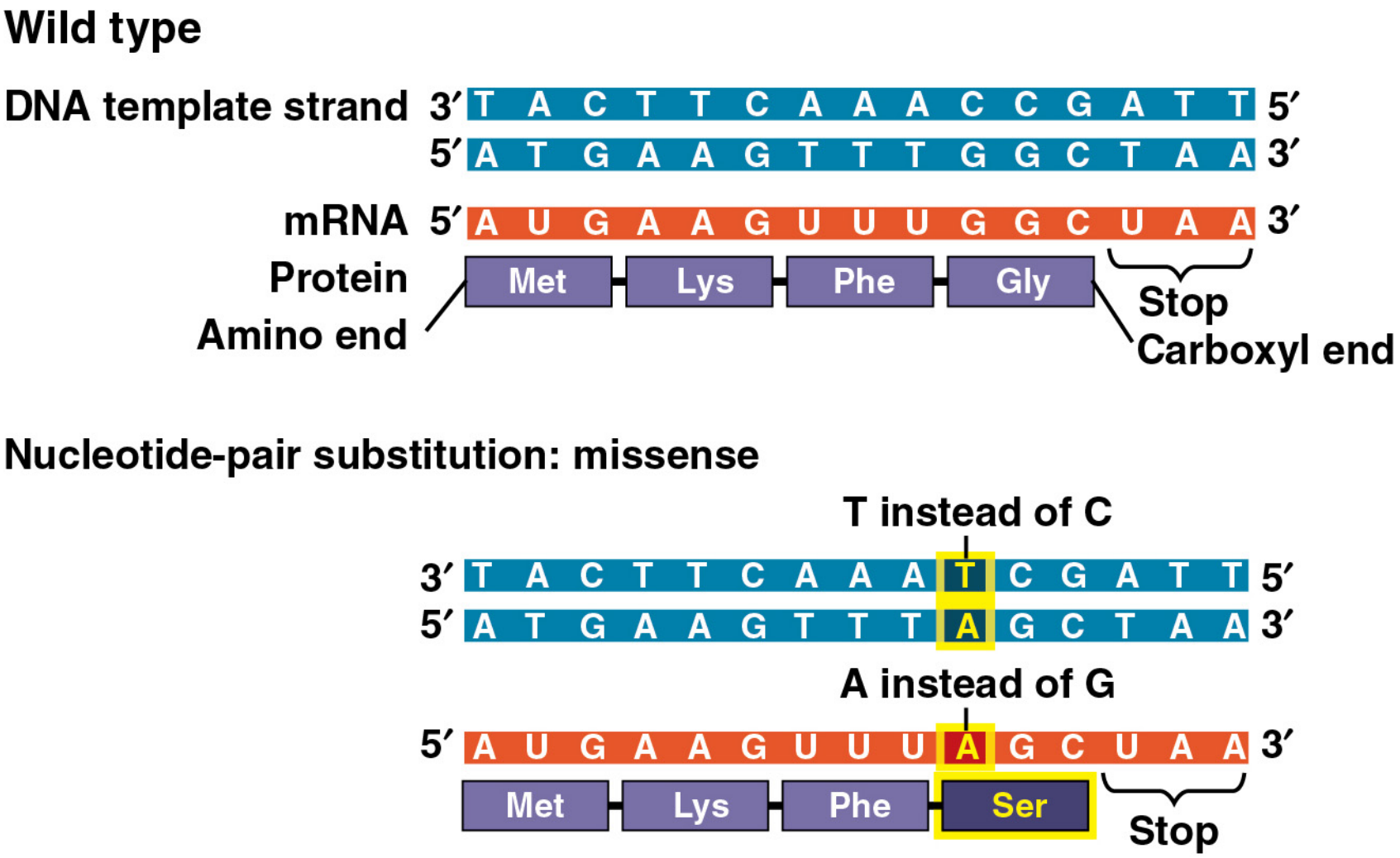
Figure 17.27a



Insertions and Deletions

- **Insertions** and **deletions** are additions or losses of nucleotide pairs in a gene
- These mutations have a disastrous effect on the resulting protein more often than substitutions do
- Insertion or deletion of nucleotides may alter the reading frame, producing a **frameshift mutation**
- Insertions or deletions outside the coding part of a gene could affect how the gene is expressed

Figure 17.27


Wild type
DNA template strand
3′ T A C T T C A A A C C G A T T 5′
5′ A T G A A G T T T G G C T A A 3′
mRNA
5′ A U G A A G U U U G G C U A A 3′
Protein
Met
Lys
Phe
Gly
Stop
Amino end
Carboxyl end
Nucleotide-pair substitution: missense
T instead of C
3′ T A C T T C A A A T C G A T T 5′
5′ A T G A A G T T T A G C T A A 3′
A instead of G
5′ A U G A A G U U U A G C U A A 3′
Met
Lys
Phe
Ser
Stop

Animation: Mutation Types

New Mutations and Mutagens

- Spontaneous mutations can occur during errors in DNA replication or recombination
- **Mutagens** are physical or chemical agents that can cause mutations
- Chemical mutagens fall into a variety of categories
- Most carcinogens (cancer-causing chemicals) are mutagens, and most mutagens are carcinogenic

Using CRISPR to Edit Genes and Correct Disease-Causing Mutations

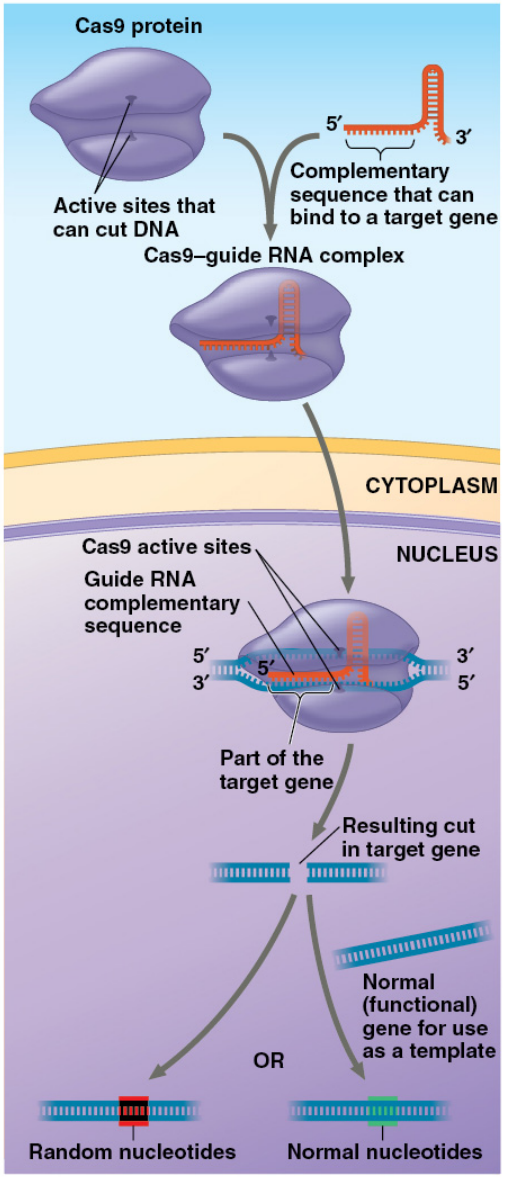
- Biologists who study disease-causing mutations have sought techniques for **gene editing**–altering genes in a specific way
- The powerful technique called **CRISPR-Cas9** is transforming the field of genetic engineering
- In bacteria, the protein Cas9 acts together with a guide RNA to help defend bacteria from viral infection

- The Cas9 protein will cut any sequence to which it is targeted
- Scientists can introduce a Cas9–guide RNA complex into a cell they wish to alter
- The guide RNA is engineered to target a gene
- Cas9 cuts both strands of the targeted gene

- The broken ends trigger a DNA repair system
- The repair enzymes remove or add some random nucleotides while joining the broken ends
- This is a way for researchers to “knock out” (disable) a given gene, to study what the gene does in an organism

- To treat genetic disease, researchers have modified this technique
- They can introduce a template with a normal (functional) copy of the gene to be corrected
- In this way, the CRISPR-Cas9 system edits the defective gene and corrects it

Figure 17.28


Cas9 protein
5′
3′
Complementary sequence that can bind to a target gene
Active sites that can cut DNA
Cas9–guide RNA complex
CYTOPLASM
NUCLEUS
Cas9 active sites
Guide RNA complementary sequence
5′
3′
5′
3′
5′
Part of the target gene
Resulting cut in target gene
Normal (functional) gene for use as a template
OR
Random nucleotides
Normal nucleotides

- Some genetic conditions like human sickle-cell disease have been somewhat successfully treated using mice
- There are still concerns about using the technique in humans
- There is the possibility of unintended effects on genes that have not been targeted
- Biologists have agreed to use extreme caution as the field moves forward

What Is a Gene? *Revisiting the Question*

- The idea of the gene has evolved through the history of genetics
- 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

- A gene can be defined as a region of DNA that can be expressed to produce a final functional product that is either a polypeptide or an RNA molecule

Figure 17.UN05

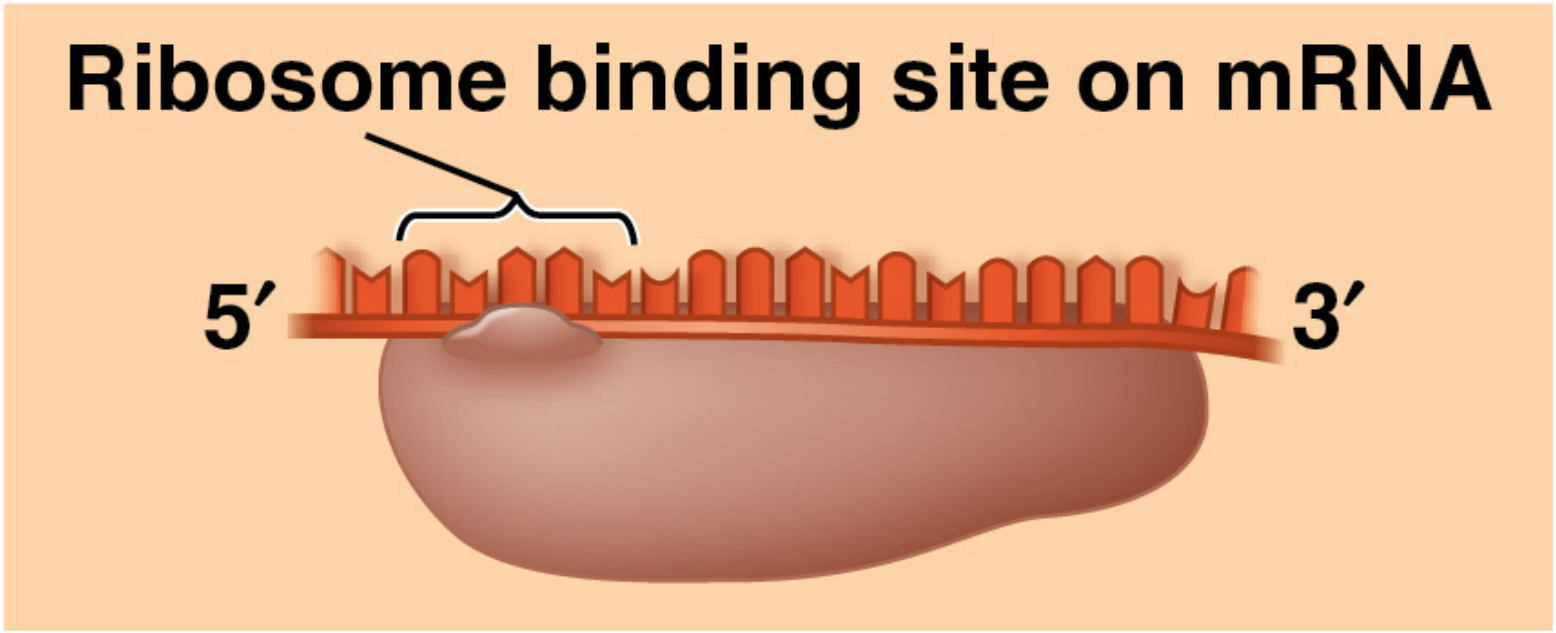
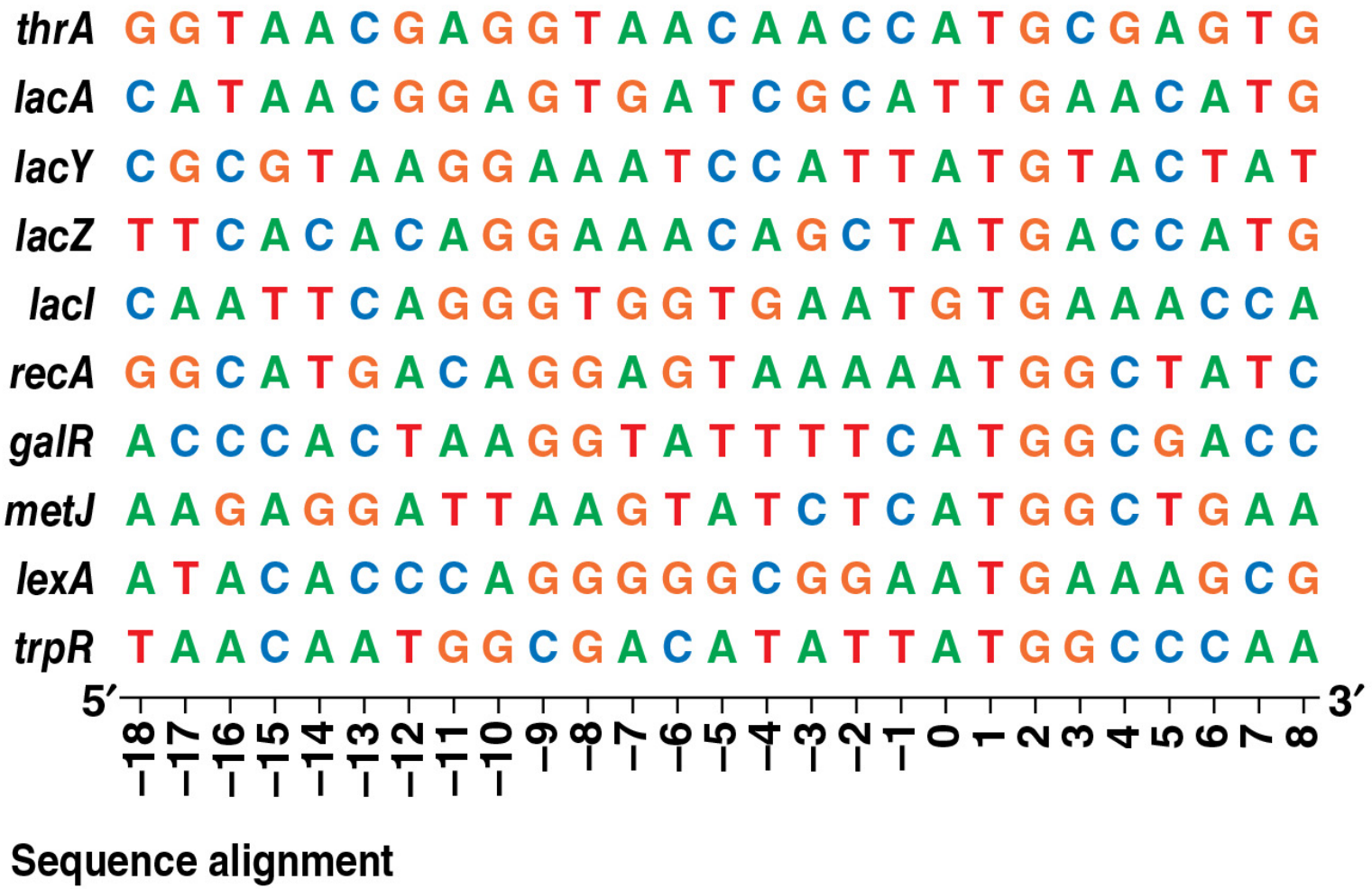
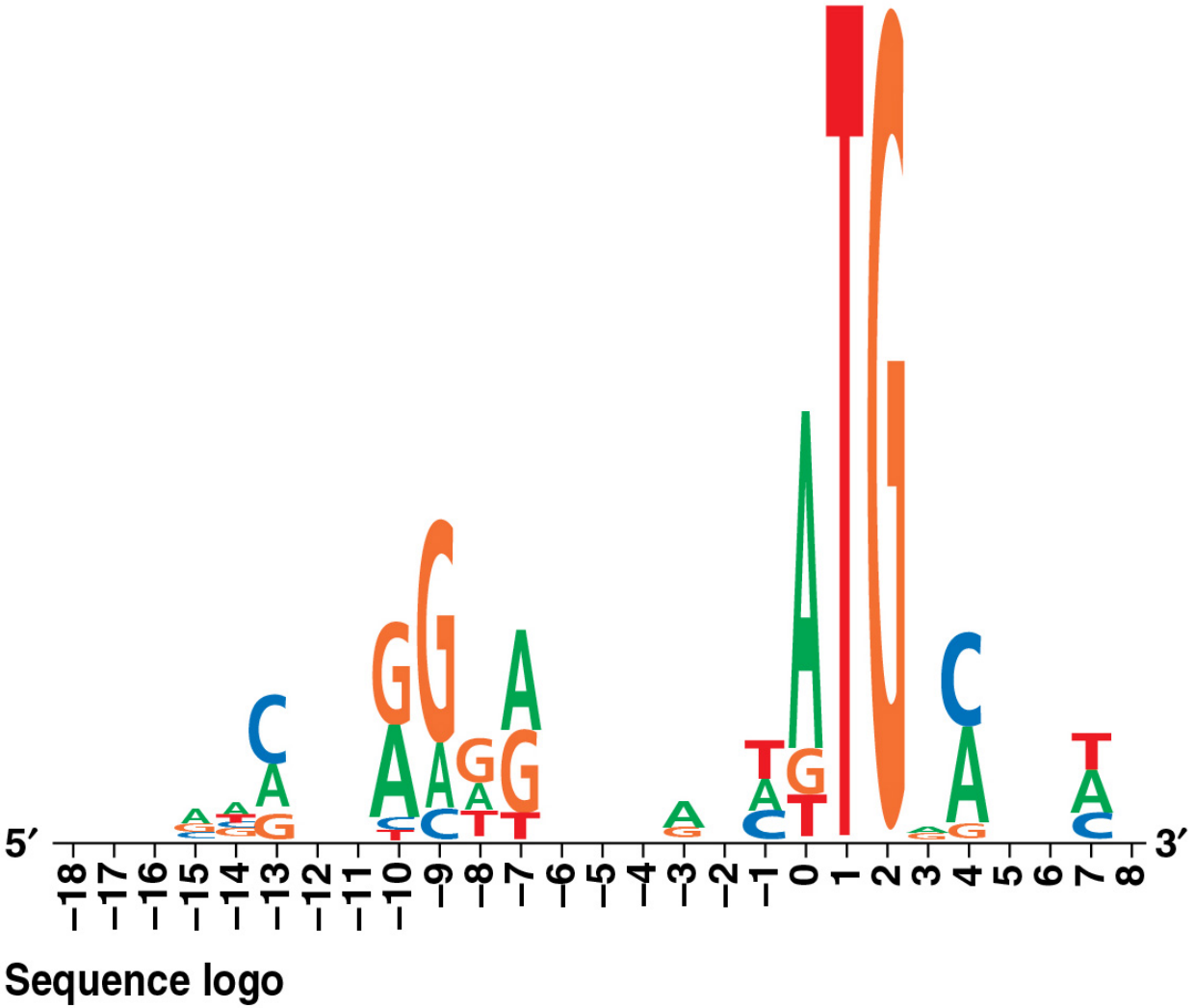


Figure 17.UN02



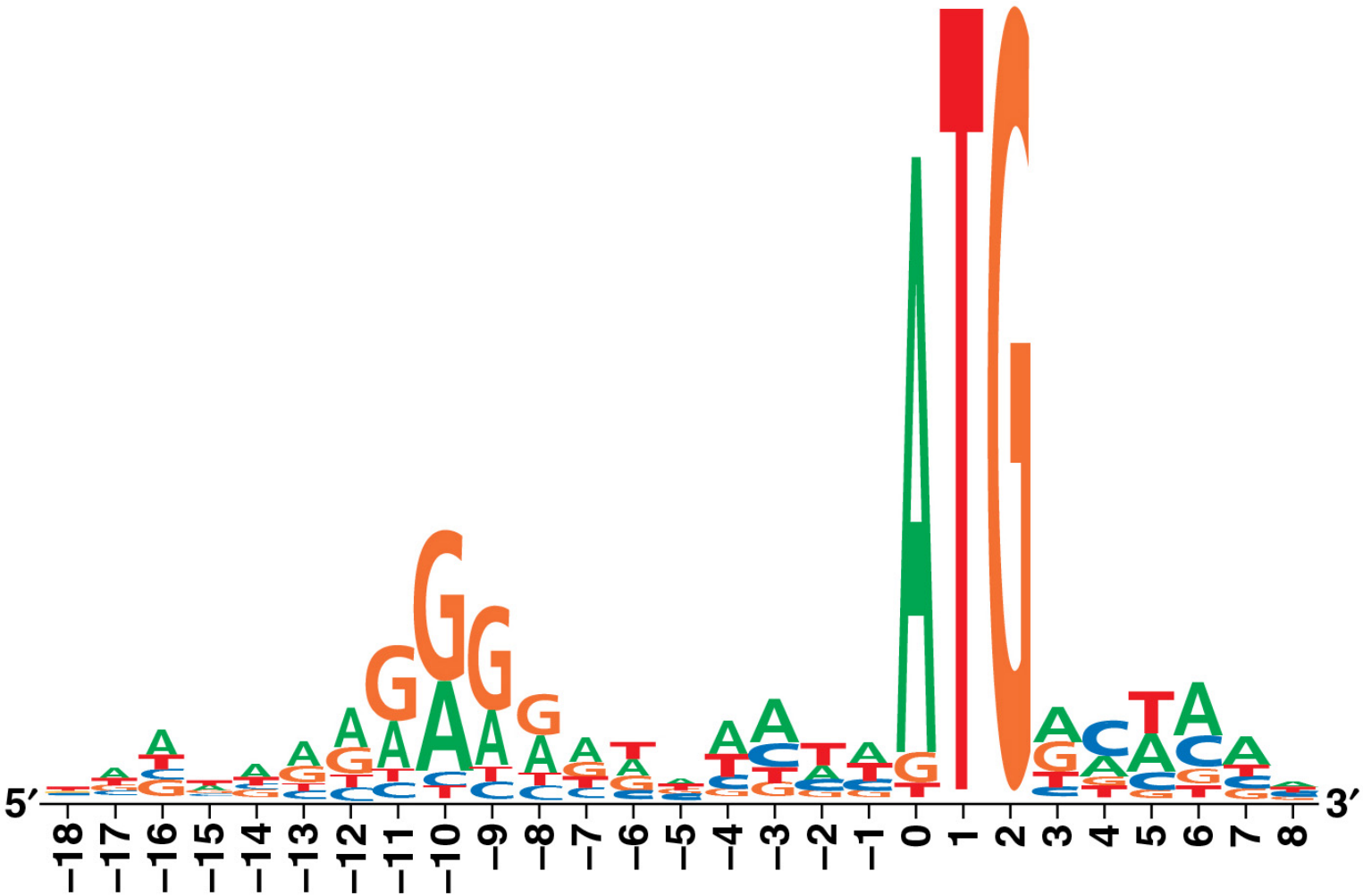
Further Reading T. D. Schneider and R. M. Stephens, Sequence logos: A new way to display consensus sequences, *Nucleic Acids Research* 18:6097–6100 (1990).

Figure 17.UN03



Further Reading T. D. Schneider and R. M. Stephens, Sequence logos: A new way to display consensus sequences, *Nucleic Acids Research* 18:6097–6100 (1990).

Figure 17.UN04



Further Reading T. D. Schneider and R. M. Stephens, Sequence logos: A new way to display consensus sequences, *Nucleic Acids Research* 18:6097–6100 (1990).

Figure 17.UN06

Figure 17.UN07

Wild-type cDNA 5′–CTG GTG GAA GCT CTC TAC CTA GTG TGC GGG GAA CGA GGC TTC TTC TAC ACA CCC AAG ACC–3′

Patient 1 cDNA 5′–CTG GTG GAA GCT CTC TAC CTA GTG TGC GGG GAA CGA GGC TGC TTC TAC ACA CCC AAG ACC–3′

Patient 2 cDNA 5′–CTG GTG GAA GCT CTC TAC CTA GTG TGC GGG GAA CGA GGC TCC TTC TAC ACA CCC AAG ACC–3′

Patient 3 cDNA 5′–CTG GTG GAA GCT CTC TAC CTA GTG TGC GGG GAA CGA GGC TTC TTG TAC ACA CCC AAG ACC–3′

Data from N. Nishi and K. Nanjo, Insulin gene mutations and diabetes, *Journal of Diabetes Investigation* 2:92–100 (2011).

Figure 17.UN09

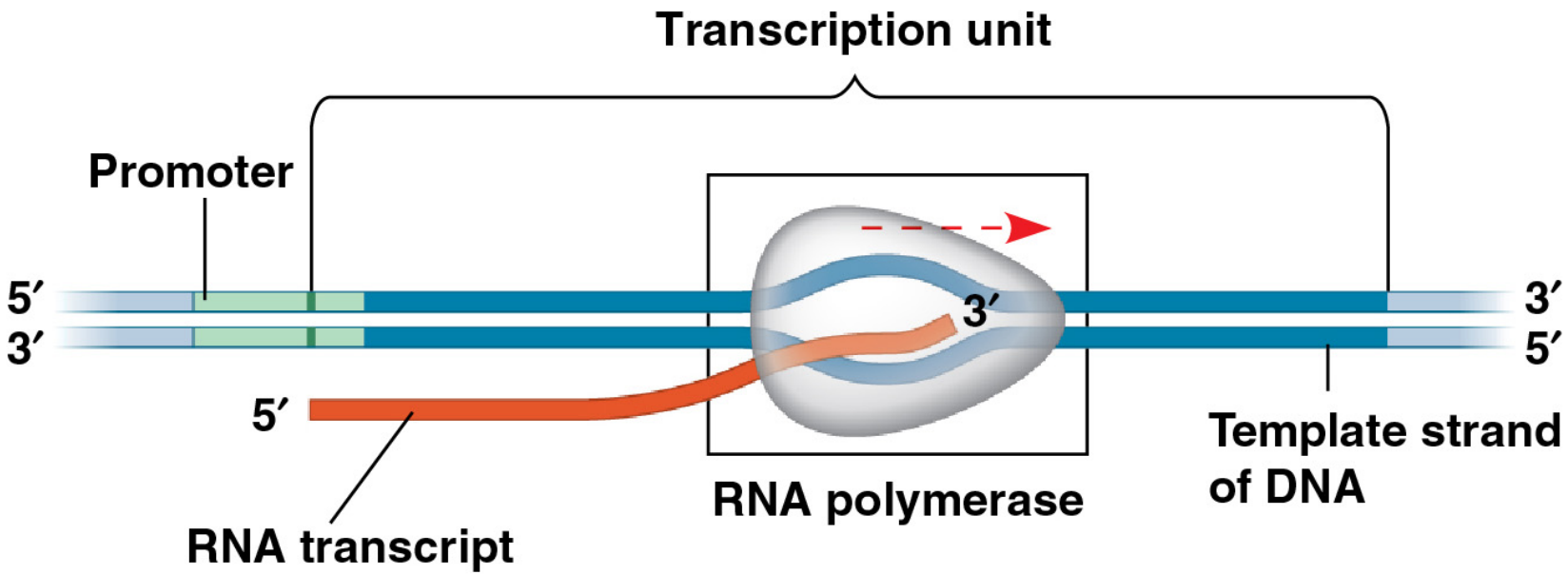


Figure 17.UN10

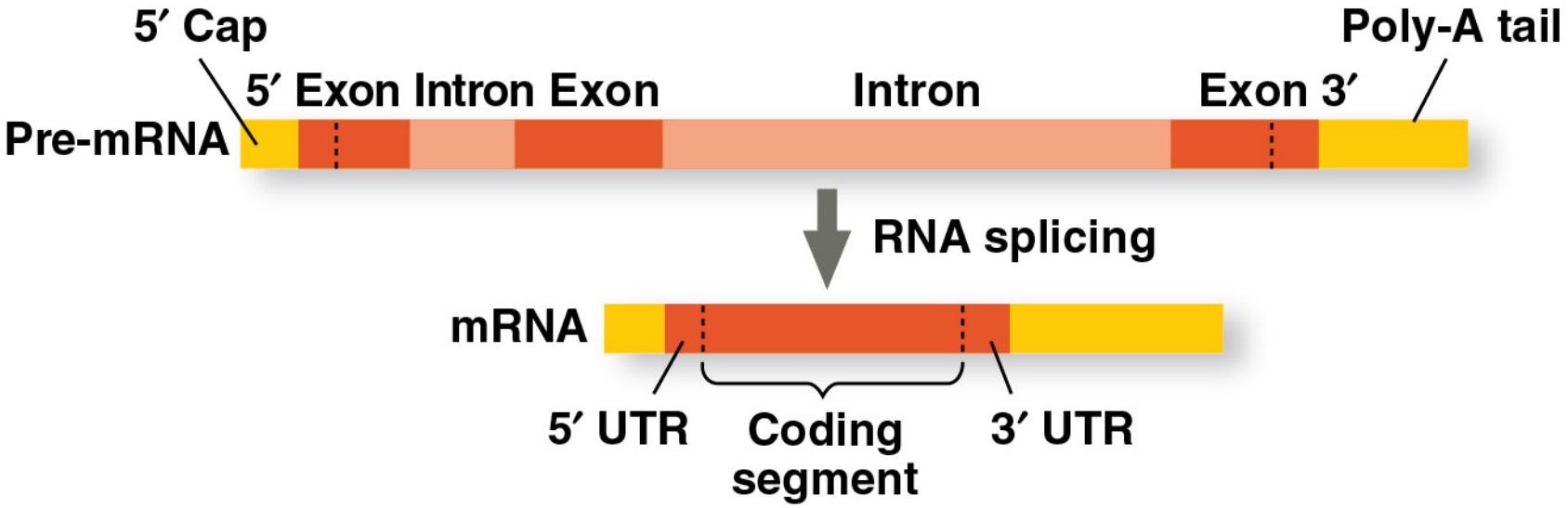


Figure 17.UN11

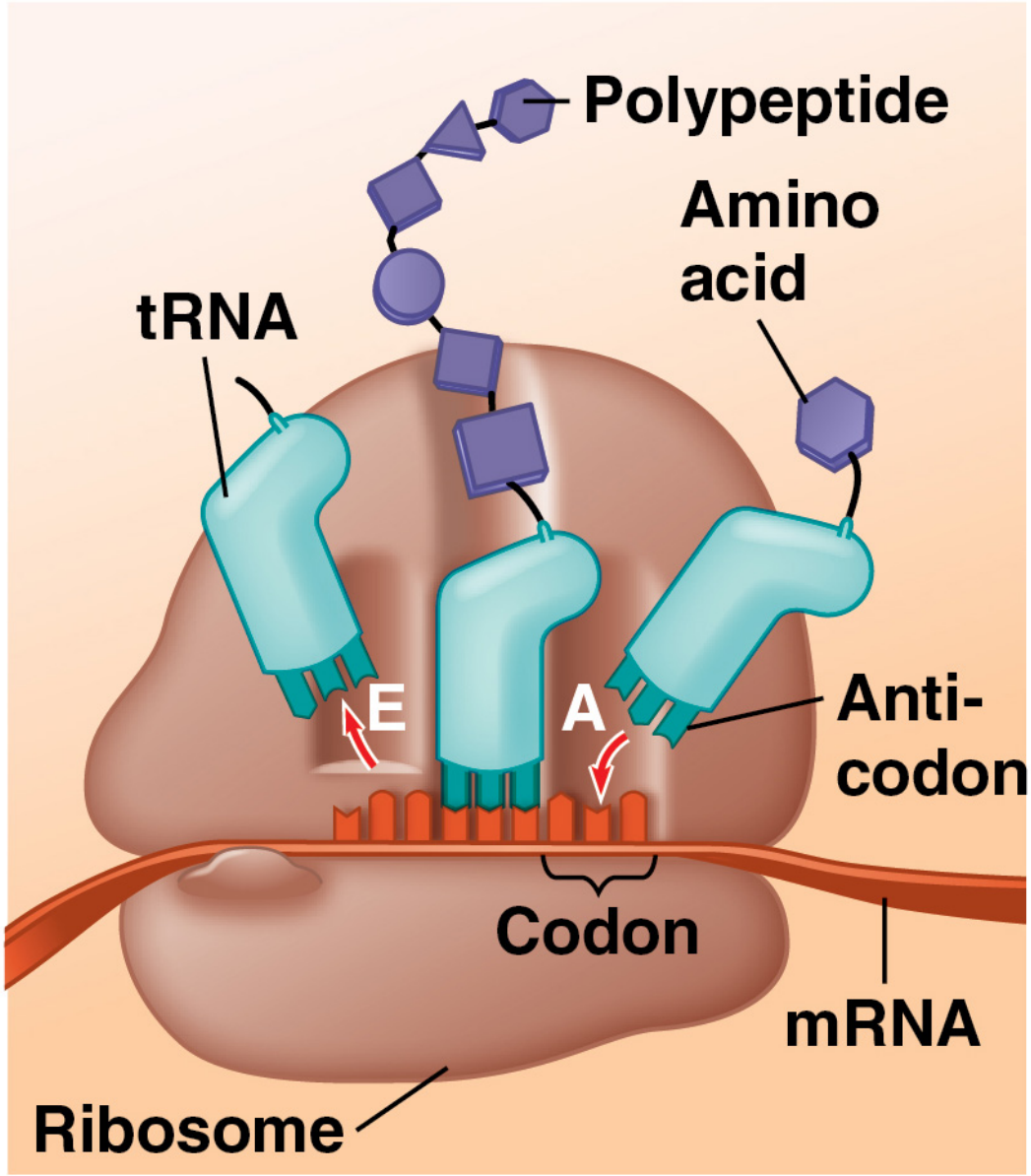


Figure 17.UN12