Chapter 22. Nucleic Acids

Sections

22.1 Types of Nucleic Acids
22.2 Nucleotides: Building Blocks of Nucleic Acids
22.3 Primary Nucleic Acid Structure
22.4 The DNA Double Helix
22.5 Replication of DNA Molecules
22.6 Overview of Protein Synthesis
22.7 Ribonucleic Acids

Chemistry at a Glance: DNA Replication
22.8 Transcription: RNA Synthesis
22.9 The Genetic Code
22.10 Anticodons and tRNA Molecules
22.11 Translation: Protein Synthesis
22.12 Mutations

Chemistry at a Glance: Protein Synthesis
22.13 Nucleic Acids and Viruses
22.14 Recombinant DNA and Genetic Engineering
22.15 The Polymerase Chain Reaction
22.16 DNA Sequencing

Why Nucleic Acid is Important to life?

- Cells in an organism are exact replicas
- Cells have information on how to make new cells
- Molecules responsible for such information are nucleic acids
  - Found in nucleus and are acidic in nature
  - A nucleic acid is a polymer in which the monomer units are nucleotides.
- Two Types of Nucleic Acids:
  - DNA: Deoxyribonucleic Acid: Found within cell nucleus
    - Storage and transfer of genetic information
    - Passed from one cell to other during cell division
  - RNA: Ribonucleic Acid: Occurs in all parts of cell
    - Primary function is to synthesize the proteins
**Nucleic Acids**

- Nucleic Acids: Polymers in which repeating unit is nucleotide
- A Nucleotide has three components:
  - Pentose Sugar: Monosaccharide
  - Phosphate Group ($\text{PO}_4^{3-}$)
  - Heterocyclic Base

**Nitrogen-Containing Heterocyclic Bases**

- There are a total five bases (four of them in most of DNA and RNAs)
- Three pyrimidine derivatives - thymine (T), cytosine (C), and uracil (U)
- Two purine derivatives - adenine (A) and guanine (G)
  - Adenine (A), guanine (G), and cytosine (C) are found in both DNA and RNA.
  - Uracil (U): found only in RNA
  - Thymine (T) found only in DNA.

**Pentose Sugar**

- Ribose is present in RNA and 2-deoxyribose is present in DNA
- Structural difference:
  - a $-\text{OH}$ group present on carbon 2’ in ribose
  - a $-\text{H}$ atom in 2-deoxyribose
- RNA and DNA differ in the identity of the sugar unit in their nucleotides.

**Phosphate**

- Phosphate - third component of a nucleotide, is derived from phosphoric acid ($\text{H}_3\text{PO}_4$)
- Under cellular pH conditions, the phosphoric acid is fully dissociated to give a hydrogen phosphate ion ($\text{HPO}_4^{2-}$)
**Nucleotide Formation**

- The formation of a nucleotide from sugar, base, and phosphate is visualized below.
- Phosphate attached to C-5' and base is attached to C-1' position of pentose.

**Nucleotide Nomenclature**

<table>
<thead>
<tr>
<th>Base</th>
<th>Sugar</th>
<th>Nucleotide Name</th>
<th>Nucleotide Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA Nucleotides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adenine</td>
<td>deoxyribose</td>
<td>deoxyadenosine 5'-monophosphate</td>
<td>dAMP</td>
</tr>
<tr>
<td>guanine</td>
<td>deoxyribose</td>
<td>deoxyguanosine 5'-monophosphate</td>
<td>dGMP</td>
</tr>
<tr>
<td>cytosine</td>
<td>deoxyribose</td>
<td>deoxycytidine 5'-monophosphate</td>
<td>dCMP</td>
</tr>
<tr>
<td>thymine</td>
<td>deoxyribose</td>
<td>deoxythymidine 5'-monophosphate</td>
<td>dTMP</td>
</tr>
<tr>
<td>RNA Nucleotides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adenine</td>
<td>ribose</td>
<td>adenine 5'-monophosphate</td>
<td>AMP</td>
</tr>
<tr>
<td>guanine</td>
<td>ribose</td>
<td>guanosine 5'-monophosphate</td>
<td>CMP</td>
</tr>
<tr>
<td>cytosine</td>
<td>ribose</td>
<td>cytidine 5'-monophosphate</td>
<td>CMP</td>
</tr>
<tr>
<td>uracil</td>
<td>ribose</td>
<td>uridine 5'-monophosphate</td>
<td>UMP</td>
</tr>
</tbody>
</table>

**Backbone structure for nucleic acid**

(a) The generalized structure of a nucleic acid. (b) The specific backbone structure for a deoxyribonucleic acid (DNA). (c) The specific backbone structure for a ribonucleic acid (RNA).

**Molecule of Adenine, a nitrogen-containing heterocyclic base present in both RNA and DNA.**
Two purine bases and three pyrimidine bases are found in the nucleotides present in nucleic acids.

<table>
<thead>
<tr>
<th>Base</th>
<th>Sugar</th>
<th>Nucleotide Name</th>
<th>Nucleotide Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA Nucleotides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adenine</td>
<td>deoxyribose</td>
<td>deoxyadenosine 5'-monophosphate</td>
<td>dAMP</td>
</tr>
<tr>
<td>guanine</td>
<td>deoxyribose</td>
<td>deoxyguanosine 5'-monophosphate</td>
<td>dGMP</td>
</tr>
<tr>
<td>cytosine</td>
<td>deoxyribose</td>
<td>deoxycytidine 5'-monophosphate</td>
<td>dCMP</td>
</tr>
<tr>
<td>thymine</td>
<td>deoxyribose</td>
<td>deoxythymidine 5'-monophosphate</td>
<td>dTMP</td>
</tr>
<tr>
<td>RNA Nucleotides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adenine</td>
<td>ribose</td>
<td>adenosine 5'-monophosphate</td>
<td>AMP</td>
</tr>
<tr>
<td>guanine</td>
<td>ribose</td>
<td>guanosine 5'-monophosphate</td>
<td>CMP</td>
</tr>
<tr>
<td>cytosine</td>
<td>ribose</td>
<td>cytidine 5'-monophosphate</td>
<td>CMP</td>
</tr>
<tr>
<td>uracil</td>
<td>ribose</td>
<td>uridine 5'-monophosphate</td>
<td>UMP</td>
</tr>
</tbody>
</table>

Fig. 22.3
The general structure of a nucleic acid in terms of nucleotide subunits.
A four-nucleotide-long segment of DNA.

Upon DNA replication the large DNA molecules interacts with histone proteins to fold long DNA molecules.

The histone–DNA complexes are called chromosomes:
- A chromosome is about 15% by mass DNA and 85% by mass protein.
- Cells of different kinds of organisms have different numbers of chromosomes.
- Example: Number of chromosomes in a human cell 46, a mosquito 6, a frog 26, a dog 78, and a turkey 82
- Chromosomes occur in matched (homologous) pairs.
  - Example: The 46 chromosomes of a human cell constitute 23 homologous pairs

A comparison of the primary structures of nucleic acids and proteins.

A schematic drawing of the DNA double helix that emphasizes the hydrogen bonding between bases on the two chains.
Hydrogen bonding in Base Pairs

DNA replication

DNA replication usually occurs at multiple sites within a molecule, and the replication is bidirectional from these sites.

One strand of DNA grows continuously in the direction of the unwinding, and the other grows in the opposite direction.

DNA Replication

DNA replication at multiple sites
What in Common Twins Have?

Identical twins share identical physical characteristics because they received identical DNA from their parents.

DNA replication… cont’d

Protein synthesis is directly under the direction of DNA
Proteins are responsible for the formation of skin, hair, enzymes, hormones, and so on
Protein synthesis can be divided into two phases.
- Transcription – A process by which DNA directs the synthesis of mRNA molecules
- Translation – A process in which mRNA is deciphered to synthesize a protein molecule

Differences Between RNA and DNA Molecules
- The sugar unit in the backbone of RNA is ribose; it is deoxyribose in DNA.
- The base thymine found in DNA is replaced by uracil in RNA
- RNA is a single-stranded molecule; DNA is double-stranded (double helix)
- RNA molecules are much smaller than DNA molecules, ranging from 75 nucleotides to a few thousand nucleotides
Types of RNA Molecules

- Heterogeneous nuclear RNA (hnRNA): Formed directly by DNA transcription.
- Post-transcription processing converts the hnRNA to mRNA
- Messenger RNA: Carries instructions for protein synthesis (genetic information) from DNA
  - The molecular mass of mRNA varies with the length of the protein
- Small nuclear RNA: Facilitates the conversion of hnRNA to mRNA.
  - Contains from 100 to 200 nucleotides
- Ribosomal RNA (rRNA): Combines with specific proteins to form ribosomes - the physical site for protein synthesis. Ribosomes have molecular masses on the order of 3 million

Types of RNA Molecules

- Transfer RNA (tRNA): Delivers amino acids to the sites for protein synthesis
  - tRNAs are the smallest (75–90 nucleotide units)

Transcription

- Transcription: A process by which DNA directs the synthesis of mRNA molecules
  - Two-step process - (1) synthesis of hnRNA and (2) editing to yield mRNA molecule
- Gene: A segment of a DNA base sequence responsible for the production of a specific hnRNA/mRNA molecule
  - Most human genes are ~1000–3500 nucleotide units long
  - Genome: All of the genetic material (the total DNA) contained in the chromosomes of an organism
  - Human genome is about 20,000–25,000 genes

Steps in the Transcription Process

- Unwinding of DNA double helix to expose some bases (a gene):
  - The unwinding process is governed by RNA polymerase
- Alignment of free ribonucleotides along the exposed DNA strand (template) forming new base pairs
- RNA polymerase catalyzes the linkage of ribonucleotides one by one to form mRNA molecule
- Transcription ends when the RNA polymerase enzyme encounters a stop signal on the DNA template:
  - The newly formed RNA molecule and the RNA polymerase enzyme are released
Post-Transcription Processing:
Formation of mRNA

- Involves conversion of hnRNA to mRNA

- Splicing: Excision of introns and joining of exons
  - Exon - a gene segment that codes for genetic information
  - Intron – a DNA segments that interrupt a genetic message
  - The splicing process is driven by snRNA

- Alternative splicing - A process by which several different protein variants are produced from a single gene
  - The process involves excision of one or more exons

Exons and Introns of RNA

Heterogenous nuclear RNA contains both exons and introns.

Transcriptome

- Transcriptome: All of the mRNA molecules that can be generated from the genetic material in a genome.
  - Transcriptome is different from a genome
  - Responsible for the biochemical complexity created by splice variants obtained by hnRNA.

RNA hairpin loop

A hairpin loop is produced when a single-stranded RNA doubles back on itself and complimentary base pairing occurs.
Classification of RNA

According to the function of RNA, it can be classified as:

**Messenger RNA** (m-RNA) synthesized on chromosome and carries genetic information to the ribosomes for protein synthesis. It has short half-life.

**Transfer RNA** (t-RNA) is a relatively small and stable molecule that carries a specific amino acid from the cytoplasm to the site of protein synthesis on ribosomes.

**Ribosomal RNA** (r-RNA) is the major component of ribosomes, constituting nearly 65%. r-RNA is responsible for protein synthesis.

**Ribozymes** are RNA molecules that have catalytic properties.

Types of RNA

Types of RNA include:

- **DNA**: The genetic material of all living organisms.
- **mRNA**: Messenger RNA, which carries genetic information from DNA to the ribosomes.
- **tRNA**: Transfer RNA, which carries specific amino acids to the ribosomes.
- **rRNA**: Ribosomal RNA, which is part of ribosomes.
- **hnRNA**: Heterogeneous nuclear RNA, which is the precursor to mRNA.
- **mRNA**: Messenger RNA, which is the final product of transcription.
- **tRNA**: Transfer RNA, which carries specific amino acids.
- **rRNA**: Ribosomal RNA, which is part of ribosomes.

**Transcription of DNA to form RNA**

The transcription of DNA to form RNA involves an unwinding of a portion of the DNA double helix.

**Exons and Introns of RNA cont'd**

An hnRNA molecule containing four exons.
### Codes for Amino Acids

<table>
<thead>
<tr>
<th>First Position (5’ end)</th>
<th>Second Position</th>
<th>Third Position (3’ end)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>Pro</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>C</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>A</td>
<td>Ile</td>
<td>Thr</td>
</tr>
<tr>
<td>Ile</td>
<td>Thr</td>
<td>Ala</td>
</tr>
<tr>
<td>Ile</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>G</td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td>Val</td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td>Val</td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td>Val</td>
<td>Val</td>
<td>Ala</td>
</tr>
</tbody>
</table>

### Aminoacyl-tRNA synthetase

An aminoacyl-tRNA synthetase has an active site for tRNA and a binding site for the particular amino acid that is to be attached to that tRNA.

### Anticodon and Codon

The interaction between anticodon and codon.

### tRNA molecule

A tRNA molecule
Ribosomes have structures that contain two subunits.

Protein Synthesis: Initiation

Initiation of protein synthesis begins with the formation of an initiation complex.

Protein Synthesis: Translation

The process of translation that occurs during protein synthesis.

Effects of Antibiotics

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Biological action</th>
</tr>
</thead>
<tbody>
<tr>
<td>chloramphenicol</td>
<td>inhibits an important enzyme (peptidyl transferase) in the large ribosomal subunit</td>
</tr>
<tr>
<td>erythromycin</td>
<td>binds to the large subunit and stops the ribosome from moving along the mRNA from one codon to the next</td>
</tr>
<tr>
<td>puromycin</td>
<td>induces premature polypeptide chain termination</td>
</tr>
<tr>
<td>streptomycin</td>
<td>inhibits initiation of protein synthesis and also causes the mRNA codons to be read incorrectly</td>
</tr>
<tr>
<td>tetracycline</td>
<td>binds to the small ribosomal subunit and inhibits the binding of incoming tRNA molecules</td>
</tr>
</tbody>
</table>
Several ribosomes can simultaneously proceed along a single strand of mRNA. Such a complex of mRNA and ribosomes is called a polysome.

**Protein Synthesis Summary**

Recombinant DNA is made by inserting a gene obtained from DNA of one organism into the DNA from another kind of organism.
Cleavage patterns resulting from the use of a restriction enzyme that cleaves DNA between G and A bases.

Cleaving DNA patterns using restriction enzymes

The “sticky ends” of the cut plasmid and the gene are complementary and combine to form recombinant DNA.

“sticky ends” of recombinants

Polymerase chain reaction process

Step 1. A DNA solution is heated to cause the base-paired double helix to unwind into single strands.

Step 2. Primers complementary to the DNA on either side of the target area of the single-stranded DNA are added.

Polymerase chain reaction process

Step 3. DNA polymerase is used to extend the primers to create segments of DNA identical to the original segment.

Step 4. The process is repeated for as many cycles as necessary, and in a short time millions of identical DNA molecules have been produced.
**Summary of Nucleic Acids**

**DNA**
- DNA contains genetic information.
- DNA contains adenine (A) and guanine (G), and thymine (T), and cytosine (C). A-T G-C
- DNA has a double helical structure.
- The bases in DNA carry the genetic information.

**RNA**
- RNA functions as genetic information-carrying intermediates in protein synthesis.
- It contains adenine (A) and guanine (G), and cytosine (C) and uracil (U).
- m-RNA carries genetic information from DNA to the ribosomes for protein synthesis.
- t-RNA transfers amino acid to the site of protein synthesis
- r-RNA is for protein synthesis.

**Selected steps in the DNA sequencing procedure for the 10-base DNA segment 5' AGCAGCTGGT 3'**

**Nucleotides**
- Nucleotides are basic units of nucleic acids DNA and RNA.
- Nucleotides include pentose, base and phosphoric acid.
- Bases include purine or pyrimidine.
- Two major purines present in nucleotides are adenine (A) and guanine (G), and three major pyrimidines are thymine (T), cytosine (C) and uracil (U).
- Ribonucleotides
  - adenosine triphosphate (ATP) stores energy.
  - NAD and NADP are important carriers of reducing power.
### Summary of Cell Construction

<table>
<thead>
<tr>
<th>Biopolymers</th>
<th>protein</th>
<th>Carbohydrates (polysaccharides)</th>
<th>DNA</th>
<th>RNA</th>
<th>lipids</th>
</tr>
</thead>
<tbody>
<tr>
<td>subunit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bonds for subunit linkage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Characteristic three-D structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Primary Structure

- A ribonucleic acid (RNA) is a nucleotide polymer in which each of the monomers contains ribose, a phosphate group, and one of the heterocyclic bases adenine, cytosine, guanine, or uracil.
- A deoxyribonucleic acid (DNA) is a nucleotide polymer in which each of the monomers contains deoxyribose, a phosphate group, and one of the heterocyclic bases adenine, cytosine, guanine, or thymine.

### Primary Structure

- **Structure**: Sequence of nucleotides in DNA or RNA
- **Primary structure** is due to changes in the bases
- Phosphodiester bond at 3' and 5' position
- 5' end has free phosphate and 3' end has a free OH group
- Sequence of bases read from 5' to 3'

### Comparison of the General Primary Structures of Nucleic Acids and Proteins

- **Backbone**: -Phosphate-Sugar- Nucleic acids
- **Backbone**: -Peptide bonds - Proteins

![Comparison of the General Primary Structures of Nucleic Acids and Proteins](image-url)
Nucleic acids have secondary and tertiary structure
- The secondary structure involves two polynucleotide chains coiled around each other in a helical fashion
- The polynucleotides run anti-parallel (opposite directions) to each other, i.e., 5' - 3' and 3' - 5'
- The bases are located at the center and hydrogen bonded (A=T and G=C)
- Base composition: %A = %T and %C = %G
  - Example: Human DNA contains 30% adenine, 30% thymine, 20% guanine and 20% cytosine

DNA Sequence: the sequence of bases on one polynucleotide is complementary to the other polynucleotide
- Complementary bases are pairs of bases in a nucleic acid structure that can hydrogen-bond to each other.
- Complementary DNA strands are strands of DNA in a double helix with base pairing such that each base is located opposite its complementary base.
- Example:
  - List of bases in sequential order in the direction from the 5' end to 3' end of the segment:
  - 5'-A-A-G-C-T-A-G-C-T-T-3'
  - Complementary strand of this sequence will be: 3'-T-T-C-G-A-T-C-G-A-T-G-A-5'

Base Pairing
- One small and one large base can fit inside the DNA strands:
  - Hydrogen bonding is stronger with A-T and G-C
  - A-T and G-C are called complementary bases

Practice Exercise
- Predict the sequence of bases in the DNA strand complementary to the single DNA strand shown below:

  5' A-A-T-G-C-A-G-C-T 3'

  Answer:
  3' T-T-A-C-G-T-C-G-A 5'
Replication: Process by which DNA molecules produce exact duplicates of themselves
- Old strands act as templates for the synthesis of new strands
- DNA polymerase checks the correct base pairing and catalyzes the formation of phosphodiester linkages
- The newly synthesized DNA has one new DNA strand and old DNA strand

DNA polymerase enzyme can only function in the 5'-to-3' direction
- Therefore one strand (top; leading strand) grows continuously in the direction of unwinding
- The lagging strand grows in segments (Okazaki fragments) in the opposite direction
- The segments are latter connected by DNA ligase
- DNA replication usually occurs at multiple sites within a molecule (origin of replication)
- DNA replication is bidirectional from these sites (replication forks)
- Multiple-site replication enables rapid DNA synthesis

Characteristics of Genetic Code

- The genetic code is highly degenerate:
  - Many amino acids are designated by more than one codon.
  - Arg, Leu, and Ser - represented by six codons.
  - Most other amino acids - represented by two codons
  - Met and Trp - have only a single codon.
  - Codons that specify the same amino acid are called synonyms
- There is a pattern to the arrangement of synonyms in the genetic code table:
  - All synonyms for an amino acid fall within a single box in unless there are more than four synonyms
  - The significance of the “single box” pattern - the first two bases are the same
  - For example, the four synonyms for Proline - CCU, CCC, CCA, and CCG.

- The genetic code is almost universal:
  - With minor exceptions the code is the same in all organisms
  - The same codon specifies the same amino acid whether the cell is a bacterial cell, a corn plant cell, or a human cell.
- An initiation codon exists:
  - The existence of “stop” codons (UAG, UAA, and UGA) suggests the existence of “start” codons.
  - The codon - coding for the amino acid methionine (AUG) functions as initiation codon.
Practice Exercise

Sections A, C, and E of the following base sequence section of a DNA template strand are exons, and sections B and D are introns.

DNA 5' CGC - CGT - AGT - TGG - CCC - GGA - GGA 3'
A B C D E

a. What is the structure of the hRNA transcribed from this template?
b. What is the structure of the mRNA obtained by splicing the hRNA?

Answers:
a. 3' GCG – GCA – UCA – ACC – GGG – CCU – CCU 5'
b. 3' GCG – ACC – CCU – CCU 5'

During protein synthesis amino acids do not directly interact with the codons of an mRNA molecule.
• tRNA molecules as intermediaries deliver amino acids to mRNA.
• Two important features of the tRNA structure
  • The 3' end of tRNA is where an amino acid is covalently bonded to the tRNA.
  • The loop opposite to the open end of tRNA is the site for a sequence of three bases called an anticodon.
• Anticodon -- a three-nucleotide sequence on a tRNA molecule that is complementary to a codon on an mRNA molecule.

Translation – a process in which mRNA codons are deciphered to synthesize a protein molecule
• Ribosome – an rRNA–protein complex - serves as the site of protein synthesis:
  • Contains 4 rRNA molecules and ~80 proteins - packed into two rRNA-protein subunits (one small and one large)
  • ~65% rRNA and 35% protein by mass
  • A ribosome’s active site – Large subunit
  • Ribosome is a RNA catalyst
  • The mRNA binds to the small subunit of the ribosome.

Five Steps of Translation Process
• Activation of tRNA: addition of specific amino acids to the 3'-OH group of tRNA.
• Initiation of protein synthesis: Begins with binding of mRNA to small ribosomal subunit such that its first codon (initiating codon AUG) occupies a site called the P site (peptidyl site)
• Elongation: Adjacent to the P site in an mRNA–ribosome complex is A site (aminoacyl site) and the next tRNA with the appropriate anticodon binds to it.
• Termination: The polypeptide continues to grow via translocation until all necessary amino acids are in place and bonded to each other.
• Post-translational processing – gives the protein the final form it needs to be fully functional
Efficiency of mRNA Utilization

- **Polysome (polyribosome):** complex of mRNA and several ribosomes
- Many ribosomes can move simultaneously along a single mRNA molecule
- The multiple use of mRNA molecules reduces the amount of resources and energy that the cell expends to synthesize needed protein.
- In the process – several ribosomes bind to a single mRNA - polysomes.

Mutation

- An error in base sequence reproduced during DNA replication
- Errors in genetic information is passed on during transcription.
- The altered information can cause changes in amino acid sequence during protein synthesis and thereby alter protein function.
- Such changes have a profound effect on an organism.

Mutagens

- Mutations are caused by mutagens
- A mutagen is a substance or agent that causes a change in the structure of a gene:
  - Radiation and chemical agents are two important types of mutagens
  - Ultraviolet, X-ray, radioactivity and cosmic radiation are mutagenic – cause cancers
  - Chemical agents can also have mutagenic effects
    - E.g., HNO₃ can convert cytosine to uracil
    - Nitrites, nitrates, and nitrosamines – can form nitrous acid in cells
- Under normal conditions mutations are repaired by repair enzymes

Viruses

- Viruses: Tiny disease causing agents with outer protein envelope and inner nucleic acid core
- They can not reproduce outside their host cells (living organisms)
- Invade their host cells to reproduce and in the process disrupt the normal cell’s operation
- Virus invade bacteria, plants animals, and humans:
  - Many human diseases are of viral origin, e.g. Common cold, smallpox, rabies, influenza, hepatitis, and
Vaccines

- Inactive virus or bacterial envelope
- Antibodies produced against inactive viral or bacterial envelopes will kill the active bacteria and viruses

Viruses

- Viruses attach to the host cell on the outside cell surface and proteins of virus envelope catalyze the breakdown of the cell membrane and forms a hole
- Viruses then inject their DNA or RNA into the host cell
- The viral genome is replicated, proteins coding for the viral envelope are produced in hundreds of copies.
- Hundreds of new viruses are produced using the host cell replicated genome and proteins in short time

Recombinant DNA Production using a Bacterial Plasmid

- DNA molecules that have been synthesized by splicing a sequence of segment DNA (usually a gene) from one organism to the DNA of another organism
- Genetic Engineering (Biotechnology):
  - The study of biochemical techniques that allow the transfer of a “foreign” gene to a host organism and produce the protein associated with the added gene
  - Bacterial strains such as E. coli inserted with circular plasmids, and/or yeast cells carrying vectors containing foreign genes are used for this purpose
  - Plasmids (double stranded DNA) replicate independently in bacteria or yeast

- Dissolution of cells:
  - E. coli cells of a specific strain containing the plasmid of interest are treated with chemicals to dissolve their membranes and release the cellular contents

- Isolation of plasmid fraction:
  - The cellular contents are fractionated to obtain plasmids

- Cleavage of plasmid DNA:
  - Restriction enzymes are used to cleave the double-stranded DNA

- Gene removal from another organism:
  - Using the same restriction enzyme the gene of interest is removed from a chromosome of another organism

- Gene–plasmid splicing:
  - The gene (from Step 4) and the opened plasmid (from Step 3) are mixed in the presence of the enzyme DNA ligase to splice them together.

- Uptake of recombinant DNA:
  - The recombinant DNA prepared in step 5 are transferred to a live E. coli culture where they can be replicated, transcribed and translated.
Clones

- Transformed cell can reproduce a large number of identical cells – clones:
  - Clones are the cells that have descended from a single cell and have identical DNA
- Given bacteria grow very fast, within few hours 1000s of clones will be produced
- Each clone can synthesize the protein directed by foreign gene it carries

The polymerase chain reaction (PCR)

- The polymerase chain reaction (PCR) is a method for rapidly producing multiple copies of a DNA nucleotide sequence (gene).
- This method allows to produce billions of copies of a specific gene in a few hours.
- PCR is very easy to carryout and the requirements are:
  - Source of gene to be copied
  - Thermostable DNA polymerase
  - Deoxynucleotide triphosphates (dATP, dGTP, dCTP and dTTP)
  - A set of two oligonucleotides with complementary sequence to the gene (primers)
  - Thermostable plastic container and
  - Source of heat

DNA sequencing

- DNA sequencing is a method by which the base sequence in a DNA molecule (or a portion of it) is determined.
- Discovered in 1977 by Fredrick Sanger
- Concept in DNA sequencing:
  - Selective interruption of polynucleotide synthesis using 2',3'-dideoxyribonucleotide triphosphates (ddNTPs).

ddNTPs Fragments

- This interruption of synthesis leads to the formation of every possible nucleotide site mixture.
- These nucleotides are labeled using radioactive dNTP during their synthesis.
- The radiolabeled nucleotides are then separated on a gel by electrophoresis
Basic steps involved in DNA sequencing

- **Step 1:** Cleavage of DNA using restriction enzymes: Restriction enzymes are used to cleave the large DNA molecule into smaller fragments (100–200 base pairs).

- **Step 2:** Separation into individual components: The mixture of small DNA fragments generated by the restriction enzymes is separated into individual components via gel electrophoresis techniques.

- **Step 3:** Separation into single strands: A given DNA fragment is separated into its two strands by chemical methods to use it as a template in step 4.