CH1131 Biomolecular Engineering

Syllabus

Chemical Basis, 2017

INTRODUCTION TO THE STUDY OF CELL & MOLECULAR BIOLOGY

LECTURE OUTLINE

Introduction

Basic Properties of Cells

I. Life – most basic property of cells; they are the smallest units to exhibit this property; plant or animal cells can be removed from organism & cultured in laboratory

A. Can grow and reproduce for long time in culture, unlike their parts, which soon deteriorate if isolated

B. Cultured cells are simpler to study than cells in body; cells grown in vitro (in culture, outside the body) have become essential tool of cell & molecular biologists

II. Cells are highly complex and organized

A. Each level of structure in cells has a great level of consistency from cell to cell – each cell type has consistent appearance in EM; organelles have particular shape & location in all individuals of species

B. Organelles have consistent macromolecular composition arranged in a predictable pattern

C. Cell structure is similar from organism to organism despite differences in higher anatomical features

1. Thus, information obtained from studying cells of one organism often has a direct application to other forms of life
2. Many of the most basic processes (protein synthesis, membrane structure, etc.) are remarkably similar in all living organisms

3. In evolutionary terms, many molecules in our cells must be very similar to those present in our primitive cellular ancestors that lived more than 3 billion years ago

III. Cells possess genetic program & the means to use it (a blueprint); encoded in collection of genes

A. Genes are the blueprint for constructing cellular structures & ultimately organisms – this vast amount of information is packaged into a set of chromosomes occupying the very small cell nucleus

1. Genes constitute the directions for running cell activities
2. Genes constitute the program for making more cells

B. Changes in genetic information from generation to generation lead to the variations that form the basis of biological evolution

IV. Cells are capable of producing more of themselves - mitosis and meiosis

A. Cells reproduce by division; process whereby “mother” cell contents are distributed to 2 “daughter” cells
B. Before division, genetic material is faithfully copied; each daughter cell gets complete & equal share of genetic information
C. Usually, daughter cells have roughly equal volume; however, during egg production, one cell gets nearly all of the cytoplasm & half of genetic material

V. Cells acquire & use energy (constant input) to develop & maintain complexity – photosynthesis, respiration

A. Most animal cells get energy prepackaged, often as glucose (released to blood by liver in humans)

B. Once in cell, glucose disassembled; most of its energy is stored as ATP & used to run cell activities

VI. Cells carry out a variety of chemical reactions - sum total of chemical reactions in cells (metabolism); to do this, cells require enzymes (molecules that greatly increase rate of chemical reactions)

VII. Cells engage in numerous mechanical activities based on dynamic, mechanical changes in cell. Most of which are initiated in the shape of "motor" proteins (require constant energy to keep working):

A. Material moved from place to place
B. Structures assembled and disassembled
C. Cells move from place to place

VIII. Cells able to respond to stimuli whether organisms are uni- or multicellular - have receptors that sense environment & initiate responses (move away from object in path or toward nutrient source)

A. Most cells covered with receptors that interact in specific ways with substances in environment

1. Receptors bind to hormones, growth factors, extracellular materials, surfaces of other cells
2. Allow ways for external agents to evoke specific responses in target cells

B. Cells may respond to specific stimuli by:

1. Altering their metabolic activities
2. Preparing for cell division
3. Moving from one place to another, or
4. Even committing suicide

IX. Cells are capable of self-regulation

A. Importance of regulatory mechanisms most evident when they break down

1. Failure of cell to correct error in DNA replication -> may lead to debilitating mutation
2. Breakdown in growth control -> may lead to cancer cell & maybe death of whole organism

B. Example: Hans Driesch, German embryologist (1891) - separate first 2 or 4 cells in sea urchin embryo -> each produces normal embryo; the cells regulated their activities to make whole embryos

C. Cell processes are a series of ordered steps – the information for these steps & product design reside in nucleic acids & construction workers for these processes/designs are primarily proteins

1. In cell, the workers act without benefit of conscious direction
2. Each step in process must occur spontaneously so that the next step is automatically triggered
3. Information needed to direct particular activity must be present within the system itself
4. Each type of cell activity requires unique set of highly complex molecular tools & machines

Two Fundamentally Different Classes of Cells: Prokaryotes and Eukaryotes

<u>I. 2 basic classes of cells</u> were distinguished by size & types of internal structures (organelles); exhibited a large fundamental evolutionary discontinuity (there are no known intermediates)

A. Prokaryotes (pro - before; karyon - nucleus) – all bacteria, cyanobacteria (blue-green algae); structurally simpler; not sure when prokaryotic cells first appeared on Earth

B. Eukaryotes (eu - true) - structurally more complex; protists, fungi, plants, animals

<u>II. Similarities between prokaryotes and eukaryotes</u> reflect the fact that eukaryotes almost certainly evolved from prokaryotic ancestors

A. Both types of cells share an identical genetic language

B. Both types of cells share a common set of metabolic pathways

C. Both types of cells share common structural features – similarly constructed plasma membrane that serves as selectively permeable barrier & cell walls (same function, different structure)

<u>III. Characteristics that distinguish prokaryotic & eukaryotic cells</u> - eukaryotic cells are much more complex internally (structurally and functionally) than prokaryotes

A. Eukaryotes have membrane-bound nucleus with complex nuclear envelope & other organelles

1. Prokaryotes have nucleoid (poorly demarcated cell region that lacks boundary membrane separating it from surrounding cytoplasm) & no membrane-bound organelles

B. Prokaryotes - relatively little DNA (0.25 - ~3 mm) coding for several hundred to several thousand proteins (1 mm of DNA = ~3 x 10^6 base pairs)

1. Although simplest eukaryotes (4.6 mm in yeast encoding ~6200 proteins) have slightly more DNA than prokaryotes; most have an order of magnitude more DNA (genetic information)

C. Eukaryotic chromosomes numerous; contain linear DNA tightly associated with protein; prokaryotes have a single, circular chromosome with DNA that is nearly "naked"

D. Eukaryotes contain an array of membranous & membrane-bound organelles that divide cytoplasm into compartments within which specialized activities take place; some examples follow:

1. Mitochondria (plants & animals) – make chemical energy available to fuel cell activities
2. Endoplasmic reticulum (plants & animals) – where many cell lipids & proteins are manufactured
3. Golgi complexes (plants & animals) – sorts, modifies, transports materials to specific cell locations
4. Variety of simple membrane-bound vesicles of varying dimensions (plants & animals)

E. Eukaryotes have many such membrane-bound structures; prokaryotes mostly devoid of them (except for infolded bacterial mesosomes & cyanobacteria photosynthetic membranes)

1. Intracytoplasmic communication smaller issue in prokaryotes due to size (simple diffusion works); in eukaryotes, interconnected channels/vesicles transport stuff around cell & outside of cell
2. Eukaryotes have cytoskeletal elements usually lacking in prokaryotes that give cell contractility, movement, support; primitive cytoskeletal filaments recently found in a few bacteria
3. Prokaryotic cytoskeleton much simpler structurally & functionally than that of eukaryotes
4. Prokaryote ribosomes smaller with fewer components than those of eukaryotes (but they essentially have the same function with similar mechanisms)
5. Both eukaryotes & prokaryotes may be surrounded by rigid, nonliving cell wall that protects, but their chemical composition is very different

F. No mitosis or meiosis in prokaryotes (binary fission instead); prokaryotes proliferate faster (double in 20 - 40 minutes; they exchange genetic information via conjugation)

1. In eukaryotes, duplicated chromosomes condense into compact structures; separated by mitotic spindle (elaborate; contains microtubules); allows daughter cells to get equal genetic material
2. In prokaryotes, no chromosome compaction & no spindle; DNA is duplicated & copies are separated by growth of intervening cell membrane
3. Prokaryotes do not reproduce sexually, but in conjugation, DNA is exchanged; the recipient almost never gets whole chromosome from donor; cell soon reverts to single chromosome

The Sizes of Cells and Their Components

I. Units of linear measure most often used to describe cell structures

A. Micrometers (µm; 10^{-6} m), nanometers (nm; 10^{-9} m)

B. Ångstroms (Å; 10^{-10} m) – often used by molecular biologists for atomic dimensions although no longer formally accepted in metric nomenclature); ~1 Å = roughly the diameter of H atom

II. Examples of dimensions of cells and cell components

A. Typical globular protein (myoglobin) - ~ 4.5 nm x 3.5 nm x 2.5 nm

B. Highly elongated proteins (collagen, myosin) - over 100 nm in length

C. DNA - ~2 nm in width

D. Large molecular complexes (ribosomes, microtubules, microfilaments) - 5 - 25 nm in diameter

1. These complexes are remarkably sophisticated nanomachines that can perform a diverse array of mechanical, chemical or electrical activities

E. Cells & organelles are more easily defined in micrometers

1. Nuclei - about 5 - 10 µm in diameter; mitochondria - about 2 µm in length

2. Bacteria - 1 to 5 µm in length; eukaryotic cells - 10 to 30 µm in length

Viruses

I. <u>Pathogens smaller and, presumably, simpler than smallest bacteria;</u> called viruses

A. Late 1800s – scientists thought infectious diseases caused by bacteria, but another agent soon found

1. Sap from sick tobacco plant found to infect other healthy plants while containing no bacteria

2. Sap still infective if forced through filter with pores smaller than the smallest known bacteria

3. Infectious agent could not be grown in culture unless living plant cells also present

B. Viruses responsible for many human diseases, some cancers - come in different shapes, sizes & constructions – AIDS, polio, influenza, cold sores, measles, a few types of cancers

1. Viruses occur in a wide variety of very different shapes, sizes & constructions, but all of them share certain common properties

<u>II. Common virus properties</u> - not considered living since need host to reproduce, metabolize, etc.

A. All are obligatory intracellular parasites (must reproduce in host cell [plant, animal, bacteria], depending on specific virus); they are macromolecular aggregates & inanimate particles

1. Alone, they are unable to reproduce, metabolize or carry on other life-associated activities
2. Thus, they are not considered to be organisms & not considered to be alive
3. Once it has attached & passed through membrane, its genetic material can alter host cell activities

B. Outside of living cell, it exists as particle, essentially a macromolecular package

1. Has small amount of genetic material (single or double stranded DNA or RNA)
2. They can have as few as 3 or 4 genes up to several hundred genes
3. The fewer the genes, the more a virus relies on enzymes/other proteins encoded by host cell genes

C. Genetic material is surrounded by protein capsule (capsid) usually made up of a specific number of subunits; efficient (need only a few genes to make capsid)

1. Numerous advantages to construction by subunits, one of most apparent being an economy of genetic information
2. If viral coat is made of many copies of one protein or a few proteins, virus needs only one or a few genes to code for its protein container
3. Capsid subunits often organized into polyhedron (a structure having planar faces [ex.: 20-sided icosahedron]) like adenovirus, which causes mammalian respiratory infections

D. Many animal viruses have capsid surrounded by lipid-containing outer envelope derived from modified host cell membrane as virus buds from host cell surface (ex.: HIV, HBV)

E. Viruses have surface proteins that bind to particular host cell surface component (specificity)

1. HIV - glycoprotein of 120,000 daltons MW (gp120) interacts with specific protein (CD4) on surface of certain white blood cells, facilitating virus entry into host cell
2. Viral & host protein interaction determines virus specificity, the hosts it can enter & infect

F. Most viruses have relatively narrow host range (certain cells of certain host like human cold & influenza viruses, which are only able to infect human respiratory epithelium cells)

1. But some can have wide host range, infecting cells from a variety of organs or species - rabies infects variety of mammalian host species (bats, dogs, humans)
2. Host cell specificity change can have dramatic effect; 1997 – 18 Hong Kong poultry workers came down with influenza & 6 died

VI. Viruses have virtues – viral gene activities mimic those of host genes & have been used in variety of ways

A. Research tool used to study host DNA replication & gene expression in their much more complex hosts

B. Used to introduce foreign genes into human cells as treatment for human diseases by gene therapy

THE CHEMICAL BASIS OF LIFE

LECTURE OUTLINE

Covalent Bonds

I. Molecular atoms are joined together by covalent bonds in which electron pairs are shared between atoms

A. Formation of a covalent bond is governed by the basic principle that atoms are most stable with a full outer electron shell

1. Number of bonds an atom forms determined by how many electrons are needed to fill outer shell
2. Outer & only shell of hydrogen & helium atoms is filled when it contains 2 electrons; outer shells of other atoms are filled when they contain 8 electrons

B. Bond formation is accompanied by energy release

1. Later reabsorption of energy by bond breaks it; C—C, C—H or C—O covalent bonds require 80 - 100 kcal/mole to break
2. This energy is quite large so these bonds are stable under most conditions

C. Atoms can be joined by bonds in which >1 pair of electrons are shared: if 2 pairs are shared -> double bond (O2); if 3 pairs shared -> triple bond (N2); no quadruple bonds are known

D. Type of bond can determine molecular shape - atoms joined by single bond can rotate relative to one another; atoms of double & triple bonds cannot

Noncovalent Bonds

I. A variety of noncovalent bonds govern interactions between molecules or different parts of a large biological molecule; such bonds are typically weaker linkages, while covalent bonds are stronger

A. Depend on attractive forces between atoms having an opposite charge

1. Involve interaction between positively & negatively charged regions within same molecule or on 2 adjacent molecules; usually weaker than covalent bonds, which are strong
2. Individual noncovalent bonds are often weak (~1 - 5 kcal/mole); they readily break & reform

3. When many of them act in concert (DNA, protein, etc.), attractive forces add up & provide structure with considerable stability

B. Noncovalent bonds mediate the dynamic interactions among molecules within the cell

II. Types of noncovalent bonds: Ionic bonds (or salt bridges)

A. Ionic bonds - result from transfer of electron(s) from 1 atom to another leading to atoms with positive & negative charges that attract each other; can hold molecules together (DNA-protein)

1. In crystal, strong; in water, ions surrounded by water, prevents attraction between them

2. Water surrounds individual ions & inhibits oppositely charged ions from approaching each other closely enough to form ionic bonds

B. Bonds between free ions not important in cells because cells are mostly water; weak ionic bonds between oppositely charged groups of large molecule are much more important

1. Ionic bonds in cell are generally weak (~3 kcal/mole) due to presence of water

2. Deep in protein core where water is excluded, they can be influential

III. Types of noncovalent bonds: <u>hydrogen (H) bonds - hydrophilic (water-loving)</u>; enhance solubility in & interactions with water

A. If H is bonded to electronegative atom (O or N), the shared electron pair is displaced toward electronegative atom so H is partially positive; H shared between two electronegative atoms

1. Bare positively charged nucleus of H can approach unshared pair of outer electrons of second electronegative atom —> an attractive (weak electrostatic) interaction (an H bond)

2. Occur between most polar molecules; important in determining structure & properties of water, also form between polar groups present in large biological molecules (like DNA)

B. Strong collectively because their strength is additive; weak individually (2 - 5 kcal/mole in aqueous solutions); a result of polar covalent bonding; makes DNA double helix very stable

IV. Types of noncovalent bonds - <u>hydrophobic (water-fearing) interactions</u>

A. Polar molecules like amino acids & sugars are said to be hydrophilic (water-loving); nonpolar molecules (fat molecules or steroids; water-fearing) are essentially insoluble in water

B. Most believe that they are not true bonds since not usually thought of as attraction between hydrophobic molecules

V. Types of noncovalent bonds - <u>van der Waals interactions (forces)</u>

A. Hydrophobic groups can form weak bonds with one another based on electrostatic interactions; due to slight perturbations of electron distributions

B. If 2 such molecules are very close together & appropriately oriented, 2 electrically neutral molecules will experience weak attractive force bonding them together (**van der Waals forces**)

1. Formation of temporary charge separation in one molecule can induce similar separation in adjacent molecule & lead to additional attractive forces among nonpolar molecules
2. Single van der Waals very weak (0.1 - 0.3 kcal/mole) & very sensitive to distance separating 2 atoms
3. Molecules must be close together & interacting portions have complementary shapes that allow close approach; many atoms of both interactants can approach each other closely
4. Important biologically as with interactions between antibodies and viral antigens

The Nature of Biological Molecules: Background

I. Organic molecules - contained in cell dry weight; once thought to only be found in living organisms; their name distinguishes them from inorganic molecules found in inanimate world

A. Chemists learned to synthesize them so some of the mystique was dispelled

B. Called them **biochemicals** (compounds made by living organisms)

II. Organic chemistry centers around carbon – both its size & electronic structure allow carbon to generate many molecules (several 100,000 known)

A. Binds to up to 4 other atoms, since it has only 4 outer-shell electrons (8 needed to fill shell)

B. Form carbon-containing backbones with long chains, which may be linear, branched or cyclic

C. Carbons can be connected by single, double (with O and N) or triple bonds (with N)

D. Compounds very stable since strength of covalent bond inversely proportional to atomic weight of elements involved; example: silicon (just below carbon in periodic table)

1. Silicon (4 outer-shell electrons) is too large for its +-charged nucleus to attract neighboring atom valence (outer-shell) electrons enough to hold such large molecules together

III. Hydrocarbons - contain only hydrogen & carbon atoms (simplest group of organic molecules)

A. As more carbons added, skeletons increase in length & structure becomes more complex

1. As get bigger, can have same formula but different structures (structural isomers) & properties

B. Fully reduced or saturated when each carbon bound to maximum number of hydrogen atoms

C. Unsaturated compounds have double or triple bonds; lack maximum number of H atoms

D. Rotation of carbons around single bonds, but not around double & triple bonds

IV. Functional groups - particular atom groupings that often behave as unit; responsible for physical properties, chemical reactivity & solubility in aqueous solutions; replace H's in hydrocarbons

A. Hydrocarbons do not occur often in living cells although they form the bulk of fossil fuels formed from the remains of ancient plants & animals

1. Many organic molecules important in biology contain chains of carbons like those in hydrocarbons but some of the hydrogens are replaced by various functional groups

B. Some major functional groups

1. Hydroxyl group - —OH

2. Carboxyl group - —COOH; acquires charge —COO-; carboxylic acids react with alcohols to form **ester bond**

3. Sulfhydryl group - —SH; react to form disulfide bonds in polypeptides

4. Amino group - —NH2; acquires charge —NH3+; react with carboxylic acids & form **amide bonds**

C. How do functional groups affect or change the properties of biochemicals?

1. Usually contain one or more electronegative atoms (N, P, O and/or S) & thus make organic molecules more polar, more water soluble & more reactive

2. Many are capable of ionization & may become positively or negatively charged

D. Example of functional group importance (ethane -> ethanol -> acetic acid -> ethyl mercaptan)

The Nature of Biological Molecules: Functional Classification of Biological Molecules

I. **Macromolecules** - form structure & carry out activities of cells; usually huge & highly organized molecules; contain from dozens to millions of carbon atoms

A. Because of their size & the intricate shapes they can assume, some can perform complex tasks with great precision & efficiency

B. Endow organisms with properties of life & set them apart chemically from inanimate world

C. Divided into 4 major categories: proteins, nucleic acids, polysaccharides, lipids - first 3 are **polymers**; made of large number of low MW building blocks (**monomers**)

D. Basic structure & function of each type of macromolecule are similar in all organisms

1. If look at special sequences of monomers making up these various macromolecules, the diversity among organisms becomes apparent

II. Macromolecule building blocks – most macromolecules in cell have short lifetime compared with cell (except DNA); steadily broken down & replaced by new macromolecules

A. Most cells contain supply (pool) of low MW precursors to build macromolecules

B. Monomers - building blocks of macromolecules (sugars/polysaccharides, amino acids/proteins, nucleotides/nucleic acids, fatty acids & glycerol/lipids)

1. Monomers joined together & form polymers by process like coupling railroad cars onto train

III. Metabolic intermediates (metabolites) – molecules in cell have complex chemical structures & must be synthesized in step-by-step sequence beginning with specific starting materials

A. In cell, each series of chemical reactions is called a **metabolic pathway**

1. Pathway starts with a compound & converts it to other ones sequentially until an end product that can be used in other reactions (like an amino acid building block of protein) is made

B. Compounds formed along pathways leading to end products might have no function per se except as a stop on the way to the end product & are called **metabolic intermediates**

IV. Molecules of miscellaneous function – vast bulk of cell dry weight is made up of macromolecules & their direct precursors

A. Vitamins – function primarily as adjuncts to proteins

B. Certain steroid or amino acid hormones

C. Molecules involved in energy storage (ATP, creatine phosphate)

D. Regulatory molecules - cyclic AMP

E. Metabolic waste products - urea

The Types of Biological Molecules: Carbohydrates

I. Carbohydrates comprise a group of substances, including simple sugars (**monosaccharides**) & larger molecules made from them

A. Serve primarily as chemical energy storehouse & durable building material for biological construction

B. Most have general formula (CH2O)n

II. The structure of simple sugars – each sugar molecule consists of carbon atom backbone linked together in linear array by single bonds

III. Linking sugars together to make larger molecules – bond joining sugars together called **glycosidic** linkage or bond (–C—O—C–); forms by reaction between C1 of one sugar & OH of another

A. Sugars can be joined by a variety of different glycosidic linkages
B. 2 monosaccharides covalently bond together to form **disaccharide**; serve primarily as readily available energy stores
1. Sucrose (table sugar) - major component of plant sap; carries chemical energy from one part of plant to another
2. Lactose (milk sugar) - fuel for early growth & development of newborns
a. Enzyme lactase that hydrolyzes it is found in membranes of cells lining intestines
b. If lose this enzyme after childhood, eating dairy products causes digestive discomfort

C. Oligosaccharides - small chains of sugars (*oligo* - few), usually attached to lipids & proteins converting them to glycolipids & glycoproteins, respectively
1. Particularly important on plasma membrane from which they project
2. They may be composed of many different combinations of sugar units & can thus play an informational role
3. They can distinguish one cell type from another & help mediate specific interactions of a cell with its surroundings
D. Polysaccharides – many, many sugars hooked together; very large molecules

IV. Polysaccharide types **-** sugars, starches, cellulose, chitin, peptidoglycan, glycosaminoglycans
A. Glycogen – branched glucose polymer mostly joined by α(1—>4) bonds

B. Starch - glucose polymer; mixture of 2 different polymers (amylose & amylopectin); plants bank their surplus chemical energy in form of starch (potatoes & cereals are primarily starch)
C. Cellulose – tough, durable structural material (cotton, linen); major plant cell wall component

Lipids
I. Composed principally of C, H & O - not macromolecules, but aggregate to form large complexes
A. Includes diverse, heterogeneous group of nonpolar biological molecules (fats, oils, phospholipids, sterols)
B. Lumped together due to solubility in organic substances (benzene, chloroform) & their insolubility in H2O, which explains many of their varied biological functions
C. Serve as fuel molecules, very rich in chemical energy (contain more [>2X] energy than carbohydrates); structural components

D. Person of average size contains ~0.5 kg of carbohydrate primarily in form of glycogen (~2000 kcal of total energy) & ~16 kg of fat (144,000 kcal of energy)
E. Since they lack polar groups & are extremely water-insoluble, stored as dry lipid droplets in cells (extremely concentrated storage fuel)

II. Triglyceride (neutral lipid, fats, triacylglycerol) - serves as lipid storage form for fuel (stored in adipocytes)
A. Formed by 3 condensation reactions, which form ester linkages (—C—O—C—) between glycerol (a polar molecule) & 3 fatty acids
B. Fatty acid chains can vary in length & degree of saturation (see below)
C. 3 fatty acids of triglyceride need not be identical but may be; if they contain more than one fatty acid species called mixed fats

III. Fatty acids - long, unbranched hydrocarbon chains with single carboxyl group at one end
A. Both hydrophobic (long C chain) & hydrophilic (carboxyl; "-" charge at physiological pH) in character (**amphipathic**); they have unusual & biologically important properties
B. Soap - in past, soap was made by heating animal fat in strong alkali (NaOH, KOH) to break fatty acid - glycerol bonds; most are now made synthetically
1. Hydrophobic end of fatty acids embed in grease; hydrophilic end interacts with water
2. Greasy materials form complexes that can be dispersed by water (micelles)

IV. Sterols and steroids – complex & characteristic 4 ringed hydrocarbon structures (4 joined rings differ in numbers & positions of double bonds & functional groups)
A. Most common & one of most important - cholesterol; a component of animal cell membrane, but not in internal membranes or in plants
1. Precursor for synthesis of many steroid hormones (testosterone, progesterone, estrogen)
2. While largely absent from plant cells (vegetable oils considered to be cholesterol-free), plants may contain lots of related compounds
B. Adrenocortical hormones
C. Sex hormones - estrogen, progesterone, testosterone, etc.
D. Vitamin D3 and the bile acids - involved in lipid digestion in the intestine

V. Phospholipids (phosphoglyceride, diacylglycerol) - glycerol + 2 fatty acids + phosphate group on third hydroxyl (often an amino group as well); highly charged at physiological pH; amphipathic
A. Major cellular function - presence in membranes (properties of which depend on phospholipids)

Proteins: General Information

I. Composed of H, C, O, N & usually S or P; very large macromolecules; polymers of amino acids

A. Traits & functions - more varied role than other organism molecules (enzymes, structural or both); execute almost all cell activities; typical cell may have ~10,000 different ones; high specificity

1. Enzymes catalyze and vastly accelerate rate of metabolic reactions
2. Cytoskeletal elements serve as structural cables, provide mechanical support in & out of cells
3. Hormones, growth factors, gene activators – wide variety of regulatory functions
4. Membrane receptors & transporters - determine what cell reacts to, what can leave, enter cell
5. Contractile elements & molecular motors – machinery for biological movements
6. Antibodies and toxins
7. Form blood clots
8. Absorb or refract light
9. Transport substances from one part of body to another

B. The wide variety of protein functions comes from the virtually unlimited shapes they can assume

1. They can exhibit a great variety of structures and thus a great variety of activities
2. Each protein has unique, highly ordered structure enabling it to carry out particular function
3. Their shapes allow them to interact selectively with other molecules (high degree of specificity)

C. Protein polymer sequences give them their unique properties

1. Many protein capabilities can be understood by examining the chemical properties of its constituent amino acids
2. 2 aspects of amino acid structure: that which is common to all & that which is unique to each

E. During protein synthesis, each amino is joined to 2 other amino acids forming a long, continuous, unbranched polymer (**polypeptide chain**); have N-terminus & C-terminus

II. Properties of R groups - determine inter- & intramolecular interactions that determine molecular structure & protein activities, respectively; give amino acids their variability

A. Polypeptide backbone is made of that part of each amino acid that is common to all of them

B. Side chain (R group) bonded to α-carbon is highly variable; this gives proteins their diverse structures & activities

C. Side chains of enzyme active sites can facilitate (catalyze) many different organic reactions

III. Four amino acid & R group categories – classified by R group character; not all of the amino acids are found in all proteins; nor are various amino acids distributed in an equivalent manner

A. <u>Polar charged</u> - contain R groups that act as stronger organic acids, bases; can form ionic bonds

1. Almost always fully charged (lysine, arginine, aspartic acid, glutamic acid) at pH 7; side chains are relatively strong organic acids & bases
2. Can form ionic bonds due to charges; histones with arginine (+-charge) bind to negatively charged phosphate groups of DNA
3. Histidine - usually only partially charged at pH 7; often important in enzyme active sites due to its ability gain or lose a proton in physiologic pH ranges

B. <u>Polar uncharged</u> - R groups weakly acidic or basic; not fully charged at pH 7; can form H bonds with other molecules like water since they have atoms with a partial negative or positive charge

1. Asparagine & glutamine [amides of aspartic & glutamic acid], threonine, serine, tyrosine

C. <u>Nonpolar</u> - R groups hydrophobic; generally lack O & N; cannot interact with water or form electrostatic bonds; vary primarily in size & shape; allows them to pack tightly into protein core

1. Alanine, valine, leucine, isoleucine, tryptophan, phenylalanine, methionine
2. Associate with one another via hydrophobic & van der Waals interactions in protein core

D. <u>The other three – glycine, proline, cysteine</u>

1. Glycine (R = H) - small R group makes backbone flexible & able to move so it is useful in protein hinges; small R group allows 2 backbones (of same or different protein) to approach closely
2. Proline – R group forms ring with amino group (making it an imino acid); hydrophobic amino acid that does not readily fit into orderly secondary structure (α-helix)
3. Cysteine – R group has reactive —SH; forms disulfide (—S—S—) bridge with other cysteines often at some distance away in polypeptide backbone or in another chain

a. Stabilize proteins especially outside cells where they get added chemical & physical stress

IV. Not all of amino acids are found in all proteins, nor are the various amino acids distributed in an equivalent manner

V. Character of amino acid R groups (ionic, polar, nonpolar) is very important to protein structure & function; side chains also affect solubility (amino acids can be separated on basis of solubility)

A. Most soluble (i.e., nonmembrane) proteins set up so polar residues are on molecule surface

1. They associate with surrounding H_2O & contribute to protein solubility in aqueous solution

B. Nonpolar residues situated predominantly in core of protein

VI. Conjugated proteins - involve another type of molecule attached covalently or noncovalently to protein; they include:

A. Nucleoproteins - protein + nucleic acids

B. Lipoproteins - protein + lipids

C. Glycoproteins - protein + carbohydrate

D. Various low-molecular-weight materials, like metals & metal-containing groups, often attached

VII. Proteins are good illustration of intimate relationship between form & function

A. Proteins are huge, complex molecules but their structure in given environment is completely defined & predictable

B. Each amino acid in protein is located in specific site within these giant molecules giving it the structure & reactivity required for the job it does

C. Protein structure described at several levels of organization – each emphasizes a different aspect & each is dependent on different types of interactions

1. 4 such levels are described: primary, secondary, tertiary & quaternary

2. The first, primary structure, concerns amino acid sequence of a protein; the latter 3 levels concern the organization of the molecules in space

Proteins: Levels of Protein Structure – Primary & Secondary Structure

<u>I. Primary (1°) structure</u> - specific linear sequence of amino acids in chain; all levels of structure are ultimately determined by the primary level

A. Number of chains that can be made = 20n, where n = number of amino acids in chain; most polypeptides have >100 aminos (some several 1000); variety of possible sequences is unlimited

B. Genome contains instructions for building them (precisely specifies amino acid sequence)

C. Amino acid sequence contains most, if not all, information needed to specify protein 3D shape & thus its function; changes in sequence resulting from mutation may not be readily tolerated

1. Example: sickle cell anemia - single change in amino acid sequence in hemoglobin molecule; valine replaces glutamic acid (nonpolar amino acid replaces charged, polar amino acid)

2. Changes in amino acid sequence caused by changes (mutations) in DNA; problems with red blood cell shape & decreased O_2-carrying capacity result from this change

3. Not all changes as big as above; related organisms show variations in sequence of same protein

D. Degree to which changes in primary sequence are tolerated depends on degree to which protein shape or critical functional residues are disturbed

E. Now know sequences of tens of thousands of proteins – first was protein hormone insulin determined by Sanger & coworkers, Cambridge, early 1950s

<u>II. Secondary (2°) structure </u>- describes polypeptide conformation (spatial organization) chain portions; preferred ones provide maximum possible number of H bonds between neighboring amino acids

A. α-helix - backbone assumes form of cylindrical, twisting spiral; backbone inside helix, R groups project outwards (Linus Pauling & Robert Corey, Cal Tech proposed both α- & β-structures)

1. Stabilized by H bonds between atoms of one peptide bond & those above & below it in spiral; H bonds parallel to molecular axis
2. Seen in X-ray diffraction patterns of actual proteins in 1950s; found in keratin from hair & various oxygen-binding proteins like myoglobin & hemoglobin; proof
3. Opposing surfaces of α-helix may have contrasting properties – in water-soluble proteins, often polar amino acids are on outside of helix & nonpolar R groups facing inward
4. Since it is coiled & held together by weak, noncovalent bonds, an α-helix can be extended in length if subjected to pulling forces

a. Example is wool (mostly α-helix) - H bonds break if pulled, stretching fibers; when tension is relieved, H bonds reformed & fiber shortens to original length

b. Human hair is less extensible because it is also stabilized by disulfide bridges

B. β-pleated sheet- consists of several polypeptide segments lying side-by-side; the backbone of each segment of polypeptide adopts a folded or pleated conformation

1. Characterized by a large number of H bonds perpendicular to polypeptide chain long axis; project across from one part of chain to another
2. β strands highly extended; sheet resists pulling (tensile) forces; very strong (ex.: silk fibroin)
3. A single fiber of spider silk (one-tenth the thickness of a human hair, is roughly 5 times stronger than a steel fiber of comparable weight
4. Spider silk being produced from cultured epithelial cells; hoping to use them in making strong, lightweight, resilient products like bulletproof vests

Proteins: Levels of Protein Structure – Tertiary Structure

I. Tertiary (3°) structure is the conformation of entire protein; results from (**intramolecular**) noncovalent interactions between R groups in same chain; virtually unlimited number of structures unlike limited 2°
  A. X-ray crystallography can be used to determine tertiary structure; ~20,000 3D structures already reported, increased pace in structure discovery per year
  B. NMR spectroscopy (not described here) – 3D structure of small proteins (<30 kDa) can also be determined by NMR spectroscopy
  C. Fibrous proteins (highly elongated shape) - long strands or flattened sheets that resist pulling or shearing forces to which they are exposed; structural materials outside cell are usually these
  D. Globular proteins – most proteins in cell; compact shape; chains folded & twisted into complex shapes; distant points brought next to each other, linked by various types of bonds; ex.: myoglobin

II. Protein domains - proteins often composed of 2 or more spatially distinct modules (domains) that fold independent of one another; often represent parts that function in semi-independent manner
  A. May bind various things (coenzyme & substrate; DNA strand & another protein) or move relatively independent of one another
  B. Proteins with >1 domain may have arisen during evolution by fusion of genes coding for different ancestral proteins; each domain representing a part that once was separate molecule
  C. Some domains may have been shuffled widely about during evolution; appear in variety of proteins whose other regions show little or no evidence of an evolutionary relationship
  D. Domain shuffling creates proteins with unique combinations of activities

III. Protein motifs - recurring protein substructures; common motifs occur in evolutionarily related proteins with similar function or unrelated ones with different function
  A. Defined arrangement of α-helices and/or β-strands
  B. Example: coiled coil - seen in some fibrous proteins (myosin); 2 or more α-helices coil about one another like entwined strands of cable
  C. Example: α/β barrel - complex; first seen in triosephosphate isomerase & up to ~10% of all enzymes

Proteins: Levels of Protein Structure -
Quaternary Structure & Multiprotein Complexes

I. Quaternary (4°) structure is the linking of polypeptide chains to form multisubunit functional protein via **intermolecular** R group interactions
  A. May be linked by disulfide bonds, but more often noncovalent bonds (hydrophobic, H bonds, etc.) like hydrophobic patches on complementary surfaces of neighboring polypeptides

B. Chains may be identical or nonidentical
   1. Protein composed of 2 identical subunits - homodimer
   2. Protein composed of 2 nonidentical subunits - heterodimer

II. Multiprotein complexes – different proteins, each with a specific function, become physically associated to form a much larger complex
   A. Example - *E. coli* pyruvate dehydrogenase complex - 60 polypeptide chains constituting 3 different enzymes; stable
      1. Its enzymes catalyze reaction series connecting 2 metabolic pathways: glycolysis & TCA cycle
   B. Because they are physically associated, the product of one enzyme can be channeled directly to next enzyme in sequence; prevents dilution in cell's aqueous medium
   C. Some associations stable; some not (transient; dynamic; associate, dissociate based on conditions); have complementary surfaces (part of one fits in pocket on other); stabilized by noncovalent bonds
      1. General rule – most proteins interact with other proteins in highly dynamic patterns;
      2. Ex.: SH3 domain – part of many different proteins involved in molecular signaling; act like knobs that allow binding to proteins with complementary handle. In this case, handle is rich in proline (a polyproline motif)
   D. A number of different structural domains like SH3 identified that function as adaptors to mediate interactions between proteins
      1. Often interactions between proteins are regulated by changes like phosphate addition to key amino acid; may greatly increase or decrease its ability to bind a protein partner
      2. Transient protein interaction important in DNA synthesis, ATP formation, RNA processing, etc.; done by molecular machines made of many interacting proteins (transient or stable)

III. Chart below summarizes features and definitions of the four levels of protein structure:

Level of Structure	Definition	Bonds Involved	Comments
Primary (1°)	Absolute sequence of amino acids from amino end to carboxyl end	Peptide bonds	All 3 higher levels are direct consequences of 1° structure (contains information about their final shape). Changes can lead to disease (ex. sickle cell) or little or no effect.
Secondary (2°)	Results from interactions between backbone portions of adjacent or nearly adjacent amino acids	H bonds	α-helix - spiral shaped; H bonding maximal and parallel to main molecular axis of helix; allows extensibility (ex.: wool & human hair). β-pleated sheet - highly flattened, extended sheetlike shape; H bonding maximal and perpendicular to main molecular axis; strong and flexible (ex.: silk fibroin). Without α or β structure adopts hinges, turns, loops or fingerlike extensions with most biological activity.
Tertiary (3°)	Results from interactions within a single chain between R groups or between R groups at a distance and backbone	H bonds, disulfide bonds, van der Waals forces, ionic bonds, hydrophobic interactions	Proteins fibrous (highly elongated like collagen) or globular (myoglobin). Protein domains - compact regions functioning semi-independently (linked by flexible part of chain serving as hinge). Protein motifs - recurring protein substructures with certain functions. Proteins flexible & can change shape.
Quaternary (4°)	Results from R group interactions between multiple protein chains (subunits) which form a functional protein unit	H bonds, disulfide bonds, van der Waals forces, ionic bonds, hydrophobic interactions	Assembly spontaneous & usually bound together by noncovalent bonds (electrostatic or hydrophobic). Homodimers - 2 identical subunits; heterodimers - at least 2 nonidentical subunits (Ex.: hemoglobin)

Nucleic Acids

I. Primarily involved in storage & transmission of genetic information; may also be structural or catalytic

A. 2 types in living organisms – deoxyribonucleic acids (**DNA**) & ribonucleic acids (**RNA**)

1. DNA – serves as genetic material of all cellular organisms; RNA plays that role in many viruses
2. Information stored in DNA used to govern cell activities through formation of RNA messages

B. Constructed as long chain (**strand**) of monomers called **nucleotides**

C. Concentrate here on RNA as representative molecule; more complex, double-stranded DNA structure will be covered in Chapter 10

II. Both DNA & RNA are composed of nucleotides (phosphate + sugar + base) connected to form polymers (**polynucleotides**)

A. Phosphate group (PO4-) - linked to 5'-carbon of sugar

B. 5-carbon sugar - ribose or deoxyribose

C. Nitrogenous base (pyrimidine -1 ring or purine -2 rings) - rings contain nitrogen; linked to 1'- carbon of sugar (adenine, guanine, thymine [not in RNA], cytosine, uracil [not in DNA])

1. Called nitrogenous bases because nitrogen atoms form part of the rings of the molecules
2. **Purines** in RNA & DNA are **adenine** & **guanine;** larger structure, consisting of 2 rings
3. **Pyrimidines** in RNA are **cytosine** & **uracil**; in DNA, **uracil** replaced by **thymine**, a pyrimidine with an extra methyl group attached to the ring; smaller, single ring structure

III. The sugar & nitrogenous base together form a **nucleoside**, so nucleotides of RNA strand are known as **ribunucleoside monophosphates**

IV. Monomers polymerize when sugar 3'-OH is linked by ester bond to 5'-phosphate of next nucleotide in chain (**3' - 5' phosphodiester linkage**); so nucleotides are joined by sugar-phosphate linkages

A. Phosphates in backbone attached to 2 sugars by ester linkages (phosphorus atom linked to 2 oxygen atoms, one from each of the 2 adjoining sugars)

B. Have hydrophilic, charged backbone (repeating P & sugar units) & nitrogenous bases as side groups

C. Bases largely hydrophobic due to ring structure

V. General structure and function of DNA/RNA

A. Sequence of bases determines specificity of DNA/RNA & encodes hereditary information for synthesis of proteins

B. RNA is usually single-stranded, but can fold back on itself to form frequent double stranded regions with local H-bond pairing (just like in DNA) & complex 3° structures

1. With DNA, it directs & carries out protein synthesis
2. Some RNAs do not carry genetic information but can be structural and/or enzymatic (rRNA)