Lecture 15: Translation and Transcription

Making proteins requires that genetic information is transcribed from DNA into RNA and then translated into protein. The overall process can be viewed as two separate processes—transcription and translation.

I. Background

A. RNA
   1. RNA is a nucleic acid
      a. Polymer nucleotides
   2. RNA structure differs from DNA
      a. Five-carbon sugar in RNA nucleotides is ribose (not deoxyribose)
      b. Pyrimidine in RNA is uracil instead of thymine
   3. The linear sequence of DNA ultimately determines the primary structure of proteins
      a. Unique combination of nucleotides code for the 20 amino acids that make up proteins

B. Terminology
   1. Transcription—synthesis of RNA from a DNA template
      a. Unique nucleotide sequence of a gene is transcribed from DNA into a complementary nucleotide sequence in messenger RNA (mRNA)
      b. mRNA carries the protein-building instructions to the cellular machinery that synthesizes protein
   2. Translation—polypeptide synthesis under the direction of mRNA
      a. Linear sequence of mRNA is translated into a linear sequence of amino acids in a polypeptide
      b. Translation occurs on a complex particle known as a ribosome.
         i. Ribosomes are comprised of ribosomal RNA (rRNA) and protein.
         ii. Ribosomes facilitate the orderly linking of amino acids into polypeptide chains

C. Differences in protein synthesis between prokaryotes and eukaryotes
   1. Prokaryotes do not have nuclei
      a. DNA is not segregated from ribosomes and protein machinery.
      b. Translation and transcription occur in rapid succession
   2. Eukaryotes have nuclear envelopes that segregate DNA segregated from ribosomes and protein machinery.
      a. Transcription takes place in the nucleus
      b. Translation takes place in the cytoplasm
      c. mRNA serves as an intermediary
      d. mRNA is modified (i.e., RNA processing) before it leaves the nucleus
D. Nucleotide triplets specify amino acids

1. There is not a one-to-one correspondence between nucleotides and amino acids
   a. There are only 4 nucleotides and 20 amino acids
   b. A two-to-one correspondence would only specify 16 amino acids ($4^2$)
   c. A three-to-one correspondence would specify 64 amino acids ($4^3$)

   (i) Nucleotide triplets are known as *codons*
   (ii) Basic unit of the genetic code
   (iii) Specify either individual amino acids or termination
2. Genes are not directly translated into amino acids
   a. DNA is transcribed into mRNA first
   b. For each gene, only one of the two DNA strands (i.e., template strand) is transcribed
   c. The complementary non-template strand is the parental strand for making a new template when DNA replicates
   d. The same DNA strand can serve as the template for some genes and the non-template strand for others
3. mRNA is complementary to the DNA strand from which it is transcribed

<table>
<thead>
<tr>
<th>DNA</th>
<th>A</th>
<th>C</th>
<th>C</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>C</th>
<th>C</th>
<th>G</th>
<th>G</th>
<th>C</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA</td>
<td>U</td>
<td>G</td>
<td>G</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>U</td>
</tr>
<tr>
<td>Protein</td>
<td>Tryptophan</td>
<td>Phenylalanine</td>
<td>Glycine</td>
<td>Cysteine</td>
<td>Arganine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. During translation, the linear sequence of codons along mRNA is translated into the linear sequence of amino acids in a polypeptide
   a. There is redundancy in the code but no ambiguity
      i. Redundancy exists because two or more codons differing only in their third base (e.g., UUU and UUC each code for phenylalanine)
      ii. Codons only code for one amino acid, therefore there is no ambiguity

II. Transcription

A. DNA-directed synthesis of RNA

1. RNA polymerase catalyzes this synthesis
   a. Two functions of RNA polymerase:
      i. Separates DNA strands
      ii. Links RNA nucleotides from the 5’ → 3’ direction as they base-pair along the DNA template
2. Types of RNA polymersae
   a. Prokaryotes have a single type of polymerase that synthesizes all types of RNA (e.g., tRNA, rRNA, mRNA)
   b. Eukaryotes have three types:
      i. RNA polymerase II catalyzes mRNA synthesis
3. Transcription unit
   a. A specific sequence of DNA that is transcribed into a single mRNA molecule
   b. Includes sequences with specific functions
      i. Region where transcription begins (initiation)
      ii. Region where transcription ends (termination)
      iii. Intervening nucleotides (i.e., gene)
   c. In eukaryotes, the transcription unit contains a single gene
      i. Resulting mRNA codes for only one peptide
   d. In prokaryotes, the transcription unit can contain several genes
      i. Resulting mRNA may code for different proteins
      ii. Functionally related

B. Stages of transcription
   1. Polymerase binding and initiation
   2. Elongation
   3. Termination

C. Polymerase binding and initiation of transcription
   1. Promotor—site where RNA polymerase binds and where transcription begins (i.e., initiation site)
      a. Approximately 100 nucleotides in eukaryotes
      b. Includes nucleotide sequences that are recognized by specific DNA-binding proteins (transcription factors) that help initiate transcription
   2. Transcription factors
      a. In eukaryotes, RNA polymerase cannot recognize the promotor without transcription factors
      b. Transcription factors bind to TATA box
         i. TATA box is a region rich in thiamine and adenine located approximately 25 nucleotides upstream from the initiation site
         ii. RNA polymerase recognizes the complex formed by the binding of the transcription factor to the TATA box
         iii. Other transcription factors associate with the RNA polymerase/TATA sequence complex
   3. Activated RNA polymerase binds to the downstream promotor, separates the strands at the initiation site and transcription begins

D. Elongation of the RNA strand
   1. Functions of RNA polymerase
a. Untwists and opens DNA to reveal approximately 10 nucleotide bases—the exposed region is the template for mRNA
b. Links incoming RNA nucleotides to the 3’ end of the elongating strand
   i. 30-60 nucleotides/second
   ii. Newly formed mRNA peels away from the template spontaneously
   iii. DNA-DNA double helix re-forms; non-template strand pairs with the template
2. Several RNA pol II can simultaneously transcribe the gene

a. Allows an increased production of particular proteins
b. Length of each strand reflects how far the enzyme has traveled from the initiation site
E. Post-transcriptional modification

1. In eukaryotes, mRNA is modified before it leaves the nucleus
   a. Includes covalent alteration of both the 3’ and 5’ ends
   b. Removal of intervening sequences

2. Terms
   a. Primary transcript
      i. Initial RNA transcribed from DNA
   b. Pre-mRNA
      i. Primary transcript that will be processed to functional mRNA

3. Alteration of pre-mRNA ends
   a. 5’ cap—modified G (guanine tri-phosphate) that is added after transcription
      i. Protects from hydrolytic enzymatic degradation
      ii. Helps small ribosomal subunit recognize attachment site on mRNA 5’ end
   b. Leading sequence—non-coding (untranslated) sequence of mRNA
      i. From 5’ end to start codon
   c. Poly (A) tail
      i. Sequence of 30-200 adenines added to the 3’ end
   d. Trailer sequence
      i. Non-coding sequence of mRNA from the stop codon to poly (A) tail
4. Split genes and RNA splicing—in eukaryotes, genes that code for proteins may not be continuous sequences
   a. Introns (intervening sequences)—non-coding sequence in DNA that intervenes between coding sequences
      i. Transcribed but not translated
      ii. Excised before leaving nucleus
      iii. Not all genes have introns
   b. Exons—coding sequence of a gene that is transcribed and expressed
   c. RNA splicing—RNA processing that removes introns and joins exons
      i. Produces mature RNA that will pass into the cytoplasm
      ii. Not limited to mRNA, also t and r
   d. Small nuclear ribonucleoproteins (snRNPs—snurps)—complex of proteins and small nuclear RNA found in the nucleus; some participate in RNA splicing
      i. Small nuclear RNA is about 300 nucleotides
      ii. Protein

III. Synthesis of protein

A. Translation—RNA—directed synthesis of a polypeptide
   1. mRNA contains a genetic message consisting of sequential codons (mRNA codon codes for one amino acid)
   2. tRNA is the interpreter between the two forms of information—base sequence in mRNA and amino acid sequence in the polypeptide

(a) The 3-D shape of a tRNA molecule is determined by hydrogen bonds formed between complementary bases.

(b) One loop contains the anticodon; these unpaired bases pair with a complementary mRNA codon. The amino acid attaches to the terminal nucleotide at the hydroxyl (OH) 3' end.

(c) This stylized diagram of an aminoacyl-tRNA shows that the amino acid attaches to tRNA by its carboxyl group, leaving its amino group exposed for peptide bond formation.

© 2007 Thomaen Higher Education
a. tRNA aligns the appropriate amino acid to form a new polypeptide
   i. Collects appropriate amino acid from cytoplasmic pool
   ii. Recognizes codon in mRNA
   iii. Each tRNA recognizes only one amino acid and its unique mRNA codon
b. Physical properties of tRNA
   i. Ability to bind amino acid and read codon depends on its structure
   ii. Site that attaches to mRNA codon
   iii. Site that attaches to amino acid
c. Anticodon—nucleotide triplet in tRNA that base pairs with complementary triplet (codon) in mRNA

<table>
<thead>
<tr>
<th>tRNA Anticodon</th>
<th>mRNA Codon</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>UUU</td>
<td>Phenylalanine</td>
</tr>
</tbody>
</table>

d. There are 45 different tRNA’s but 64 codon ($4^3$)
   i. Some tRNA’s can recognize 2-3 mRNA codon specifying the same amino acid
   ii. Wobble—base pairing rules are relaxed between the $3^{rd}$ base of the mRNA and the corresponding anticodon
   iii. $5'$ end ($3^{rd}$ base) can hydrogen bond with more than one kind of base in the $3^{rd}$ position ($3'$) of the codon
   iv. Some tRNAs contain a modified base called inosine (I) which is in the anticodon’s wobble position
   v. Inosine can base pair with U, C and A (not G)

<table>
<thead>
<tr>
<th>tRNA Anticodon</th>
<th>mRNA Codon</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCI</td>
<td>GGU, GGA, GGC</td>
<td>Glycine</td>
</tr>
</tbody>
</table>

B Aminoacyl-tRNA synthetase
1. Amino acids must first be correctly linked with tRNA before the tRNA anticodon pairs with its complementary mRNA codon.
2. Aminoacyl-tRNA synthetase catalyzes the attachment of the amino acid to its tRNA
   a. Each amino acid has a unique aminoacyl-tRNA synthetase
   b. Reaction requires energy (i.e., endergonic reaction) provided by the hydrolysis of ATP
   c. Reaction forming aminoacyl-tRNA complex requires two steps:
      i. Activation of amino acid with AMP—enzyme binds AMP to amino acid with net release of pyrophosphate (i.e., ATP to AMP + pyrophosphate)
      ii. tRNA is covalently bonded to amino acid, displacing AMP
   d. Aminoacyl-tRNA complex releases from the enzyme and transfers its amino acid to growing polypeptide on the ribosome.
C. Ribosomes
1. Ribosomes coordinate the pairing of tRNA anticodons to mRNA codons.
2. Structure
   a. Two subunits—small and large
      i. Subunits are separated when not involved in protein synthesis
   b. Comprised of 60% ribosomal RNA (rRNA) and 40% protein
   c. Constructed in the nucleolus
   d. Enters cytoplasm through pores in the nucleus
   e. Assembled into functional ribosomes only when attached to mRNA
   f. Four binding sites:

   ![Diagram of ribosomes with binding sites](Image)

(b) The mRNA passes through a groove between the two ribosomal subunits. Each ribosome contains three binding sites for tRNA molecules.
i. mRNA  
ii. P site—holds the tRNA carrying the polypeptide  
iii. A site—holds the tRNA carrying the next amino acid to be added  
iv. E site—discharged tRNA exits from this site  
g. Ribosomes do not catalyze the synthesis of bonds between the amino acid and the growing polypeptide; this is done by other enzymes (see below)  

IV. Building a peptide  

A. Translation occurs in three steps:  
   1. Initiation  
   2. Elongation  
   3. Termination  

B. Initiation  
   1. Initiation brings together the mRNA and the first amino acid (always methionine) and the two ribosomal subunits  
      a. Small ribosomal subunit is bonded to mRNA and initiator tRNA  
      b. In eukaryotes, 5' cap of the mRNA aids in binding the leader sequence to the small ribosomal subunit  
      c. Small ribosomal subunit facilitates binding of initiator tRNA to the start code
The small ribosomal subunit binds to the mRNA at the AUG start codon. The leader sequence upstream of the AUG sequence helps the ribosome identify the AUG sequence.

i. AUG is the start code
ii. AUG is the codon for methionine (i.e., that’s why the first amino acid is always methionine)

d. *Initiation Complex*—small ribosomal subunit, initiator tRNA and mRNA
The initiator tRNA binds to the start codon, and one of the initiation factors is released.

Requires:

1. Initiator factors that are bound to small ribosomal subunit
2. One GTP molecule that when hydrolyzed drives the attachment of the large ribosomal subunit
3. Functional translational complex—large ribosomal subunit binds to initiation complex (i.e., to the small ribosome)
3 When the large ribosomal subunit binds to the small subunit and the remaining initiation factors are released, the initiation complex is complete.

i. Initiation factors attached to the small subunit are released, allowing the large subunit to bind with small subunit
ii. Initiator tRNA fits into the P site
iii. Vacant A site is ready for the next aminoacyl-tRNA

C. Elongation
1. Three step cycle to add amino acids one by one to the initial amino acid and subsequent polypeptide
1. **Codon recognition**
   a. mRNA codon in A site of the ribosome forms hydrogen bonds with the anticodon of the entering tRNA carrying the next amino acid

2. **Codon recognition**
   a. mRNA codon in A site of the ribosome forms hydrogen bonds with the anticodon of the entering tRNA carrying the next amino acid

3. **Peptide bond is formed**
   a. Peptide bond is formed between the amino acid in the P site and neighboring amino acid in the A site
   b. Formation of peptide bond is catalyzed by peptidyl transferase
   c. Polypeptide separates from its tRNA and is transferred to the new amino acid carried by the tRNA in the A site

4. **Translocation**
   a. Codon recognition
   b. Peptide bond is formed
   c. Translocation

5. **Translocation**
   a. Codon recognition
   b. Peptide bond is formed
   c. Translocation

6. **Translocation**
   a. Codon recognition
   b. Peptide bond is formed
   c. Translocation
4. Translocation—tRNA in the A site carrying the growing peptide is translocated to the P site while the tRNA in the P site is simultaneously translocated to the E site (where it exits the ribosome complex)
   a. mRNA and tRNA move as a unit maintaining codon and anticodon bonding and bringing the next codon into the A site
   b. mRNA is moved through the ribosome in the 5’ to 3’ direction
   c. Requires the hydrolysis of GTP for energy
   d. Elongation takes about 1/10 sec per amino acid

D. Termination
1. Elongation continues until a termination codon is reached

   - Termination codon (stop codon)—base triplet (codon) on mRNA that signals translation to stop
     a. UAA, UAG, UGA
     b. Stop codons do not code for amino acids

2. When reached, release factor binds to the codon and initiates the following events:
   a. Bond between the polypeptide and tRNA in the P site is hydrolyzed
   b. Polypeptide and tRNA are released from the ribosome
   c. Translocation complex dissociates

E. Polyribosomes
1. Cluster of ribosomes simultaneously translating an mRNA
2. Once the ribosome complex passes the initiation codon, a second ribosome can attach to the leader sequence
   a. Multiple ribosomes can translate an mRNA at once

V. Post-translational modification
1. Biological activity of a protein is conferred by its three-dimensional conformation.
2. Genes determine protein primary structure (i.e., linear sequence of amino acids)
3. Primary structure determines how a polypeptide chain will spontaneously coil and fold to form a three-dimensional shape (i.e., secondary and tertiary conformation)
4. Some proteins undergo post-translational modification to become functional
   a. Types of modification
      i. Chemical modification—sugars, lipids, phosphate groups and other molecules attached to specific amino acids in the polypeptide
      ii. Chain length modification—one or more amino acids may be cleaved from the end of a peptide
      iii. Single peptide may be divided into two or more pieces (e.g., proinsulin to insulin)
      iv. Two or more peptides may be joined as subunits of a protein and determine quaternary structure

VI. Point Mutations

A. Terms:
   1. Mutation—change in the genetic material of a cell
   2. Point mutation—mutation limited to one or a few base pairs in a single gene

B. Types of point mutations:
   1. Substitution
      a. Base-pair substitution—replacement of one base-pair with another; nucleotide and its partner in the complementary DNA are replaced with another pair of nucleotides according to base-pairing rules
         i. Silent substitution—because of redundancy in the genetic code, a base pair change may simply transform one codon into another that codes for the same amino acid
ii. **Conservative substitution**—change of base pair results in a change in amino acid but the new amino acid has similar properties or the substitution occurs in a non-essential area of the protein

iii. Base pair changes may change a single amino acid in a crucial area, altering protein activity either enhancing (rare) protein activity or impairing protein activity

b. Base-pair substitutions are usually *missense mutations* or *nonsense mutations*

i. Missense mutation—base-pair substitution that alters an amino acid codon (sense codon) to a new codon that codes for a different amino acid; these are translated; most common type of base-pair substitution

ii. Nonsense mutation—base-pair substitution that changes an amino acid codon to a termination codon, or vice versa; usually results in a non-functioning protein

2. Insertions or deletions—usually have negative effects
a. Base-pair insertion—insert one or more base-pairs into a gene
b. Base-pair deletion—removes one or more base-pairs out of a gene
c. If insertion or deletion is not a multiple of 3 nucleotides, the base-pair insertion or deletion results in a *frame shift mutation*
d. Frame shift mutation—insertion or deletion of base pairs that results in a change in reading frame
   i. Change causes nucleotide to be improperly grouped
   ii. Results in missense that will ultimately result in nonsense
   iii. Translated protein is usually not functional