Chapter 4.1

Nucleic Acids, the Genetic Code, and the Synthesis of Macromolecules
Genomes – genes – DNA – nucleotides

**Core components**
- DNA: deoxyNucleotides
- RNA: riboNucleotides

**Key concepts:**
- Fidelity
- Speed
- Accurate regulation

**Annex:**
- Deoxyribonucleic acid (DNA) contains the information prescribing the amino acid sequence of proteins
- This information is arranged in units termed genes
- Ribonucleic acid (RNA) serves in the cellular machinery that chooses and links amino acids in the correct sequence
- The central dogma: DNA -> RNA -> Protein
- DNA and RNA are polymers of nucleotide subunits

**Other types of RNA**
- Pre mRNA
- mRNA
- Coding sequence (CDS)
- Protein
This lecture:

- **Basics**

- **Transcription** (synthesis of RNA from a DNA template)

- **Translation** (translating an RNA message into primary protein structure)
All nucleotides have a common structure (pentose sugar phosphate + base)

**Basics**

The five principal bases

**PURINES**

- Adenine (A)
- Guanine (G)

**PYRIMIDINES**

- Thymine (T)
- Cytosine (C)
- Uracil (U)

A, G, T, C are present in DNA (DeoxyriboNucleic Acid)
A, G, U, C are present in RNA (RiboNucleic Acid)
Nucleotide subunits are linked together by phosphodiester bonds.
Greatest biological advance of the 20th Century!

In 1962 James Watson (1928– ), Francis Crick (1916–2004), and Maurice Wilkins (1916–2004) jointly received the Nobel Prize in medicine or physiology.
Native DNA is a double helix of complementary anti-parallel chains.

Right-handed DNA double helix (10.1 bp/turn)

Hydrogen bonding between complementary base pairs (A-T or G-C)
And Van der Waals forces between the stacks of bases hold the two strands together.
DNA can undergo reversible strand separation

Analysis of DNA denaturation:

(a) Absorption of 260 nm light:

- Single-stranded DNA
- Double-stranded DNA

(b) Percentage of G-C pairs vs. $T_m$ (°C)

$T_m$ = melting temperature
Many DNA molecules are circular and local unwinding of circular DNA can produce supercoiling.
RNA molecules are generally single-stranded and exhibit varied conformations.

**Primary structure:**

5’-AUGCCGUGACC-3’

**Secondary structure**

(a) Secondary structure

**Tertiary structure**

(b) Tertiary structure
Chain elongation of nucleic acids proceeds by sequential addition of monomeric subunits

Possible modifications of nucleic acids following chain formation

- Both DNA and RNA chains are produced by copying of template DNA strands
- Nucleic acid strands (poly-nucleotides) grow by the addition of one nucleotide at a time, and always in the 5' -> 3' direction
- RNA polymerases can initiate strand growth but DNA polymerases require a primer strand
- The primary poly-nucleotide product is often modified

+ chemical modification

Conformational modification (allosteri)
Example of conformational modification of DNA

TBP binds to the minor groove of specific DNA sequences rich in A and T, untwisting and sharply bending the double helix.

This conformational change is required for activation of almost a third of all eucaryotic genes.
**Some important gene-features**

**Terminology:**
- Nontemplate strand/upper strand/\(+\) strand/\textit{coding strand}/Watson strand
- Template strand/lower strand/\(-\) strand/noncoding strand/Crick strand

**Transcription**
Transcription (RNA synthesis)

The principle of the transcription bubble
Transcription of DNA
= RNA synthesis

**INITIATION**
- RNA polymerase binds to promoter sequence in duplex DNA. "Closed complex"
- Polymerase melts duplex DNA near transcription start site, forming a transcription bubble. "Open complex"
- Polymerase catalyzes phosphodiester linkage of two initial rNTPs.

**ELONGATION**
- Polymerase advances 3' → 5' down template strand, melting duplex DNA and adding rNTPs to growing RNA.

**TERMINATION**
- At transcription stop site, polymerase releases completed RNA and dissociates from DNA.
Gene organization, transcription, and translation in prokaryotes

Transcription

(a) Prokaryotes

E. coli genome

trp operon

Start site for trp mRNA synthesis

Transcription

trp mRNA

Start sites for protein synthesis

Translation

Proteins
Gene organization, transcription, and translation in eukaryotes
Overview of RNA processing in eukaryotes

- Capping
- Poly-adenylation
- Splicing (leaving only exons)
Splicing can give different proteins from the same gene (=alternative splicing)

NB: remember proteins: motifs and modules!
Many RNA polymerase complexes can transcribe the same gene simultaneously.
Translation (synthesis of Protein)

Translation involves:
- RNA (mRNA, tRNA, rRNA)
- Codons-anticodons
- Activated tRNAs (aa-tRNA)
- Initiation factors
- Ribosomes
- Termination factors
The three roles of RNA in protein synthesis

- Three types of RNA molecules perform different but complementary roles in protein synthesis (translation):
  - **Messenger RNA (mRNA)** carries information copied from DNA in the form of a series of three base “words” termed codons.
  - **Transfer RNA (tRNA)** deciphers the code and delivers the specified amino acid.
  - **Ribosomal RNA (rRNA)** associates with a set of proteins to form ribosomes, structures that function as protein-synthesizing machines.
The genetic code is a triplet code
( = always made of a string of 3-base-sequences (=codons))

<table>
<thead>
<tr>
<th>FIRST POSITION (5' END)</th>
<th>SECOND POSITION</th>
<th>THIRD POSITION (3' END)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>U</strong></td>
<td><strong>C</strong></td>
<td><strong>A</strong></td>
</tr>
<tr>
<td>Phe</td>
<td>Ser</td>
<td>Tyr</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
<td>Stop</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
<td>Stop</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td><strong>Pro</strong></td>
<td><strong>His</strong></td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>His</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>Gln</td>
</tr>
<tr>
<td>Leu (Met)*</td>
<td>Pro</td>
<td>Gln</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>Ile</strong></td>
<td><strong>Thr</strong></td>
</tr>
<tr>
<td><strong>Ile</strong></td>
<td>Thr</td>
<td>Asn</td>
</tr>
<tr>
<td><strong>Ile</strong></td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>Met (Start)</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td><strong>G</strong></td>
<td><strong>Val</strong></td>
<td><strong>Ala</strong></td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Asp</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>Val (Met)*</td>
<td>Ala</td>
<td>Glu</td>
</tr>
</tbody>
</table>

- And can be read in different frames:

```
5’ GCU UGU UUA CGA AUU A - mRNA
   Ala Cys Leu Arg Ile - Polypeptide
```

```
5’ GCC GUU UAC GAA UUA - mRNA
   Leu Val Tyr Glu Leu - Polypeptide
```
The folded structure of tRNA specifies its decoding function.
Translation is a two-step decoding process

** activation of tRNA **

** Codon-anticodon recognition **

(= decoding)

Net Result: Phenylalanine Is Selected by Its Codon
Ribosome structure in prokaryotes & eukaryotes

Translation

<table>
<thead>
<tr>
<th>Prokaryotic</th>
<th>rRNA</th>
<th>Proteins</th>
<th>Subunits</th>
<th>Assembled ribosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23S (2900 rNTs)</td>
<td>5S (120 rNTs)</td>
<td>Total: 31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16S (1500 rNTs)</td>
<td></td>
<td></td>
<td>50S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70S</td>
</tr>
<tr>
<td>Eukaryotic (vertebrate)</td>
<td>28S: 5.8S (4800 rNTs, 160 rNTs)</td>
<td>5S (120 rNTs)</td>
<td>Total: 50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18S (1900 rNTs)</td>
<td></td>
<td></td>
<td>60S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80S</td>
</tr>
</tbody>
</table>
Initiation of translation (eucaryotes)

Sequential addition of activating components to the 40S – eIF3 complex forms initiation complex

Scanning of the mRNA to position the initiation complex at the start-codon

Association of the large ribosomal subunit (60S) forms an 80S ribosome ready to translate the mRNA

Two initiation factors, eIF2 and eIF5 are GTP-binding proteins whose GTP is hydrolyzed during translation initiation.

*NB: the precise timing of release of the eIF’s is not yet well characterized*
Elongation of translation (eucaryotes)

A ternary complex carrying the second amino acid (aa₂) coded by the mRNA binds the ribosomal A site.

Hydrolysis of eIF1α-bound GTP → conformational change → large rRNA subunit catalyzes peptide bond between Met₁ and aa₂.

Hydrolysis of EF2-bound GTP → conformational change → translocation along the mRNA → tRNA^{Met} (empty) moves to E site, and the tRNA with the bound peptide to the P site.

The elongation complex is now ready for the next cycle (back to (1)).

In the second cycle the empty tRNA is released from the E (exit) site by the conformational change induced by the hydrolysis of eIF1α-bound GTP.
Termination of translation (eucaryotes)

When the ribosome reaches a stop-codon (AUG, UGA, UAA) **release factor** eRF1 enters the complex together with eRF3-GTP.

Hydrolysis of GTP → cleavage of the peptide chain from the tRNA in the P site and release of the tRNAs and the ribosomal subunits.
Model of ribosome

based on computer-derived images of cryoelectron microscopic images and on chemical cross-linking studies
Simultaneous translation by multiple ribosomes and their rapid recycling increases the efficiency of protein synthesis.

The native protein is then processed and folded as described in Chpt 3.