Updated: 4 October 2019

Print version

CEE 370 Environmental Engineering Principles

Lecture #14

Environmental Biology III: Cell structure/function

Reading: Mihelcic & Zimmerman, Chapter 5

Davis & Masten, Chapter 3

Genotype vs Phenotype

- An organism's **genotype** is the set of genes that it carries; its genetic makeup
  - Inscribed in a code in DNA molecules. Every cell contains a complete genetic description of the whole organism encompassing it. Moreover, they do not merely describe but are part of an elaborate cellular machinery to cause a body part to develop the form of those descriptions.
- An organism's **phenotype** is all of its observable characteristics—which are influenced both by its genotype and by the environment
  - The genotype is a major determinant of the phenotypic attributes of the organism. But, genes are not exclusively responsible for a person's phenotype. The environment also plays an essential role. In general phenotypic traits are specified or "determined" by a combination of genetic and environmental factors

Genomes

- the term *genome* refers to the complete complement of DNA for a given species
  - the human genome consists of 46 chromosomes.
- every cell (except sex cells and mature red blood cells) contains the complete genome of an organism
  - Cells from the different parts of an organism have the same DNA
    - Distinction: The portion of the DNA that is transcribed and translated into protein

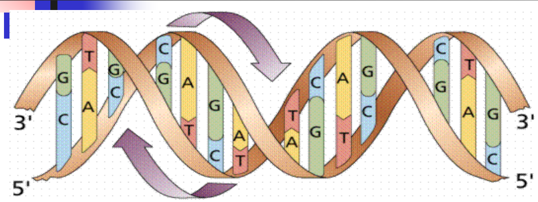
Cells, genome, gene and DNA

- Overall function of genome: Control the generation of molecules (mostly proteins) that will
  - Regulate the metabolism of a cell and its response to the environment, and
  - Provide structural integrity.
- Analogy
  - Nucleotide => letter
  - Gene => sentence
  - Contig => chapter
  - Chromosome => book
    - Gender, hair/eye color, …
    - Disorders: down syndrome, turner syndrome, …
    - Chromosome number varies for species
    - We have 46 (23 + 23) chromosomes
  - Complete genome => volumes of encyclopedia

Functions of Genes

- **Signal transduction:** sensing a physical signal and turning into a chemical signal
- **Structural support:** creating the shape and pliability of a cell or set of cells
- **Enzymatic catalysis:** accelerating chemical transformations otherwise too slow.
- **Transport:** getting things into and out of separated compartments
- **Movement**: contracting in order to pull things together or push things apart.
- **Transcription control**: deciding when other genes should be turned ON/OFF
- **Trafficking**: affecting where different elements end up inside the cell

A gene codes for a protein



DNA

transcription

mRNA

translation

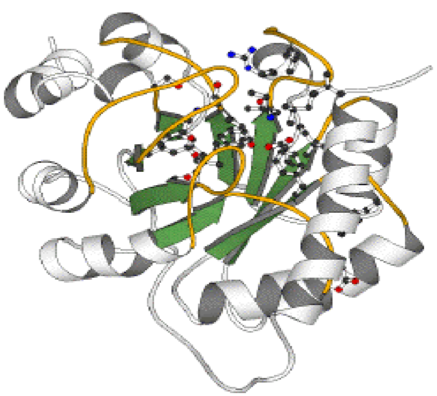
Protein

CCTGAGCCAACTATTGATGAA

CCUGAGCCAACUAUUGAUGAA

Condon (3 bases) codes for one amino acid

PEPTIDE



Exons & introns

- Most eukaryotic genes have exons (portions that will be put in the mRNA) and introns (that are *normally* spliced out and not in mRNA)
  - Some introns may have a promoter-like control of the transcription process
  - If an intron is not spliced out then an alternative splicing product is created.
  - Various tissue types can flexibly alter their gene products through alternative splicing
- Post-splicing (in Eukaryotes)
  - The generated mRNA is exported (through nuclear pore complexes) to the cytoplasm
  - In the cytoplasm, the ribosonal complex (containing hundreds of proteins and special function RNA molecules) acts to generate the protein on the basis of the mRNA code.

Introns and Exons

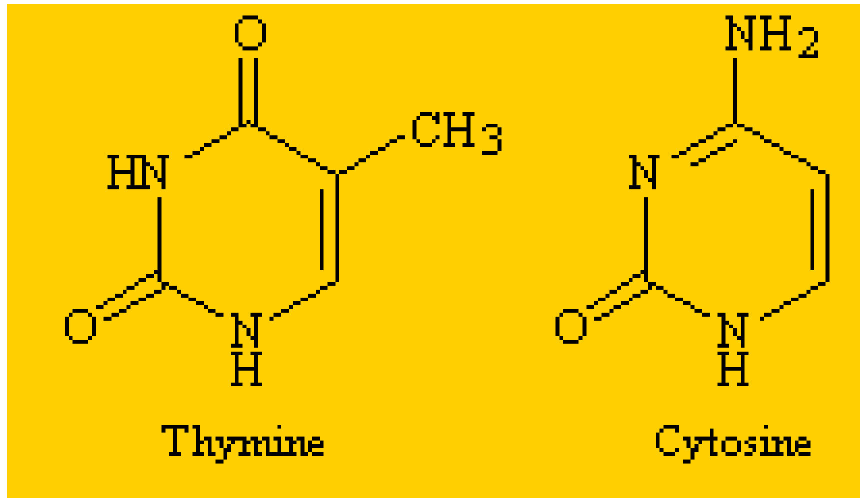
- Humans have about 35,000 genes containing 40,000,000 DNA bases
  - But this is only ~3% of total DNA in genome.
- Remaining 2,960,000,000 bases for control information. (e.g. when, where, how long, etc...)

Structure of DNA

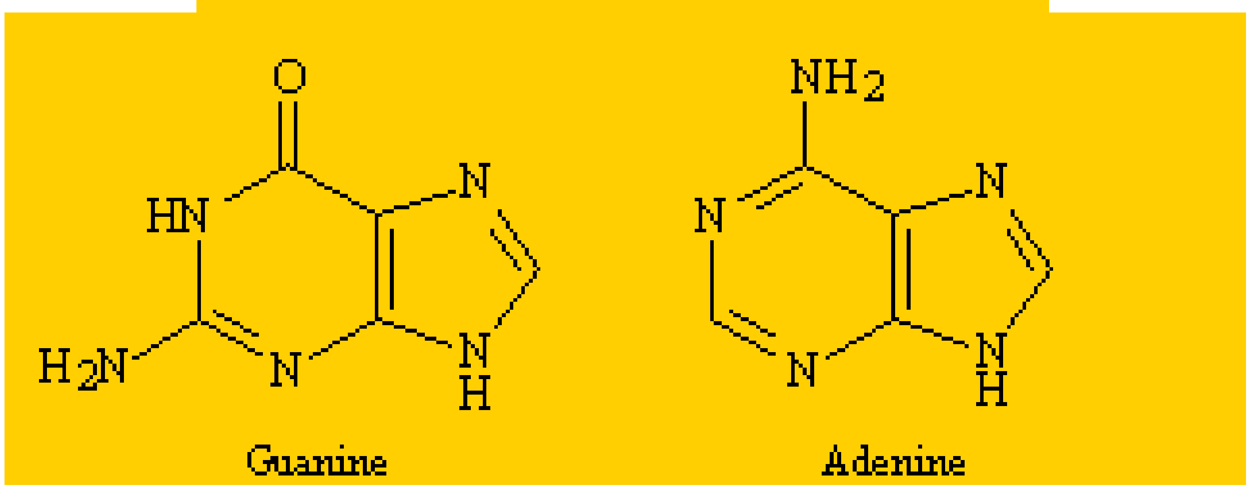
- Made up of 4 different building blocks (so called nucleotide bases), each an almost planar nitrogenic organic compound
  - Adenine (A)
  - Thymine (T)
  - Guanine (G)
  - Cytosine (C)
  - Base pairs (A -- T, C -- G)

Chemical Structure of Nucleotides

Purines

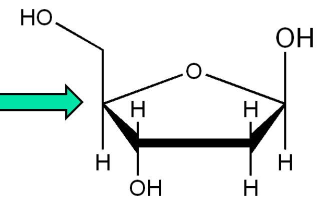


Pyrmidines



Structure of DNA -- 2

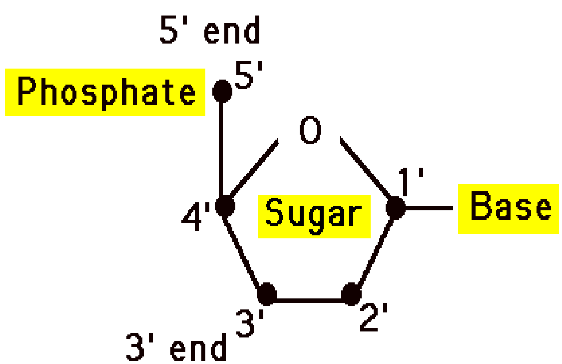
- Base pairs (A -- T,C -- G) are attached to a sugar phosphate backbone to form one of 2 strands of a DNA molecule.
  - Phosphate ($(PO_4)^{-3}$)
  - Deoxyribose



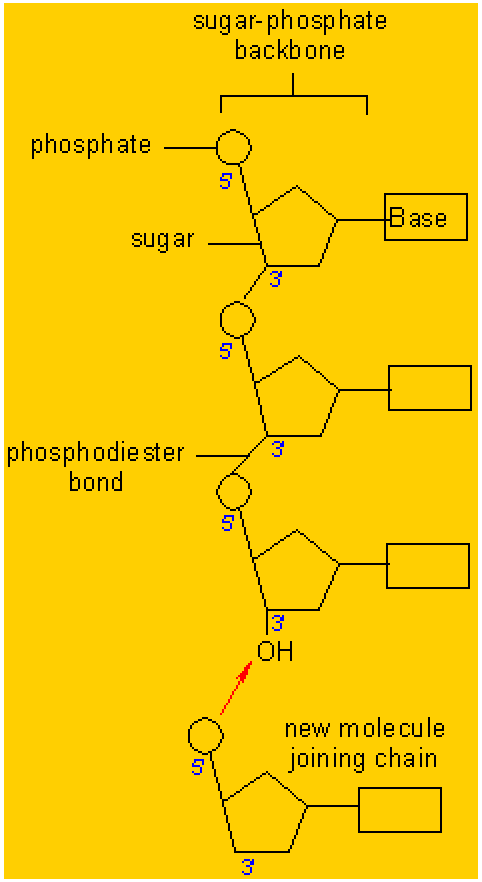
- Two strands are bonded together by the base pairs (A – T, C – G).
- Results in mirror image or complementary strands, each is twisted (or helical), and when bonded they form a double helix.
- Direction of each strand (5′ meaning beginning or 3′ meaning end of the strand)
  - 5′ and 3′ refer to position of bases in relation to the sugar molecule in the DNA backbone.
  - Are important reference points to navigate the genome.
  - 2 complementary strands are oriented in opposite direction to each other.

Nucleotides to Long Chains

- Nucleotides comprise
  - A base
  - A pentose sugar
  - A phosphate

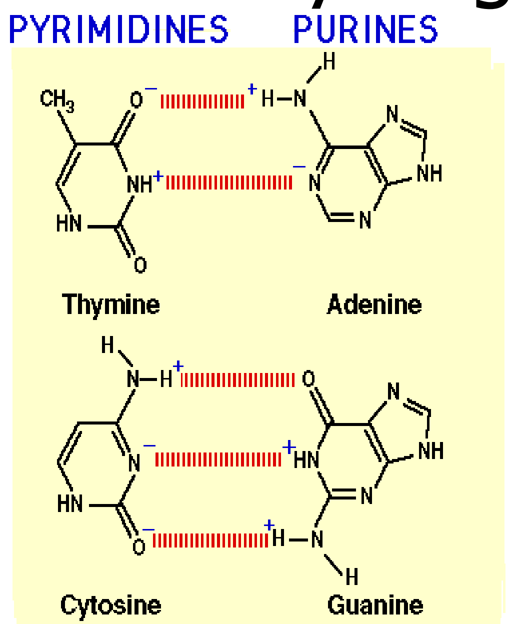


nucleotide

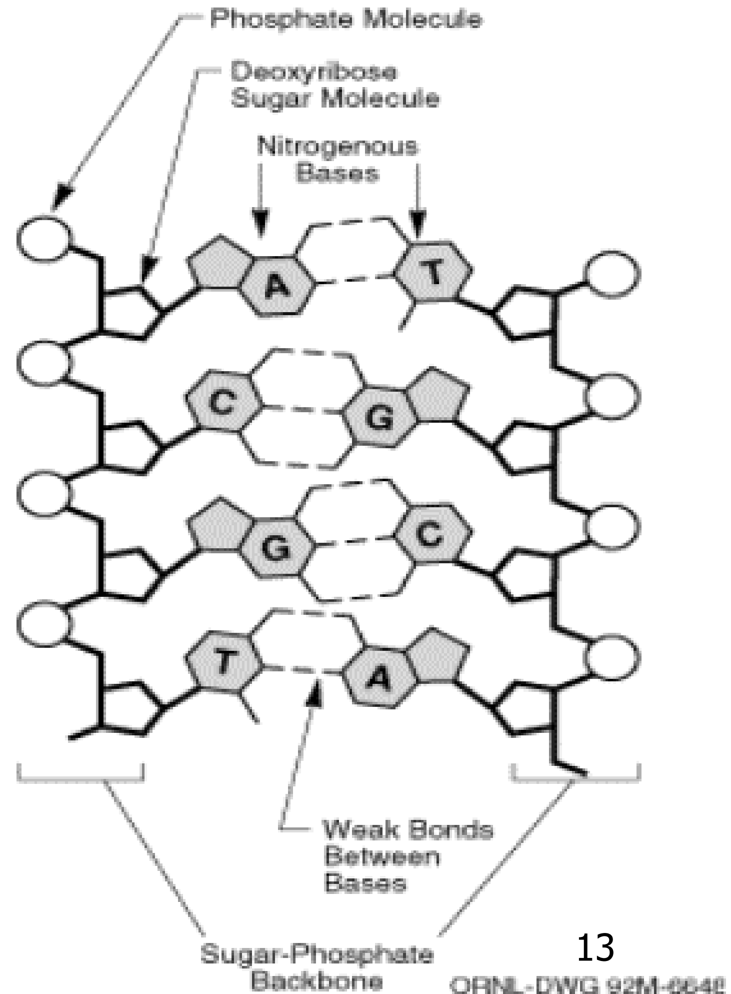


Long Chains to double helix

- Base pairs form hydrogen bonds
  - T=A
  - C=G

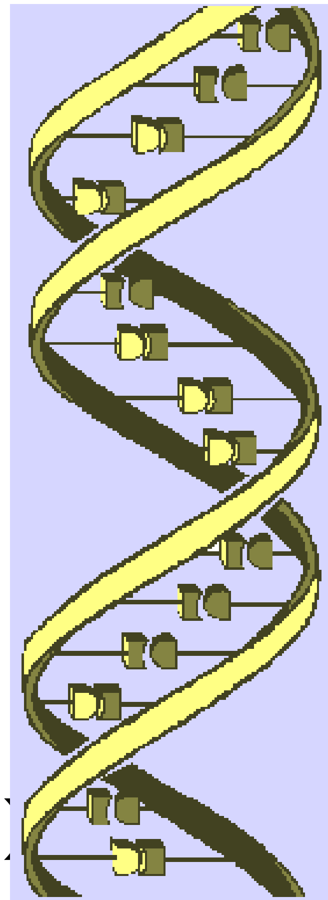


- Linking chains

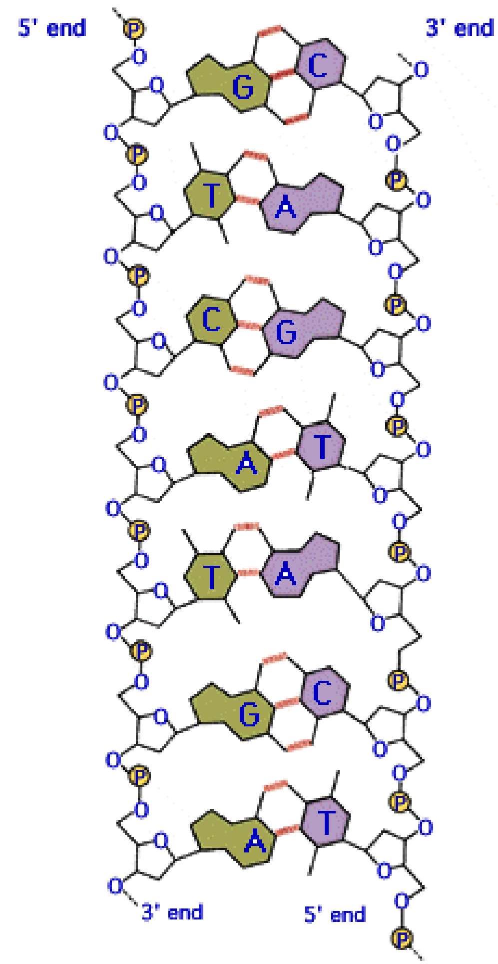


Double Helix I

- Double stranded, helix (Watson & Crick)
- Complementary
  - A-T
  - G-C
- Antiparallel
  - 3′ -> 5′ (downstream)
  - 5′ -> 3′ (upstream)

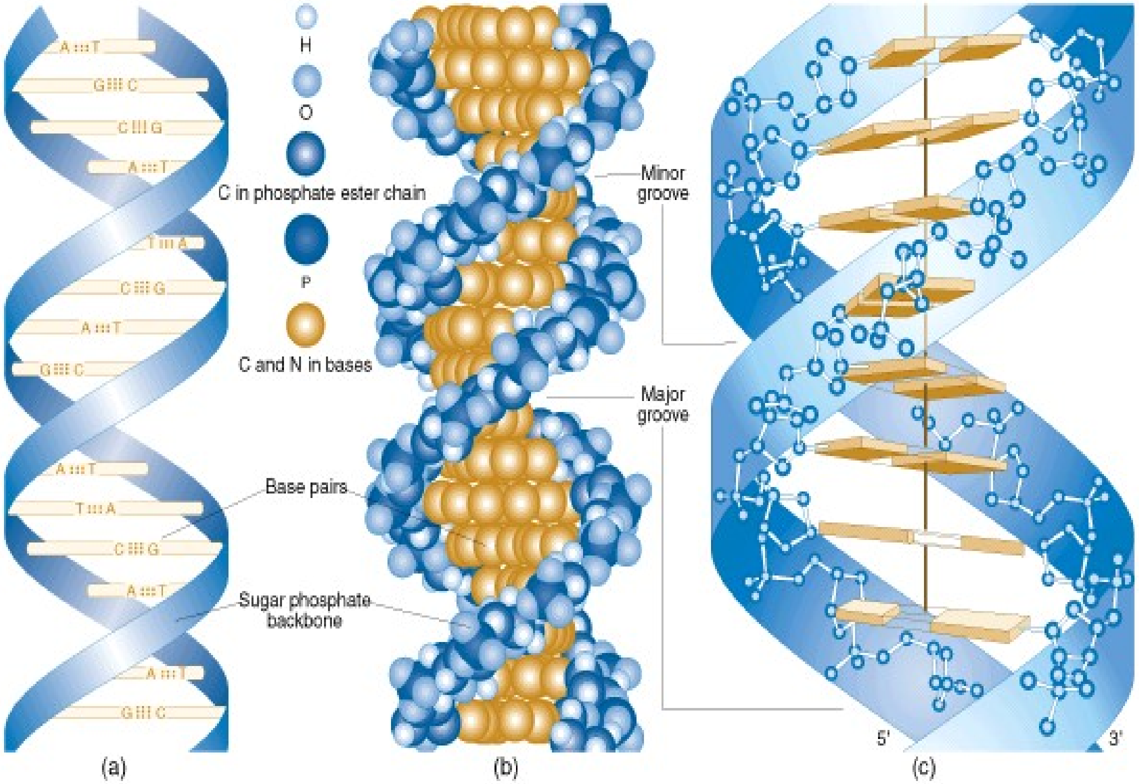


DNA showing the double helix



Double Helix II

DNA molecules usually consist of two strands arranged in the famous double helix



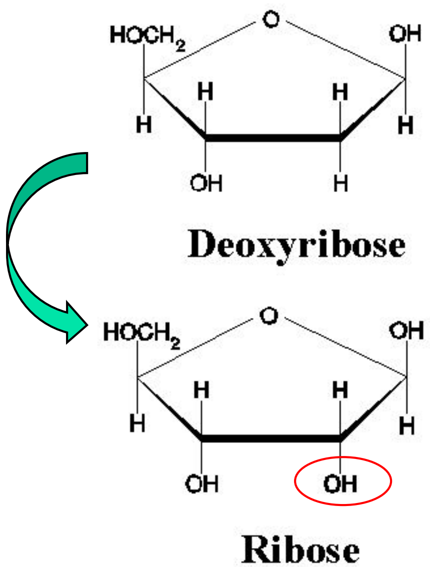
Transcription of DNA to RNA

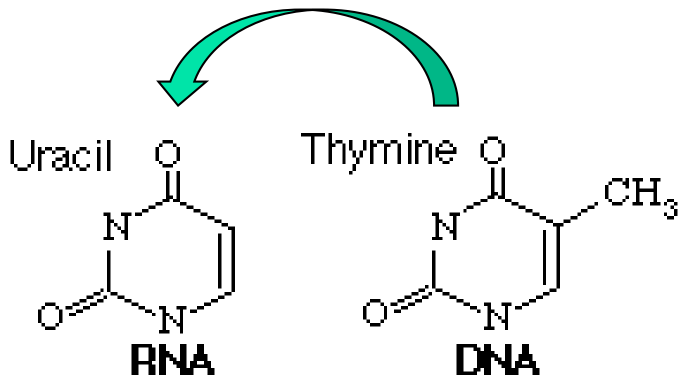
- Why transcription:
  - (For genome) to direct or effect changes in the cytoplasm of the cell
  - Need to generate new proteins to populate the cytosol (heteregenous intracellular soup of the cytoplasm)
- Note: DNA is in the nucleus, while proteins are needed in the cytoplasm, where many of the cell’s functions are performed.
- Coding region of the DNA is copied to a more transient molecule called RNA
  - Gene is a single segment of the coding region that is transcribed into RNA
  - Generation of RNA from DNA (in the nucleus) is done trough a process called transcription

RNA – Ribonucleic acid

- In RNA the base *Thymine* (T) is replaced by *Uracil* (U). The other difference to DNA is that the sugar (*Pentose*) will be *Ribose* instead of *Deoxiribose*. Ribose has an *additional hydroxyl group.*

Bases:
Cytosine - C
Guanine - G
Adenine - A
Uracil - U





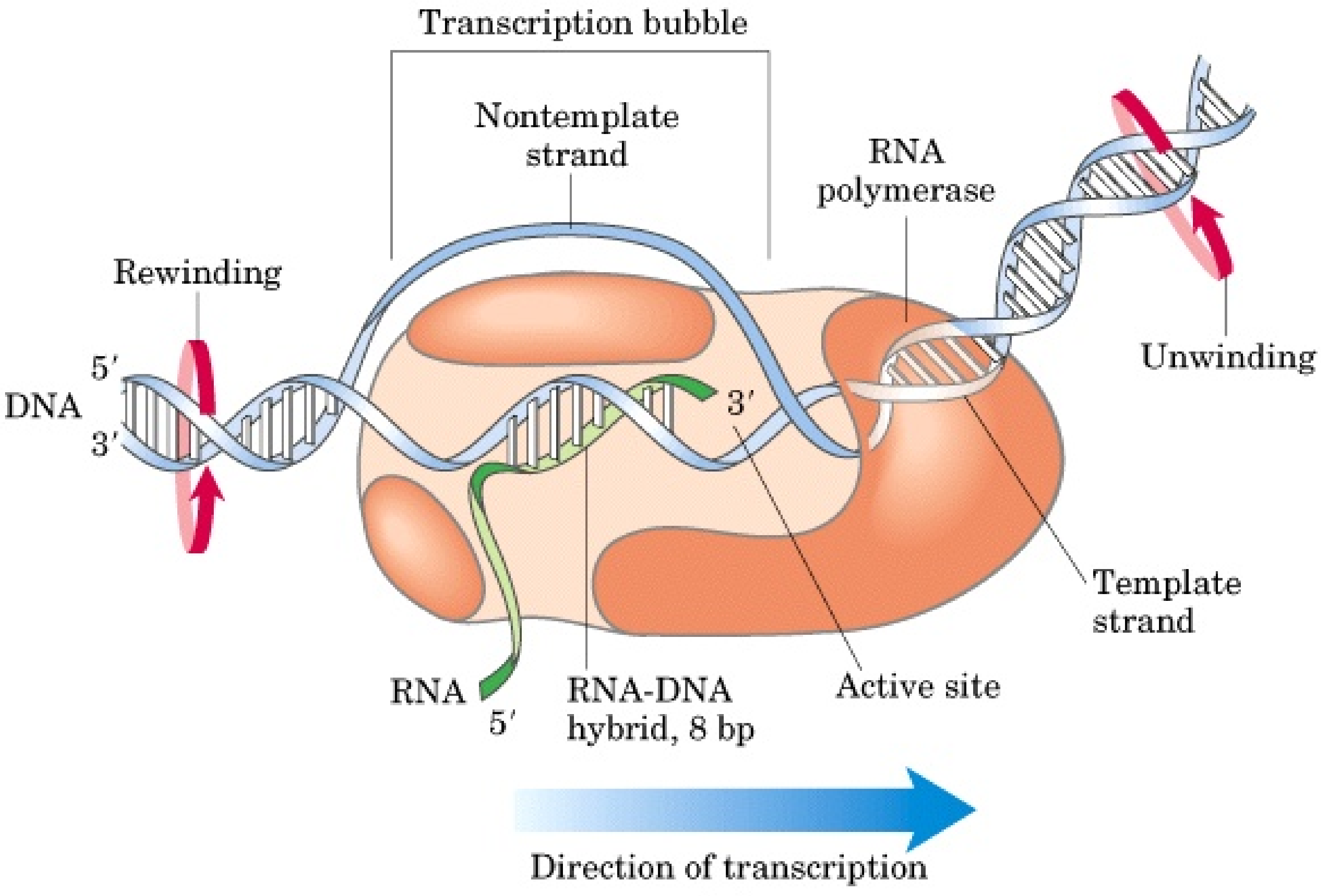
RNA transmits genetic information from DNA (via transcription) into proteins (by translation).
RNA is almost exclusively found in the single-stranded form.

RNA – Ribonucleic acid

RNA plays several roles in biology:

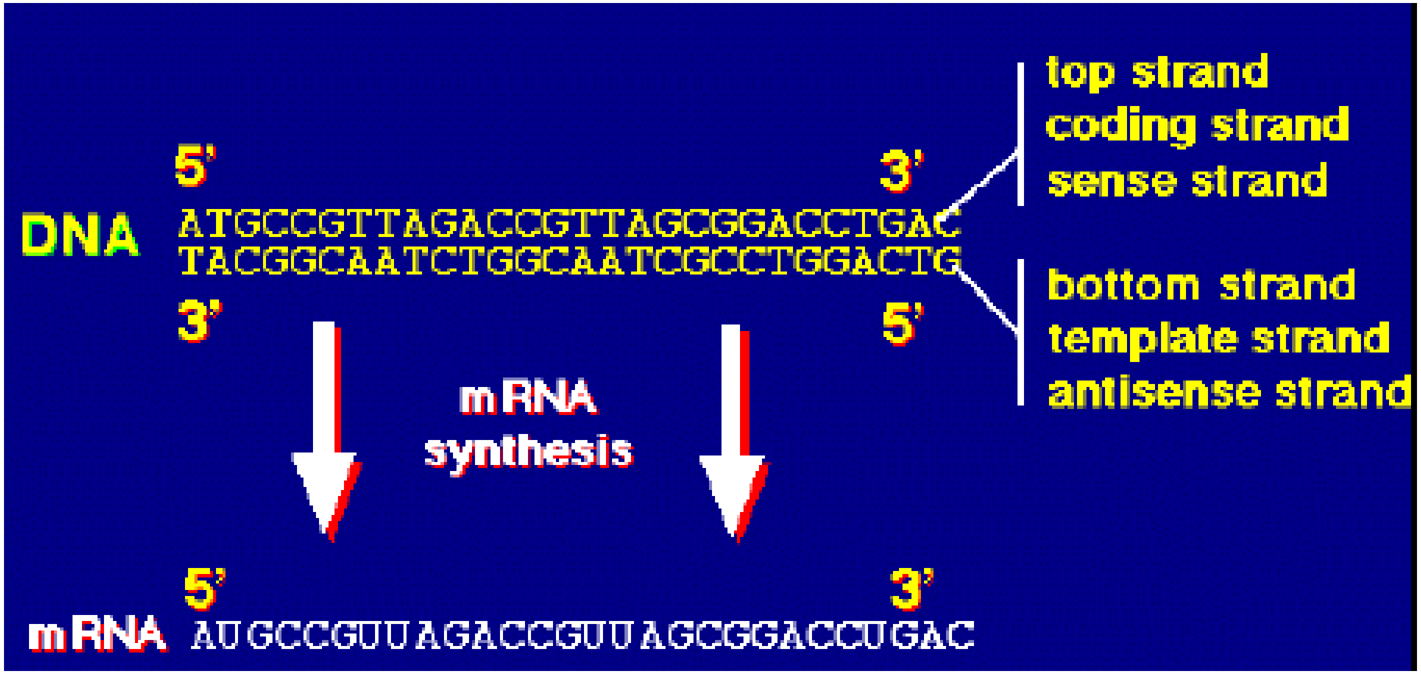
- **Messenger RNA** (**mRNA**) is transcribed directly from a gene's DNA and is used to encode proteins.
- RNA genes are genes that encode functional RNA molecules; in contrast to mRNA, these RNA do not code for proteins. The best-known examples of RNA genes are **transfer RNA** (**tRNA**) and **ribosomal RNA** (**rRNA**). Both forms participate in the process of translation, but many others exist.
- RNA forms the genetic material (genomes) of some kinds of viruses.
- **Double-stranded RNA** (**dsRNA**) is used as the genetic material of some RNA viruses and is involved in some cellular processes, such as RNA interference.

Transcription

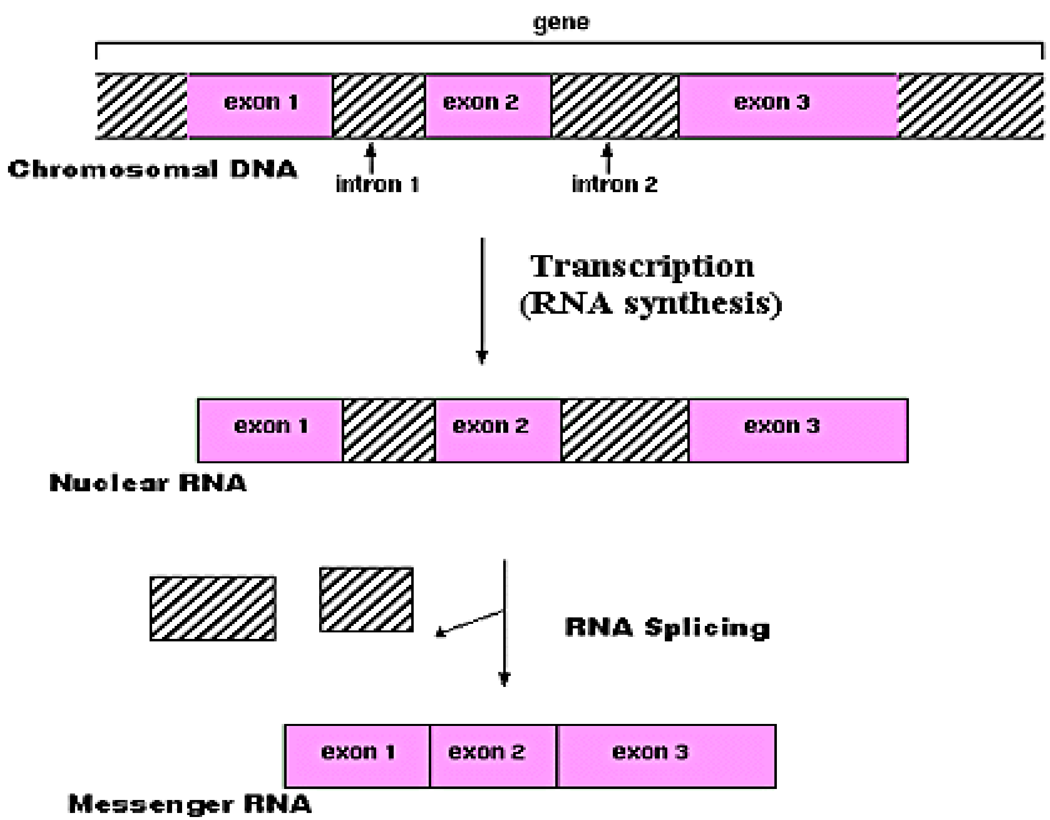


(a)

Transcription



mRNA splicing

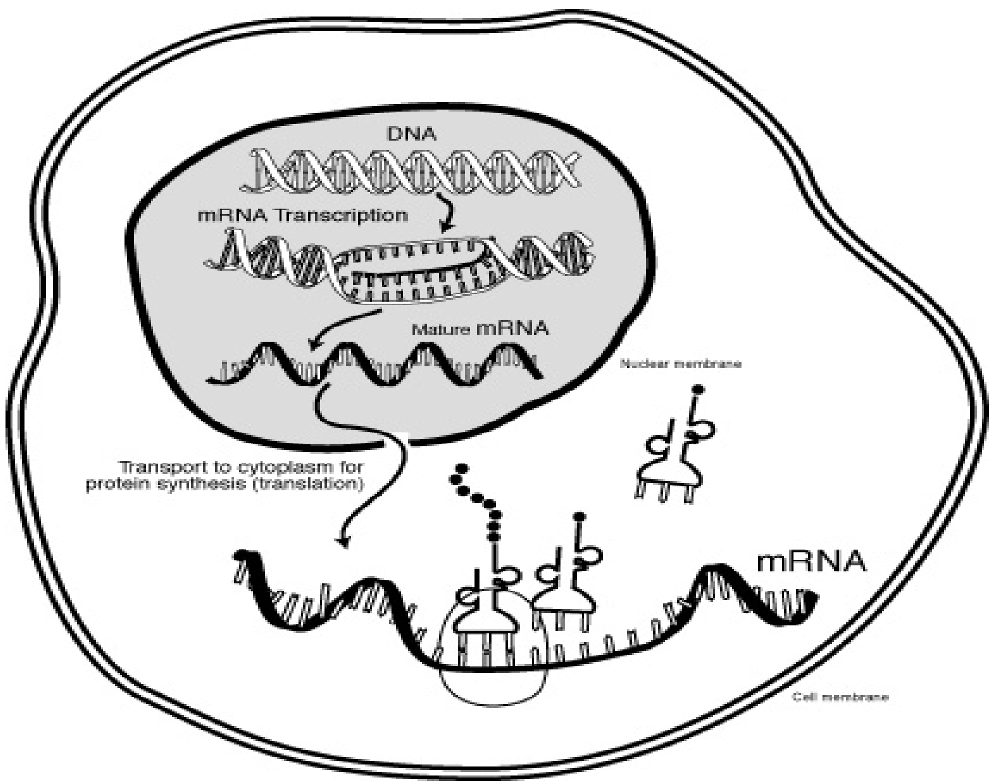


Translation

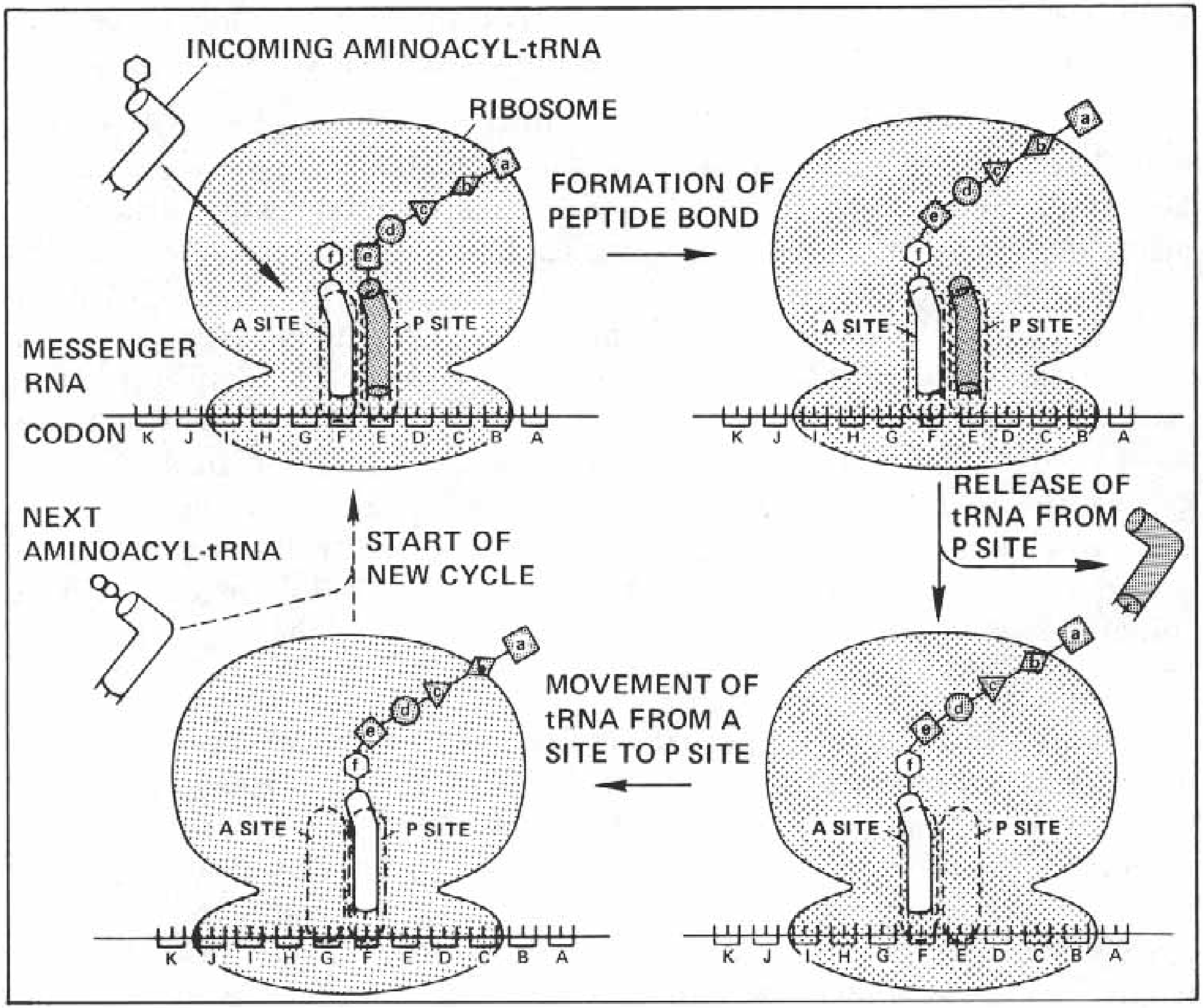
- *ribosomes* are the machines that synthesize proteins from mRNA
- the grouping of codons is called the *reading frame*
- translation begins with the *start codon*
- translation ends with the *stop codon*

Translation

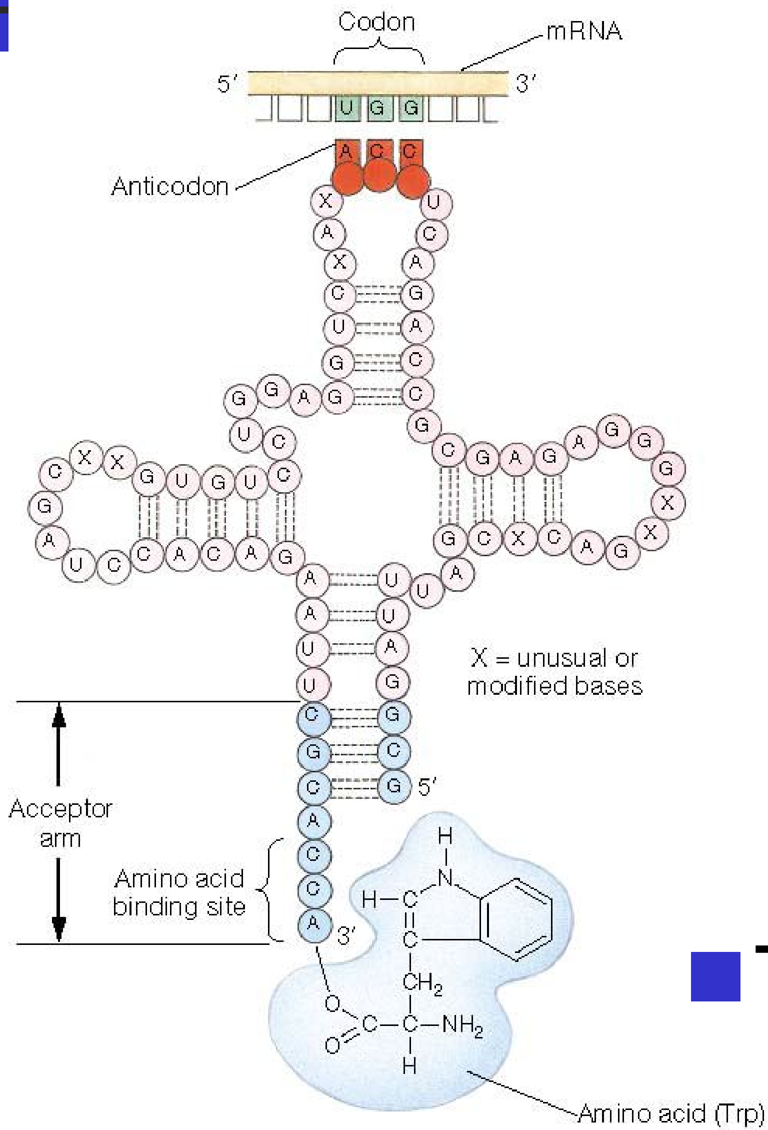
- Uses mRNA as template to make proteins
- Occurs in ribosomes
- One codon corresponds to one amino acid



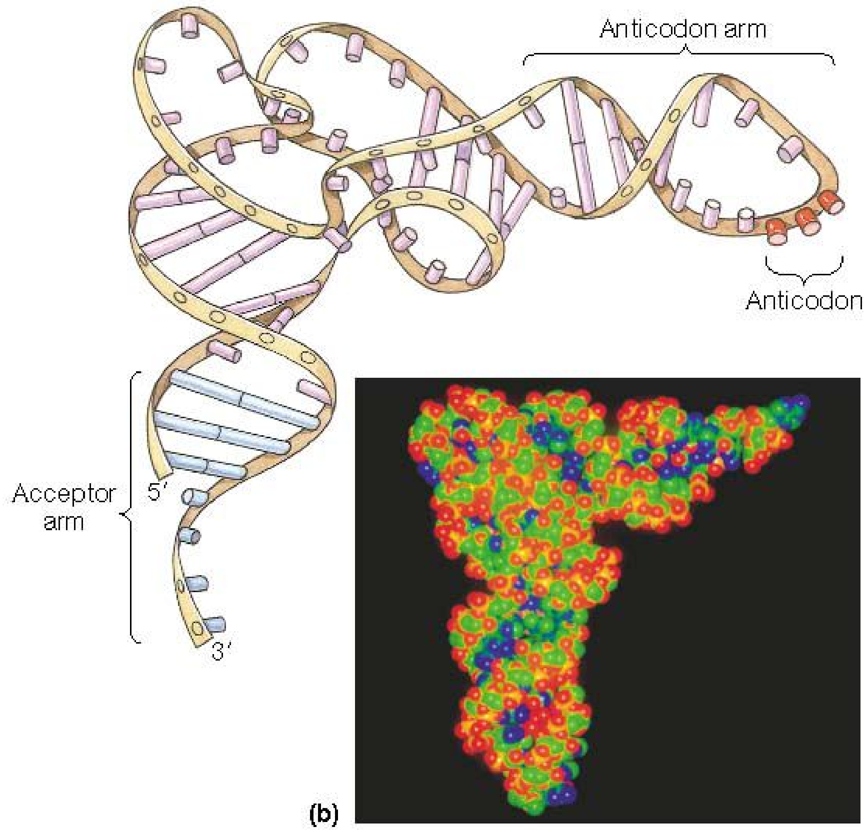
Translation



Translation



(a)

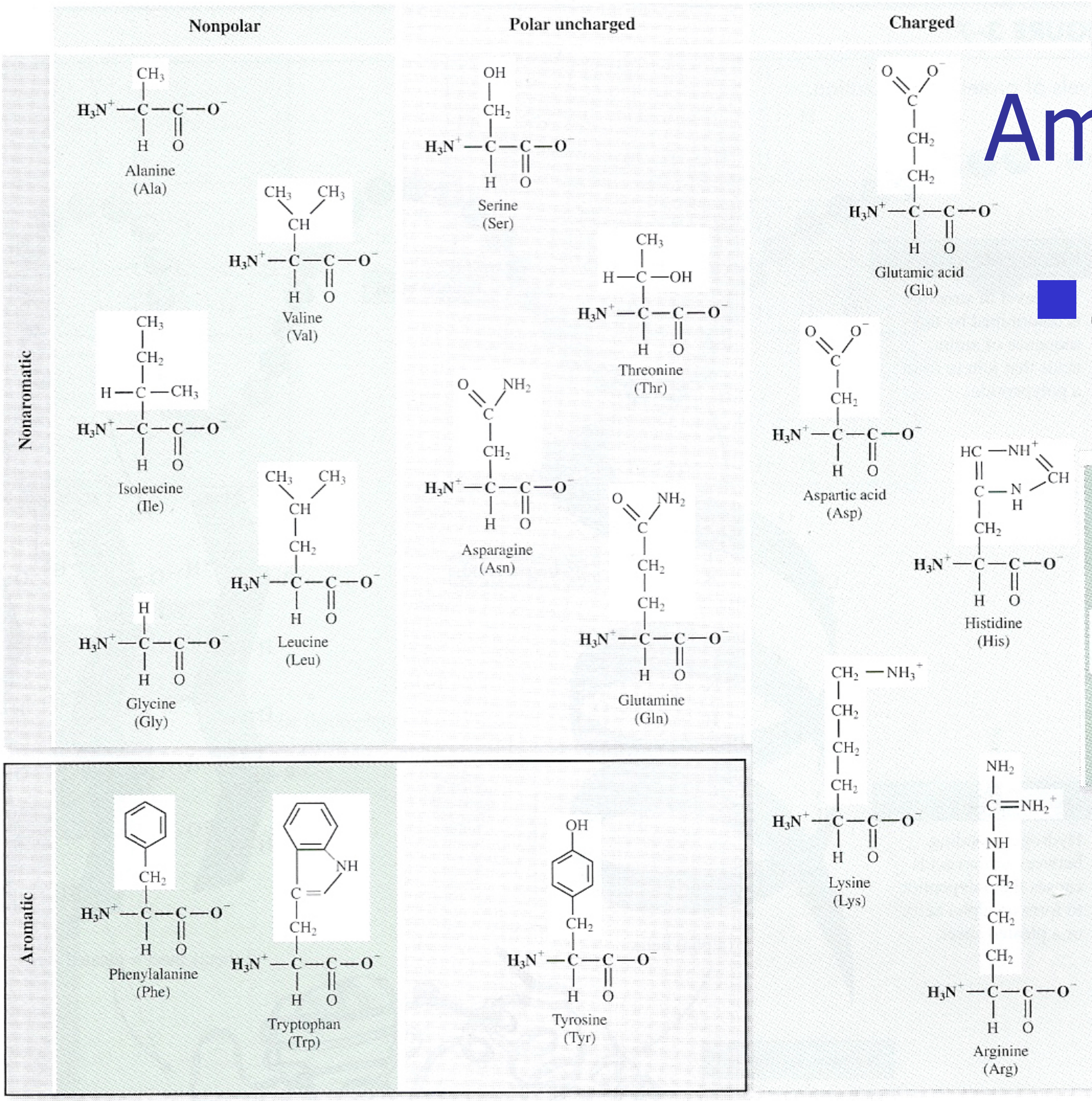


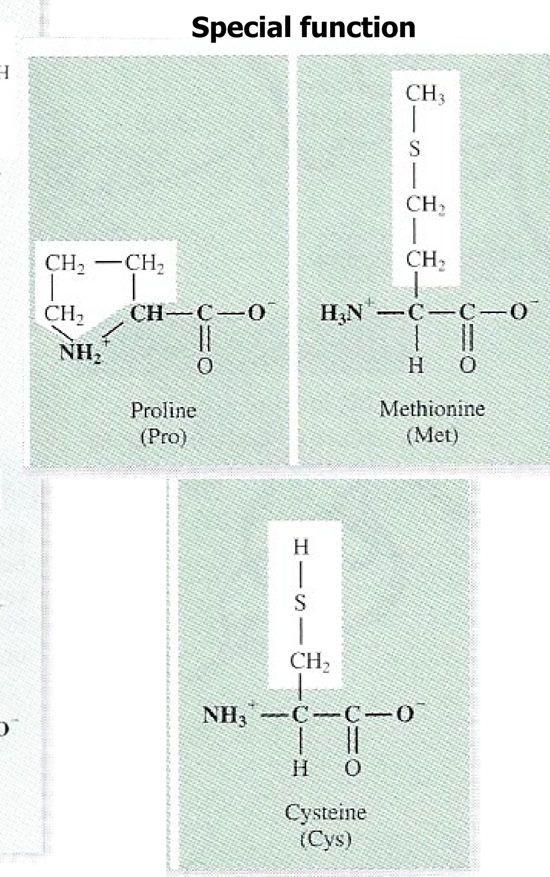
(b)

- Transfer RNA or tRNA

Amino Acids

- 20 essential





The Genetic Code

Second letter

First letter	U	C	A	G	Third letter
U	UUU, UUC: Phe UUA, UUG: Leu	UCU, UCC, UCA, UCG: Ser	UAU, UAC: Tyr UAA: Stop UAG: Stop	UGU, UGC: Cys UGA: Stop UGG: Trp	U C A G
C	CUU, CUC, CUA, CUG: Leu	CCU, CCC, CCA, CCG: Pro	CAU, CAC: His CAA, CAG: Gln	CGU, CGC, CGA, CGG: Arg	U C A G
A	AUU, AUC, AUA: Ile AUG: Met	ACU, ACC, ACA, ACG: Thr	AAU, AAC: Asn AAA, AAG: Lys	AGU, AGC: Ser AGA, AGG: Arg	U C A G
G	GUU, GUC, GUA, GUG: Val	GCU, GCC, GCA, GCG: Ala	GAU, GAC: Asp GAA, GAG: Glu	GGU, GGC, GGA, GGG: Gly	U C A G

Proteins

- proteins are molecules composed of one or more *polypeptides*
- a polypeptide is a polymer composed of *amino acids*
- cells build their proteins from 20 different amino acids
- a polypeptide can be thought of as a string
- composed from a 20-character alphabet

Proteins – Amino Acids

Name	1-letter code	Triplet
Glycine	G	GGT,GGC,GGA,GGG
Alanine	A	GCT,GCC,GCA,GCG
Valine	V	GTT,GTC,GTA,GTG
Leucine	L	TTG,TTA,CTT,CTC,CTA,CTG
Isoleucine	I	ATT,ATC,ATA
Histidine	H	CAT,CAC
Serine	S	TCT,TCC,TCA,TCG,AGT,AGC
Threonine	T	ACT,ACC,ACA,ACG
Cysteine	C	TGT,TGC
Methionine	M	ATG
Glutamic Acid	E	GAA,GAG
Aspartic Acid	D	GAT,GAC,AAT,AAC
Lysine	K	AAA,AAG
Arginine	R	CGT,CGC,CGA,CGG,AGA,AGG
Asparagine	N	AAT,AAC
Glutamine	Q	CAA,CAG
Phenylalanine	F	TTT,TTC
Tyrosine	Y	TAT,TAC
Tryptophan	W	TGG
Proline	P	CCT,CCC,CCA,CCG
Terminator (Stop)	*	TAA,TAG,TGA

Protein-Sequence (Alphabet: ACDEFGHIKLMNPQRSTVWY):

MENFQKVEKIGEGTYGVVY
KARNKLTGEVVALKKIRLDT
ETEGVPSTAIREISLLK...

- a typical human cell contains about 100 million proteins of about 10,000 types

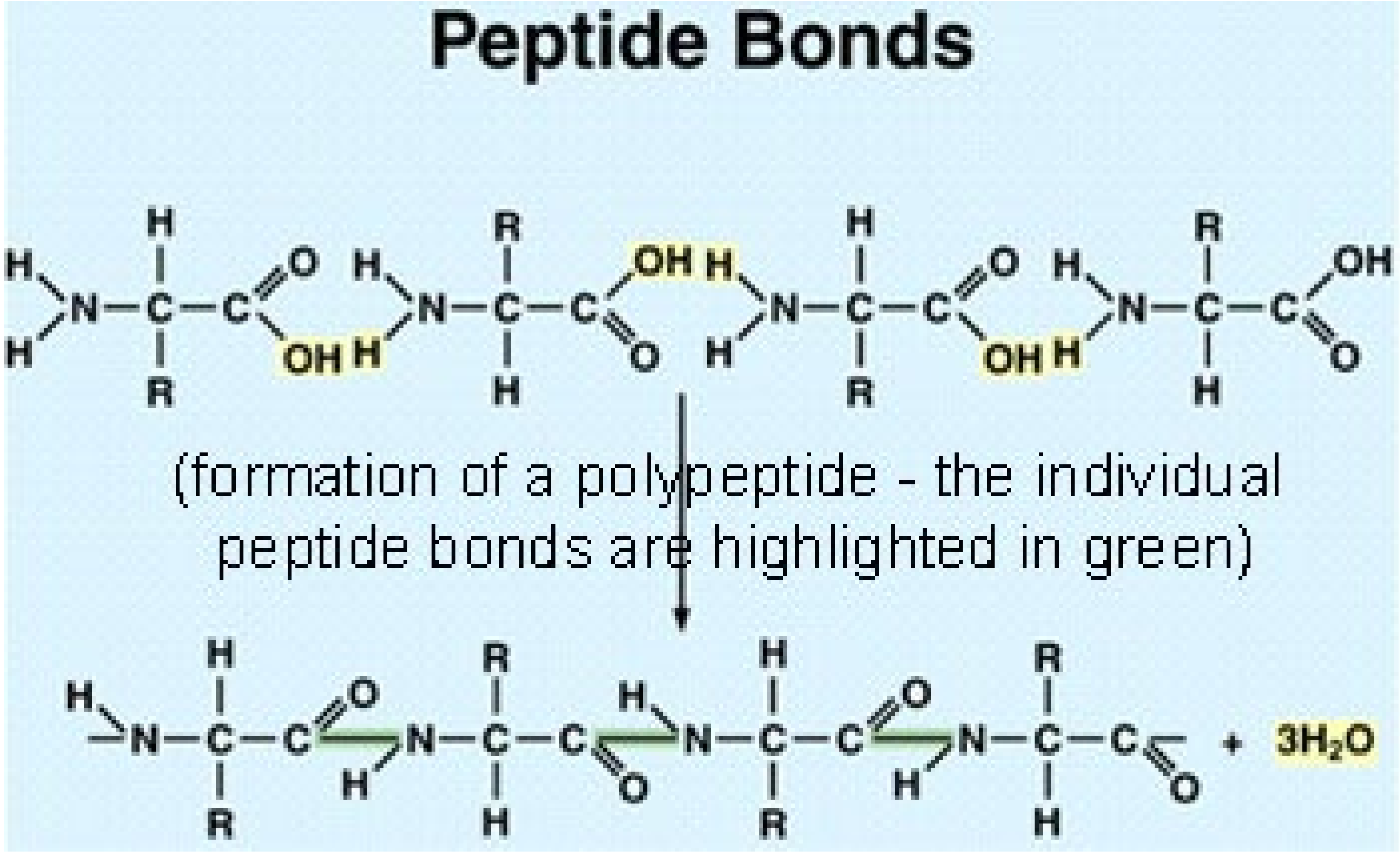
Proteins

- Huge proportion of cell (after water)
- Many functions:
  - Structure (e.g. collagen in bone)
  - Enzymes
  - Transmembrane receptors
  - Hormones
- Four levels of structure

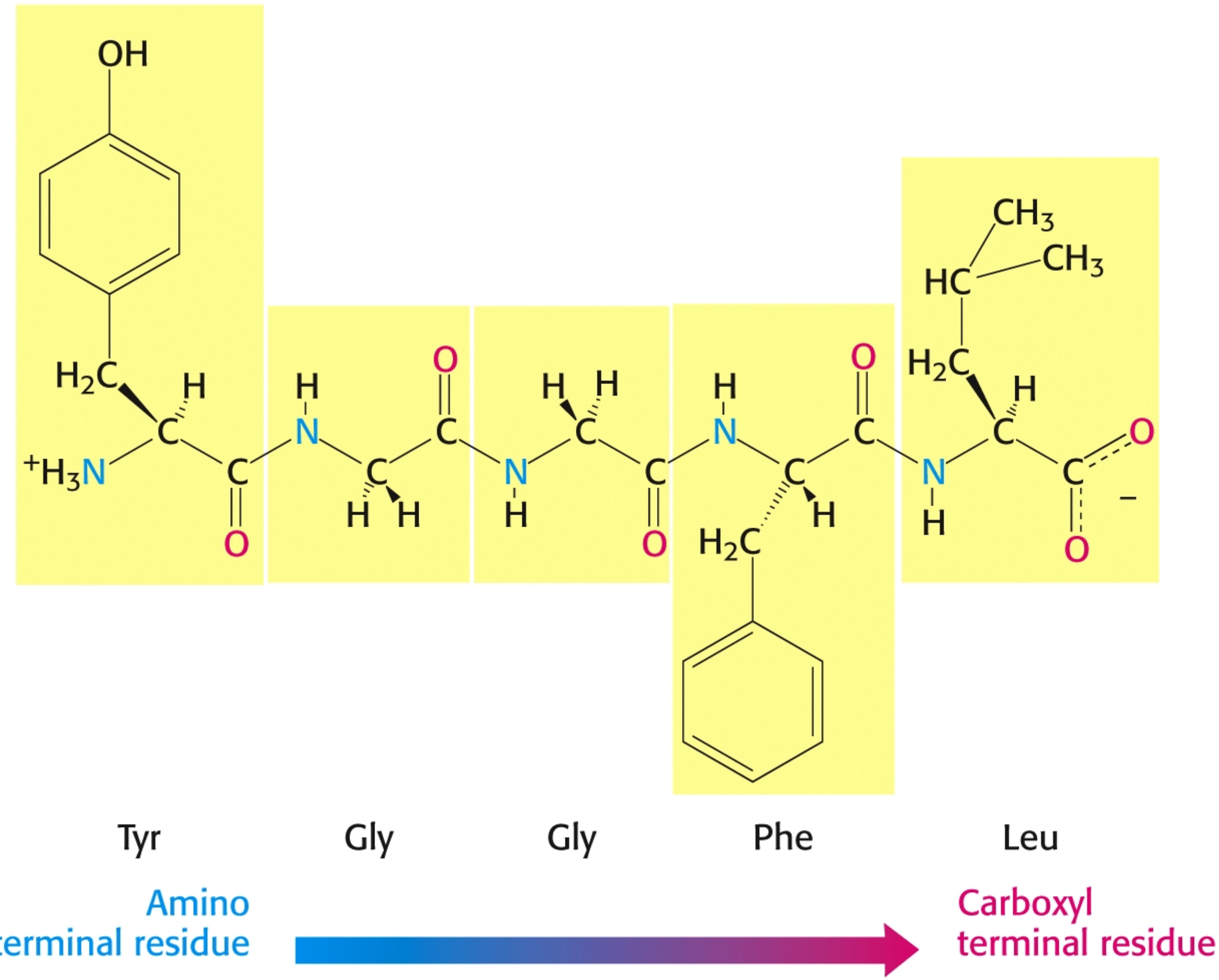
Protein Functions

- structural support
- storage of amino acids
- transport of other substances
- coordination of an organism’s activities
- response of cell to chemical stimuli
- movement
- protection against disease
- selective acceleration of chemical reactions

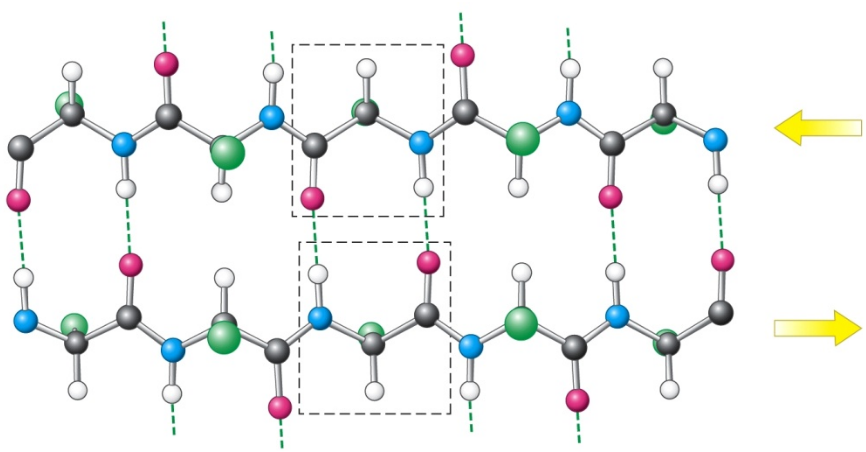
Peptide Bonds



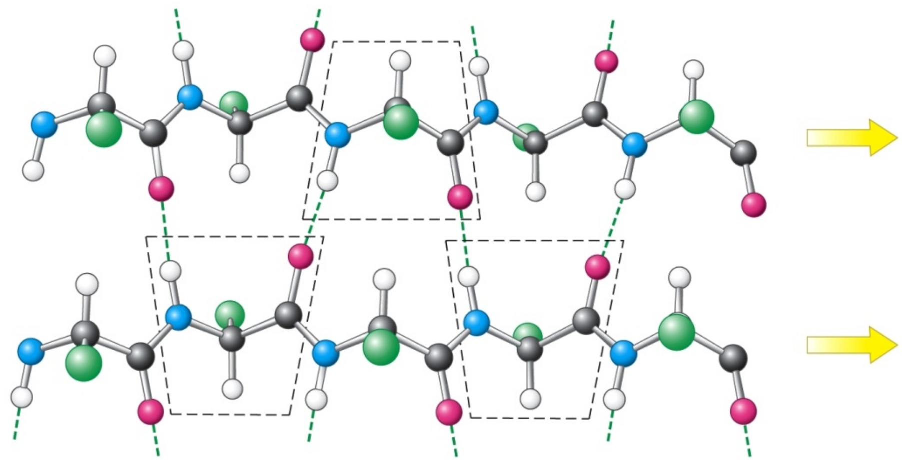
Direction of Protein Sequence



Secondary Structure: Beta sheet



anti-parallel

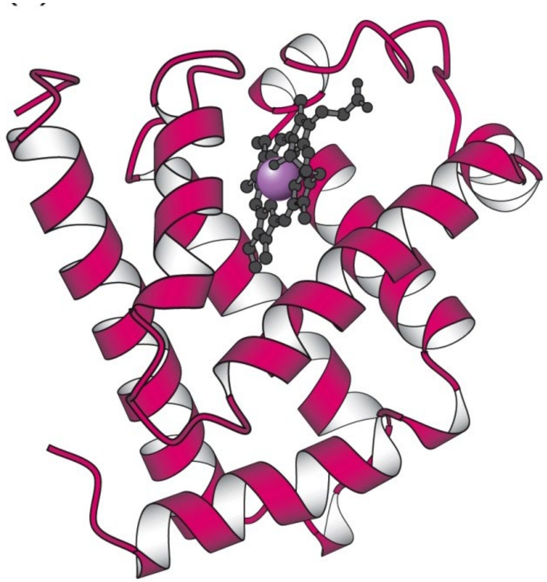


parallel

Phi = -135
Psi = 135

Tertiary Structure

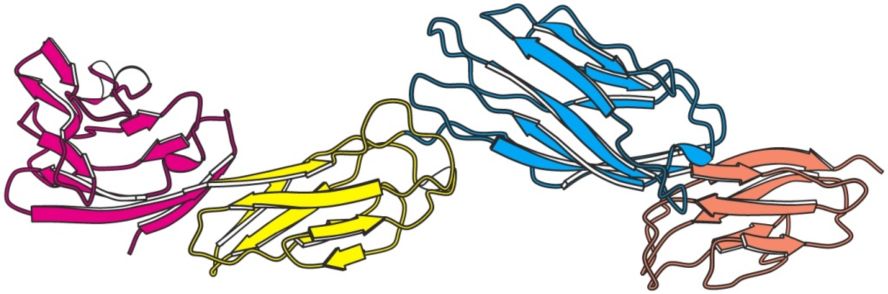
- 3-d structure of a polypeptide sequence
  - interactions between non-local atoms



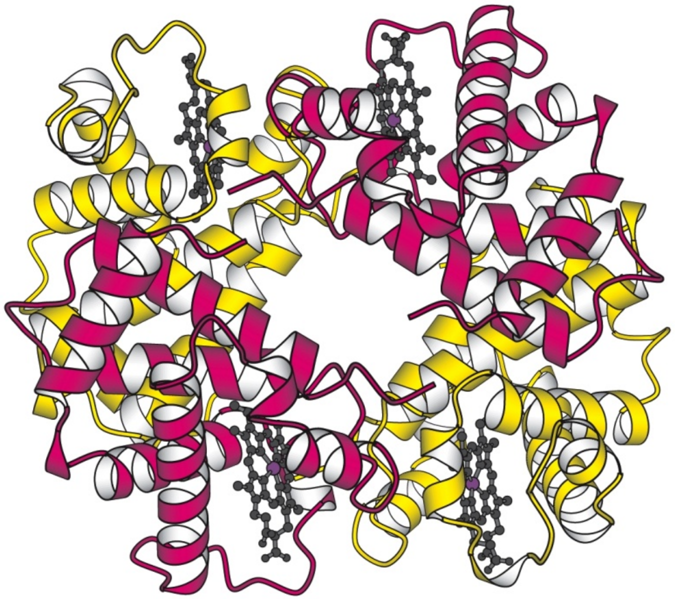
tertiary structure of
myoglobin

Quaternary Structure

- Arrangement of protein subunits



quaternary structure
of Cro



human hemoglobin
tetramer

Proteins

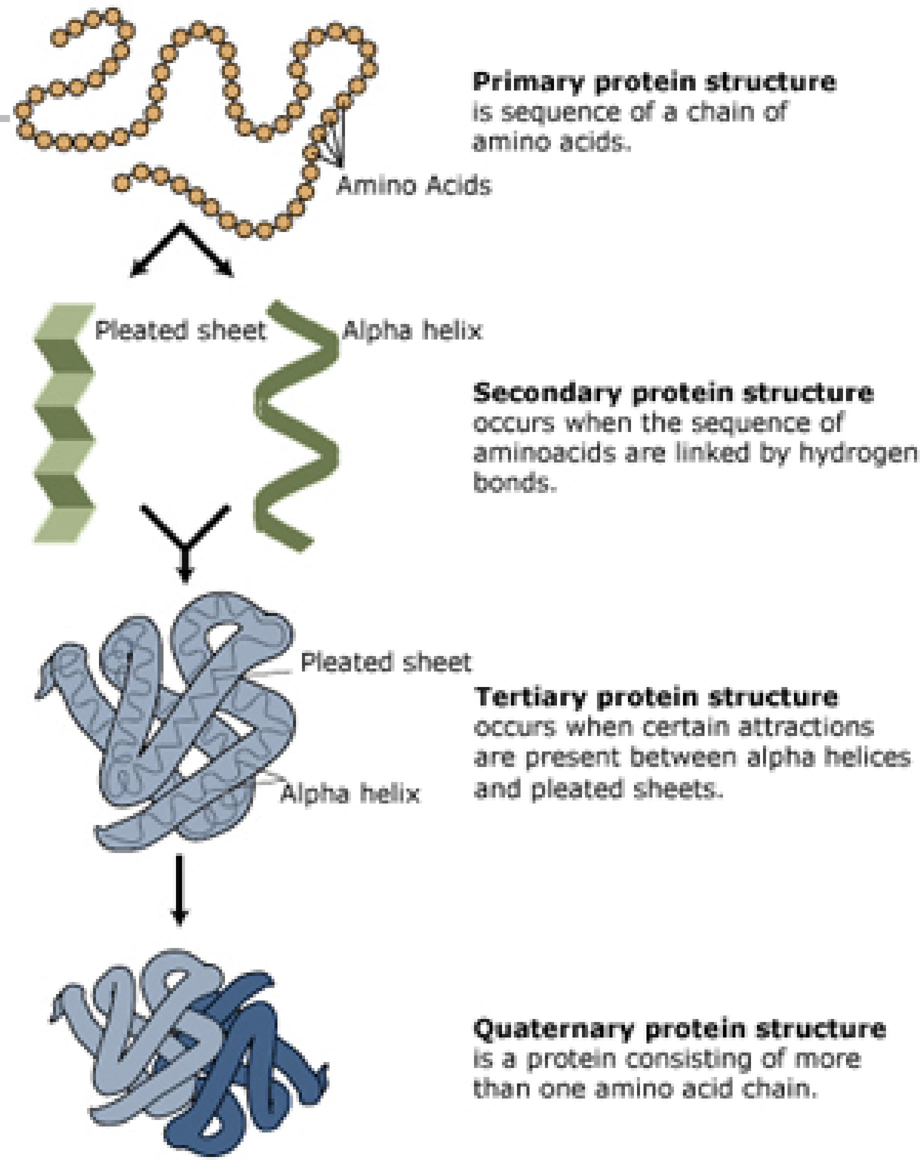


Image adapted from: National Human Genome Research Institute.

Translocation of proteins

- A newly formed protein needs to be translocated to the right place to perform its function (such as structural protein in the cytoskeleton, as a cell membrane receptor, as a hormone that is to be secreted by the cell, etc.)
- Signal peptide (header): part of the polypeptide that is one of the determinant of its location and handling

Transcriptional programs

- Initiation of the transcription process can be caused by external events or by a programmed event within the cell.
- External events
  - Piezoelectric forces generated in bones through walking can gradually stimulate osteoblastic and osteoclastic transcriptional activity to cause bone remodelling; Heat shock
  - Appearance or disappearance of new micro or macronutrients around the cell; binding of distantly secreted hormones
- Internally programmed sequences of transcriptional expression (eg. clock and per genes)
- Pathological internal derangements of the cell
  - Self-repair or damage detection programs can trigger apoptosis (self-destruction) under conditions such as irreparable DNA damage

Biological function of proteins

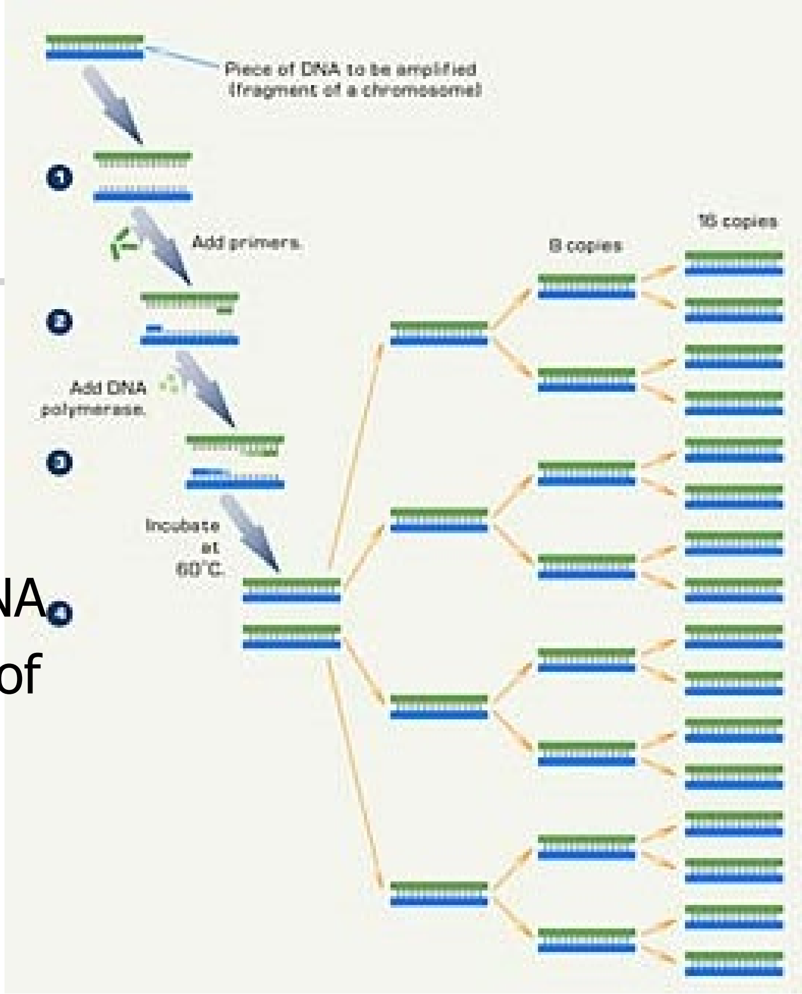
- **Enzyme catalysis**: DNA polymerases, lactate dehydrogenase, trypsin
- **Transport**: hemoglobin, membrane transporters, serum albumin
- **Storage**: ovalbumin, egg-white protein, ferritin
- **Motion**: myosin, actin, tubulin, flagellar proteins
- **Structural and mechanical support**: collagen, elastin, keratin, viral coat proteins
- **Defense**: antibodies, complement factors, blood clotting factors, protease inhibitors
- **Signal transduction**: receptors, ion channels, rhodopsin, G proteins, signalling cascade proteins
- **Control of growth, differentiation and metabolism**: repressor proteins, growth factors, cytokines, bone morphogenic proteins, peptide hormones, cell adhesion proteins
- **Toxins**: snake venoms, cholera toxin

Gene expression studies

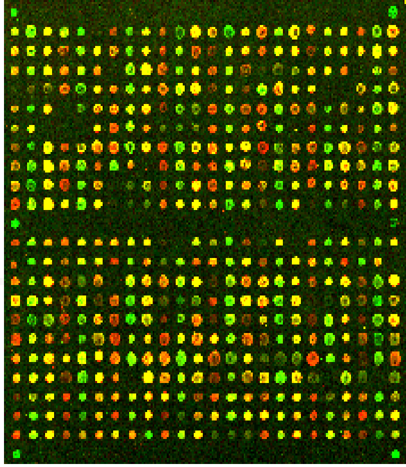
- Allow you to understand how a gene is regulated in a tissue or a cell type.
- Most useful way of studying gene expression is by measuring the levels of mRNA produced from a particular gene in a particular tissue.
- Application: to understand certain biological process it is useful to study the differences in gene expression which occur during such processes. E.g.
  - It is of interest to know which genes are induced or repressed, say in the aquatic environment to know if a particular contaminant can be degraded
- Some techniques for analyzing DNA or mRNA level of a single gene or to quantify gene expression
  - Polymerase Chain Reaction (PCR)
  - DNA microarrays
  - Proteomics (analysis of the protein synthesis that results from gene expression)

PCR

- Polymerase Chain Reaction
  - Used to amplify (increase concentration) of a single copy of DNA
  - Primers select for the DNA segment of interest
  - Thermal cycling with polymerase makes copies of it
- Allows identification and quantification of microorganisms that are not easily cultured



DNA microarrays



- Consist of thousands of DNA probes corresponding to different genes arranged as an array.
- Each probe (sometimes consisting of a short sequences of synthetic DNA) is complementary to a different mRNA (or cDNA)
- mRNA isolated from a tissue or cell type is converted to fluorescently labeled mRNA or cDNA and is used to hybridize the array.
- All expressed genes in the sample will bind to one probe of the array and generate a fluorescent signal.
- A DNA microarray can interrogate the level of transcription of several thousand of different genes from one sample in one experiment. (One DNA microarray experiment reveals the mRNA levels of 1000s of genes from one tissue or cell type at one time point)
- Particularly useful when studying the effect of environmental factors on gene expression.
- A fingernail size chip can interrogate 10,000 different transcripts. Chip has 30-40 different probes; half of them are designed to perfectly match 20 nucleotide stretches of the gene and the other half contains a mismatch as a control to test for specificity of the hybridization signal.

Detecting Mutagens

- Radiation and certain chemical compounds are "mutagens": they cause mutation.
- Cancer is caused by somatic mutations, and so mutagens are also carcinogens.
- Testing for mutagenicity is a key step is development of pharmaceutical drugs.
- Simple test using bacteria (Salmonella, a close relative of E. coli) developed by Bruce Ames: the "Ames test".

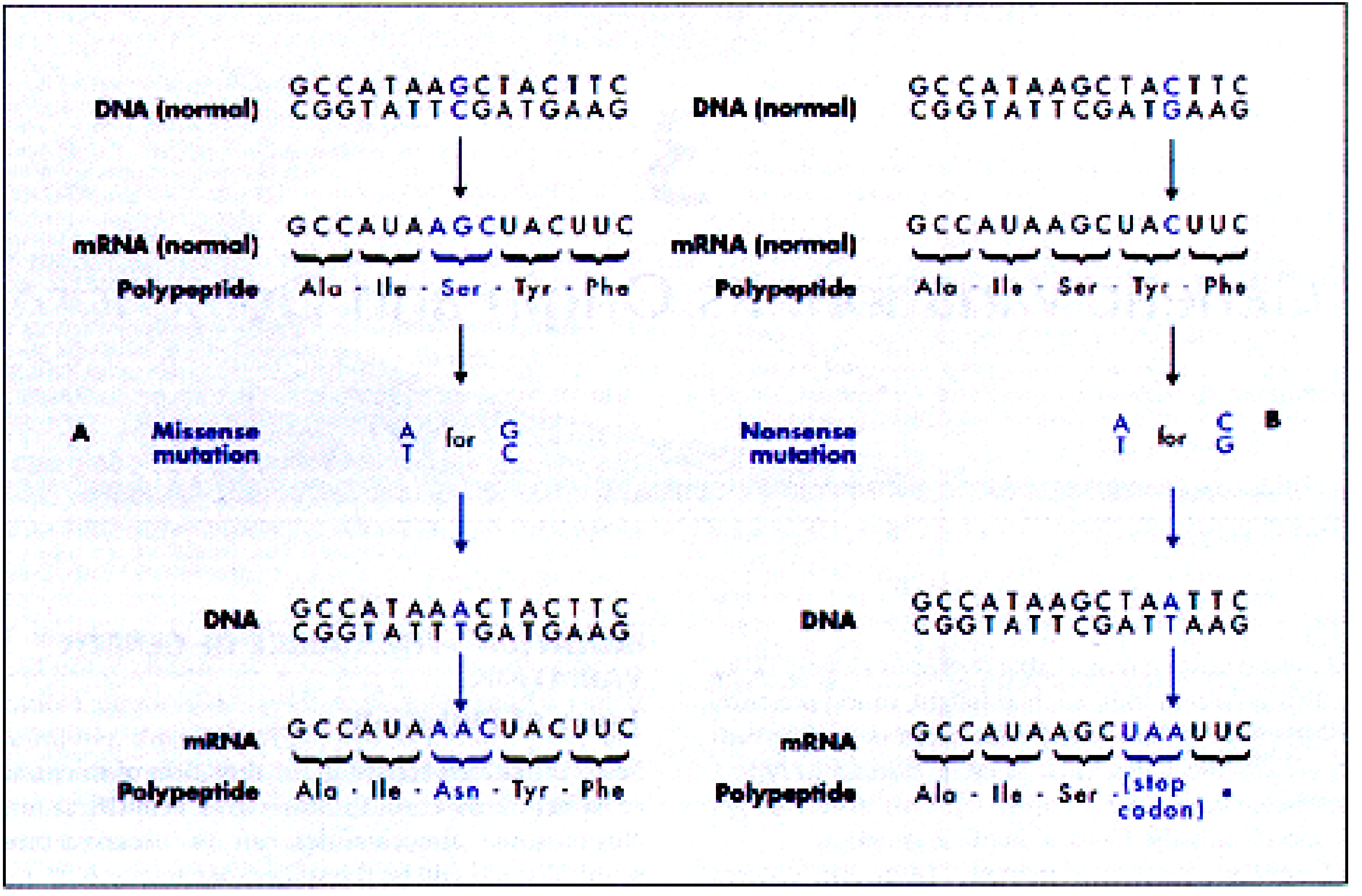


FIG. 3-1 Missense mutations **(A)** produce a single amino acid change, whereas nonsense mutations **(B)** produce a stop codon in the mRNA. Stop codons terminate translation of the polypeptide.

Frameshift Mutation

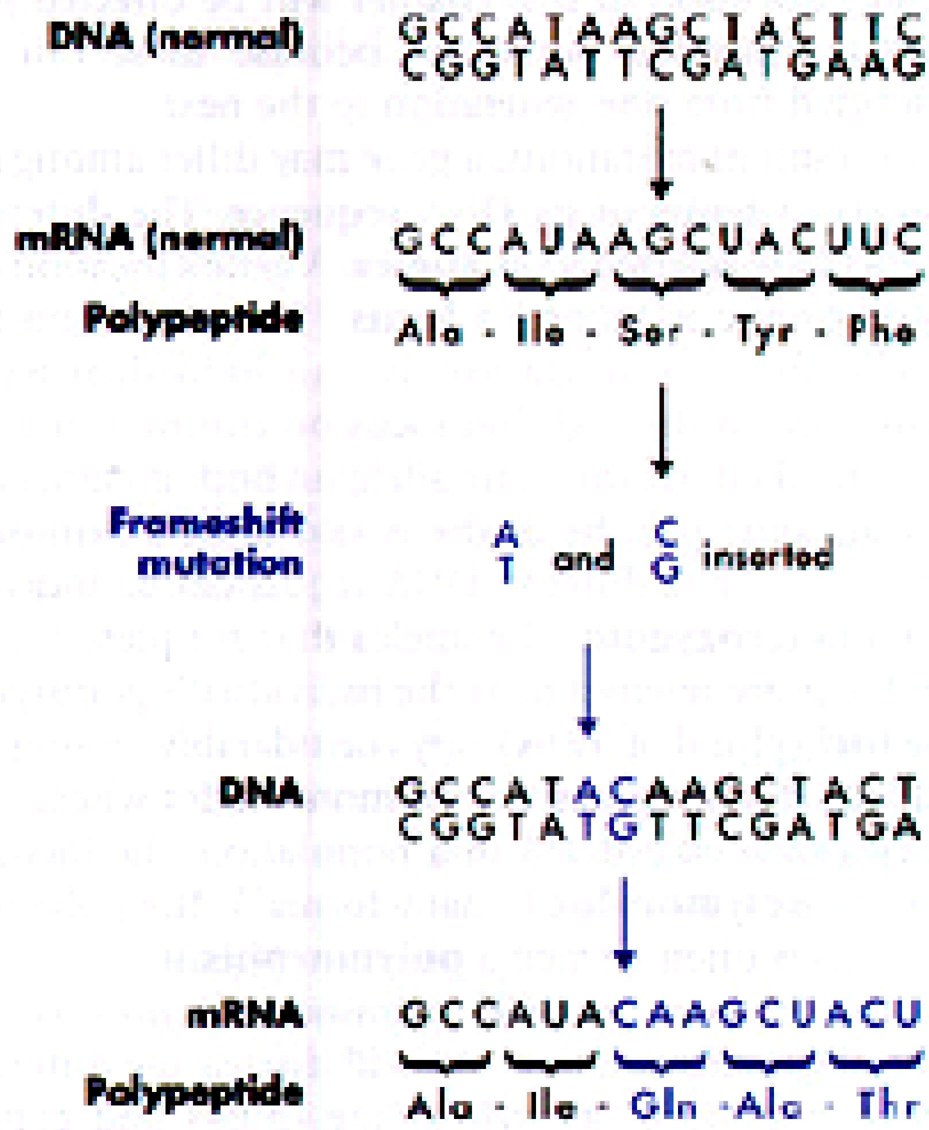
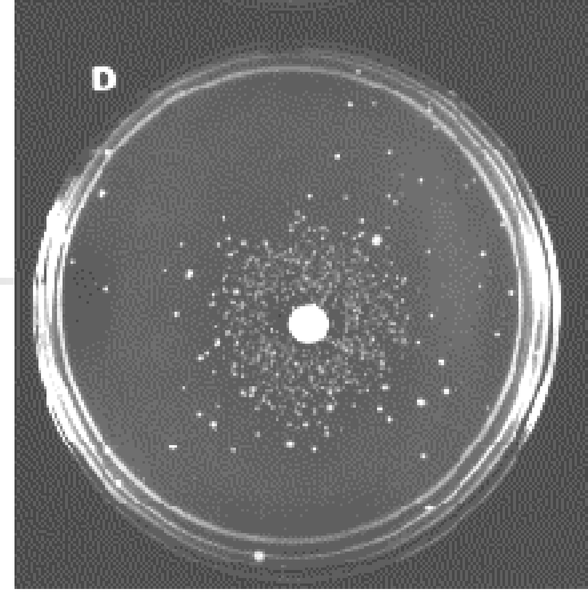


FIG. 3-2 Frameshift mutations result from the addition or deletion of a number of bases that is not a multiple of three. This alters all of the codons downstream from the site of insertion or deletion.

Ames Test



- An important test for environmental mutagens
- Start with Salmonella that are unable to make their own histidine. They will only grow if histidine is added to the growth medium.
- Add compound to be tested to growth medium, count number of colonies growing. These are revertants, which have been mutated back to wild type capable of making histidine.
- In many cases, mutagens need to be activated, converted to mutagenic state, by enzymes in the liver that are meant to detoxify dangerous compounds. Liver extracts are often added to the growth medium to accomplish this.
- Test isn't perfect: Salmonella are prokaryotes, and we have complex biochemistries that modify foreign compounds. But, it is a good initial screen.

- To next lecture