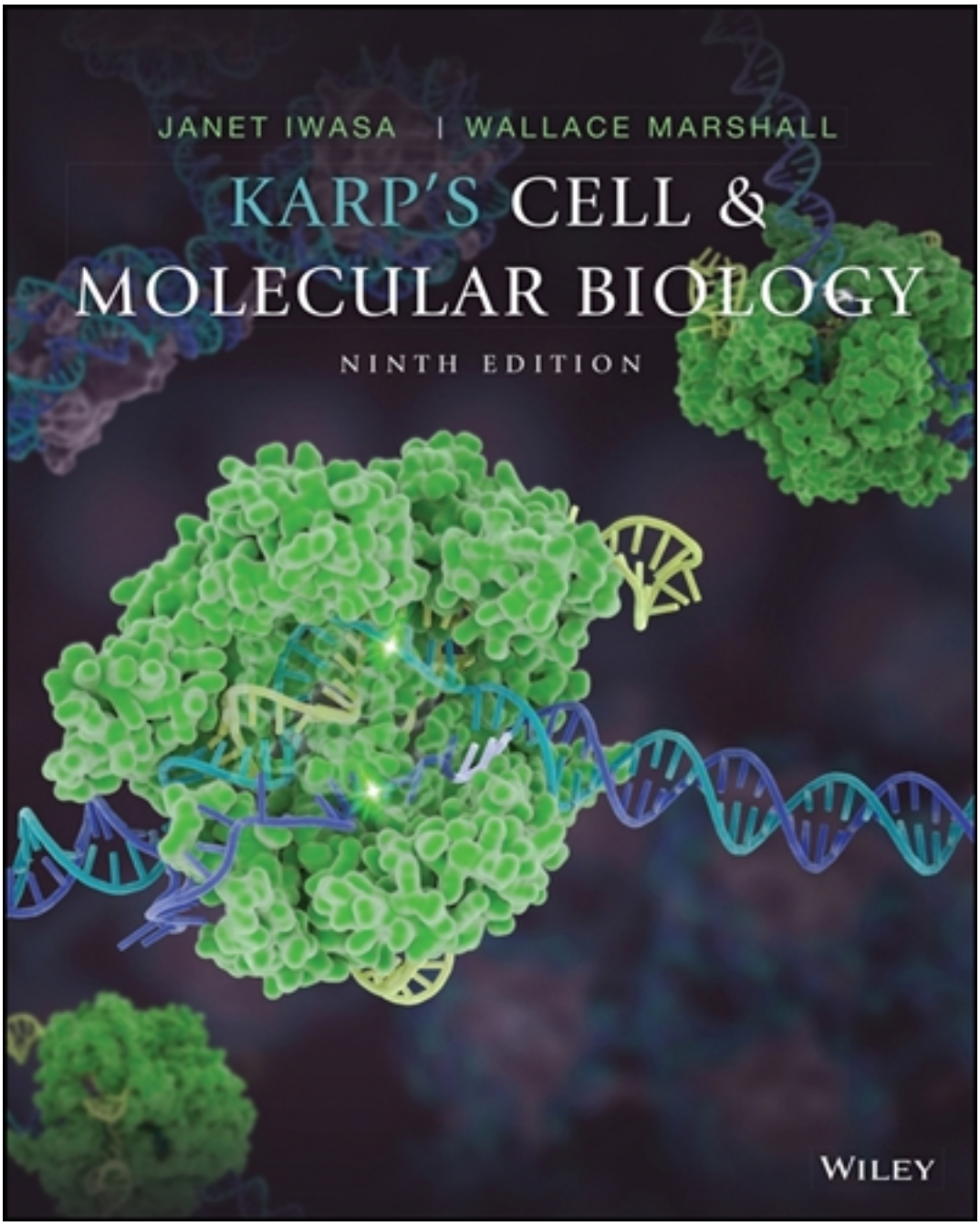
Solutions for Karp's Cell and Molecular Biology 9th Edition by Karp

CLICK HERE TO ACCESS COMPLETE Solutions



Solutions

2 The Chemical Basis of Life

CASE STUDY: Good vs. Bad Fats?

There is much information in the popular press touting the virtues of this oil versus that oil, or this margarine versus that margarine. Beyond the obvious marketing strategies used by many companies, is there any scientific basis for why one fat is better or worse for you than another? Furthermore, why do things like French fries and potato chips that taste so good make us fat? If we look into the nutritional information regarding several of the biological building blocks (proteins, carbohydrates, and fats) that we get from our food, one will find that 1 g of protein or carbohydrate contains 4 calories whereas 1 g of fat contains 9 calories. That means we need much less fat to meet our body's energy requirements than we do for protein or carbohydrates. Fat is readily stored in our body tissues whereas carbohydrates and proteins must be metabolized to be converted into fat. Despite the bad press that fats often receive, they are essential building blocks for our cells and serve multiple physiologically important roles.

But what is the difference between fats? This comes down to the structure of the fatty acid side chains and how the fats are processed in our bodies. Fats can be saturated in which their hydrocarbon chains have no double bonds. Unsaturated fats have double bonds along the length. These double bonds create kinks in the side chains, which inhibit the ability of the chains to pack tightly together. This in turn will affect the temperature at which the transition between liquid and solid occurs. Saturated fatty acids are typically solid at room temperature, whereas unsaturated fatty acids are liquid. In some cases, saturated fatty acids may remain solid at human body temperature, as well. In addition saturated fats are readily converted into LDL cholesterol, high levels of which have been associated with heart disease.

Questions:

1. You are about to make a delicious batch of chocolate chip cookies. You debate between using butter versus margarine. Given that butter has a lot of saturated fat and margarine has a higher percentage of unsaturated fat, which is the healthier choice and why?

Answer: Margarine is the better choice because of its reduced levels of saturated fat. One misnomer is that the fat that clogs the arteries. In reality it is the altered metabolism of fats and the creation of cholesterol that appears to be the major culprit.

2. The biological properties of molecules are also very important in cooking. If the recipe for those chocolate chip cookies calls for 2 sticks of butter softened at room temperature, will the texture of the cookie dough be the same if you substitute 2 sticks of margarine that has no saturated fat? Why or why not? Provide an explanation based on your knowledge of the structure of the fatty acid side chains.

Answer: *The texture of the cookie dough will actually be softer with the margarine. Because the butter is rich in saturated fats it is more solid at room temperature due to the tight packing of the hydrocarbon*

side chains. In contrast the melting temperature of the margarine will be lower because the unsaturated side chains cannot pack so tightly together.

Where can I learn more?

1. https://www.ncbi.nlm.nih.gov/pubmed/27648593

2. Butter vs. Margarine. Harvard Health Publishing. https://www.health.harvard.edu/staying-healthy/butter-vs-margarine

1. The pH of cola beverages is around 3, while the pH of battery acid is around 1. Which has a higher concentration of protons? By how much? If the concentration of protons in wine is approximately one-tenth that in cola, what is the pH of wine?
   Ans: Since the pH of battery acid is lower, it has a higher concentration of protons. Since pH is proportional to the log base ten of the proton concentration, a difference of two pH units corresponds to a hundred-fold difference in proton concentration, so the concentration is 100 times higher in the battery acid. Since the concentration of proteins is one tenth that in cola, its pH has to be one pH unit higher, so the pH of wine should be approximately 4.

   Difficulty: Medium
   Section Reference:
   2.4

2. Sickle cell anemia results from a substitution of a valine for a glutamic acid. What do you expect the effect might be if the mutation were to have placed a leucine at that site? An aspartic acid?
   Ans: Since leucine is similar in structure to valine, one would expect the protein to have similar abnormal properties. Since aspartic acid is similar to glutamic acid, which is the proper amino acid, relatively little effect might be expected, but this could only be confirmed by site-directed mutagenesis.

   Difficulty: Difficult
   Section Reference:
   2.8

3. Of the following amino acids, glycine, isoleucine, and lysine, which would you expect to be the most soluble in an acidic aqueous solution? Which the least?
   Ans: Lysine would be the most soluble and isoleucine the least.

   Difficulty: Medium
   Section Reference:
   2.8

4. How many structural isomers could be formed from a molecule with the formula C_5H_{12}? C_4H_8?
   Ans: Four (one straight chain and three branched). Three (two straight chain and one branched), ignoring cis-trans isomerism.

   Difficulty: Difficult
   Section Reference:
   2.6

5. Glyceraldehyde is the only three-carbon aldotetrose, and it can exist as two stereoisomers. What is the structure of dihydroxyacetone, the only ketotriose? How many stereoisomers does it form?
   Ans: The structure of DHAP is on p. 106. Dihydroxyacetone is similar to DHAP except that the phosphate group is replaced with a hydroxyl. Because the middle carbon has a double bonded oxygen, it is not a chiral center, hence dihydroxyacetone does not have stereoisomers.

   Difficulty: Difficult
   Section Reference:
   2.6

6. Bacteria are known to change the kinds of fatty acids they produce as the temperature of their environment changes. What types of changes in fatty acids would you expect as the temperature drops? Why would this be adaptive?
   Ans: As the temperature drops, cells face a challenge of maintaining the fluidity of their membranes. Thus, to adapt to decreasing temperatures, we can predict that bacteria should produce fatty acids that are shorter or more unsaturated, thus allowing the membranes to remain fluid at lower temperatures.

   Difficulty: Medium
   Section Reference:
   2.7

7. In the polypeptide backbone C-C-N-C-C- N-C-C-NH_2, identify the alpha carbons
   Ans: -C-C*-N-C-C*-N-C-C*-NH_2. Carbons with an asterisk are alpha carbons.

   Difficulty: Easy
   Section Reference:
   2.8

8. Which of the following are true? Increasing the pH of a solution would (1) suppress the dissociation of a carboxylic acid, (2) increase the charge on an amino group, (3) increase the dissociation of a carboxylic acid, (4) suppress the charge on an amino group.
   Ans: 3 and 4. A rise in pH, would increase the loss of a proton from a -COOH group and from an -NH_3+ group.

   Difficulty: Medium
   Section Reference:
   2.8

9. Which of the four classes of amino acids has side chains with the greatest hydrogen-bond-forming potential? Which has the greatest potential to form ionic bonds? Hydrophobic interactions?
   Ans: Polar, uncharged. Polar, charged. Nonpolar.

   Difficulty: Easy
   Section Reference:
   2.8

10. If the three enzymes of the pyruvate dehydrogenase complex existed as physically separate proteins rather than as a complex, what effect might this have on the rate of reactions catalyzed by these enzymes?
    Ans: It would be expected to slow the rates of the reaction because the products of the first two reactions would diffuse into the surrounding medium rather than be passed directly into the active sites of the second and third enzymes.

    Difficulty: Difficult
    Section Reference:
    2.11

11. Would you agree that neither ribonuclease nor myoglobin had quaternary structure? Why or why not?
    Ans: Yes. Both consist of a single polypeptide chain.

    Difficulty: Easy
    Section Reference:
    2.10

12. How many different tripeptides are possible? How many carboxyl terminals of polypeptide chains are present in a molecule of hemoglobin?
    Ans: Because there are three positions in the chain, and each position can have any of 20 amino acids, the number of possible chains is 8,000 (20^3). See the accompanying online Quantitative Tutorial video for explanation of the first answer. Since each hemoglobin molecule contains four polypeptide chains, each of which has its own N and C terminus, there are a total of 4 carboxyl terminals, one per chain.

    Difficulty: Difficult
    Section Reference:
    2.8

13. You have isolated a pentapeptide composed of four glycine residues and one lysine residue that resides at the C-terminus of the peptide.Using the information provided in the legend of Figure 2.27, if the pK of the side chain of lysine is 10 and the pK of the terminal carboxyl group is 4, what is the structure of the peptide at pH 7? At pH 12?
    Ans: At pH 7, the amino group of lysine will be positively charged and the carboxyl group negatively charged. At pH 12, the amino group will be uncharged and the carboxyl group will be negatively charged.

    Difficulty: Difficult
    Section Reference:
    2.8

14. The side chains of glutamic acid (pK 4.3) and arginine (pK 12.5) can form an ionic bond under certain conditions. Describe the relevant portions of the side chains and indicate whether or not an ionic bond could form at the following: (a) pH 4; (b) pH 7; (c) pH 12; (d) pH 13.
    Ans: Glutamic acid has a pK of 4.3. Therefore, above pH 4.3, the side chain is predominantly negatively charged and below pH 4.3 uncharged. Arginine has a pK of 12.5. Therefore, below pH 12.5, the side chain is predominantly positively charged and above pH 12.5 uncharged. For an ionic bond to form, both side chains must be charged. At pH 4, arginine is charged, but glutamic acid is uncharged; at pH 13, glutamic acid is charged, but arginine is uncharged. However, at pH 7 and 12, both groups will be charged and an ionic bond can form.

    Difficulty: Medium
    Section Reference:
    2.10

15. Would you expect a solution of high salt to be able to denature ribonuclease? Why or why not?
    Ans: No. The disulfide bonds would not be broken.

    Difficulty: Medium
    Section Reference:
    2.10

16. You have read in the Human Perspective that (1) mutations in the *PRNP* gene can make a polypeptide more likely to fold into the PrPSc conformation, thus causing CJD and (2) exposure to the PrPSc prion can lead to an infection that also causes CJD. How can you explain the occurrence of rare sporadic cases of the disease in persons who have no genetic propensity for it?

    Ans: There are two possibilities. 1) Even though the normal polypeptide normally folds into PrP^{C}, it is possible that on rare occasions it might fold abnormally to form PrP^{Sc}. If this were to occur, the formation of one or a few of the abnormal molecules could lead to a chain reaction that converted normal PrP molecules to the PrP^{Sc} form. 2) A mutation in a gene could arise in a somatic cell that led to the formation of the abnormal PrP^{Sc} in that cell. If that protein was released into the extracellular environment, it could be picked up by other cells whose PrP^{C} could then be converted to the abnormal form, thus causing the disease.

    Difficulty: Difficult
    Section Reference:
    2.13

17. Persons who are born with Down syndrome have an extra (third) copy of chromosome #21 in their cells. Chromosome #21 contains the gene that encodes the APP protein. Why do you suppose that individuals with Down syndrome develop Alzheimers disease at an early age?

    Ans: If CJD results from a rare, sporadic misfolding of the PrP protein, then the longer a person lives, the more likely the event will occur, and thus the older they are likely to be when they are struck by the disease.

    Difficulty: Medium
    Section Reference:
    2.13

18. We saw on page 74 how evolution has led to the existence of protein families composed of related molecules with similar functions. A few examples are also known where proteins with very similar functions have primary and tertiary structures that show no evidence of evolutionary relationship. Subtilisin and trypsin, for example, are two protein-digesting enzymes (proteases) that show no evidence they are homologous despite the fact they utilize the same mechanism for attacking their substrates.How can this coincidence be explained?
    Ans: Convergent evolution. These examples indicate that a similar grouping of reactive side chains within the active site of a protein can evolve from ancestral proteins that are not related. Another example is provided in Figure 2.37.

    Difficulty: Medium
    Section Reference:
    2.10

19. Would you agree with the statement that many different amino acid sequences can fold into the same basic tertiary structure? What data can you cite as evidence for your position.
    Ans: Yes. Many different sequences generate the same globin fold that is depicted in Figure 2.34a.

    Difficulty: Easy
    Section Reference:
    2.10

20. In the words of one scientist: ìThe first question any structural biologist asks upon being told that a new [protein] structure has been solved is no longer 'What does it *look* like?; it is now 'What does it look *like*?' î What do you suppose he meant by this statement?
    Ans: Answers will vary.

    Difficulty: Medium
    Section Reference:
    2.10

Chapter 2: Review Questions
Karp, 8th edition

Section 2.1

1) Oxygen atoms have eight protons in their nucleus. How many electrons do they have? How many orbitals are in the inner electron shell? How many electrons are in the outer shell? How many more electrons can the outer shell hold before it is filled?

Answer: 8 electrons. 1 orbital. 6 electrons in outer shell. 2 more electrons to fill outer shell.

2) Compare and contrast: a sodium atom and a sodium ion; a single bond and a double bond; an atom of weak and strong electronegativity; the electron distribution around an oxygen atom bound to another oxygen atom and an oxygen atom bound to two hydrogen atoms.

Answer:

- **Sodium atom/sodium ion**: A sodium atom has no net charge while a sodium ion has a net charge of +1.

- **Single bond/double bond**: Covalent bonds where two atoms share 1 pair of electrons are single bonds; double bonds occur when 2 pairs of electrons are shared.

- **Atom/weak electronegativity/strong electronegativity**: An atom has a positively charged nucleus. The greater the positive charge of the nucleus of an atom, the greater attraction it has for shared electrons and the stronger its electronegativity. Nitrogen and oxygen are strongly electronegative while hydrogen or lithium are weakly electronegative.

- **Electron distribution around O_2 versus H_2O**: The two O atoms in O_2 will share the electrons equally, whereas shared electrons in the H_2O molecule will be located closer to the O atom.

Section 2.3

1) Describe some of the properties that distinguish covalent and non-covalent bonds.

Answer: Covalent bonds depend on shared electrons between atoms and are strong; they require a large amount of energy to break. Non-covalent bonds do not depend on shared electrons and instead depend on attractive forces between atoms having an opposite charge. Individual non-covalent bonds are weak, but when there are many of these bonds, the resulting structure can be quite strong.

2) Why do polar molecules, such as table sugar, dissolve so readily in water? Why do fat droplets form on the surface of an aqueous solution? Why does sweating help cool the body?

Answer: Polar molecules dissolve readily in water because they have charged regions that attract the poles of the water molecules. Fat droplets are comprised of nonpolar molecules that will be

forced together when exposed to water. As a result fat droplets will form on top of an aqueous solution. Heat is required for water to evaporate and this heat can be obtained from the body. As a result, sweating cools the body.

Section 2.4

1) If you were to add hydrochloric acid to water, what effect would this have on the hydrogen ion concentration? On the pH? On the ionic charge of any proteins in the solution?

Answer: Increases. pH decreases. Ionic charge of proteins in solution will increase.

2) What is the relationship between a base and its conjugate acid?

Answer: A base that accepts a proton forms a conjugate acid.

Section 2.5

1) What properties of a carbon atom are critical to life?

Answer: Its ability to form an enormous number of molecules is due to its size and electronic structure.

2) Draw the structures of four different functional groups. How would each of these groups alter the solubility of a molecule in water?

Answer:

1. Methyl. Decrease.
2. Hydroxyl. Increase.
3. Phosphate. Increase.
4. Amino. Increase.

Section 2.6

1) Name three polysaccharides composed of polymers of glucose. How do these macromolecules differ from one another?

Answer: Glycogen, starch, and cellulose. Glycogen is a branched polymer of glucose with an α-1,4 linkage or a with an α-1,6 linkage, starch molecules are a less-branched structure with an α-1,4 linkage or a with an α-1,6 linkage, and cellulose molecules are unbranched and highly extended with a β-1,4 linkage.

Section 2.7

1) Describe the properties of three different types of lipid molecules. What are their respective biological roles?

Answer:

- **Fats:** Glycerol linked by ester bonds to 3 fatty acids. Are used as an extremely concentrated means of storing energy over longer periods of time.

- **Steroids**: A ringed hydrocarbon skeleton. A major component of animal cell membranes and a precursor for many steroid hormones.

- **Phospholipids**: Resembles a fat except only 2 fatty acids and a phosphate group. An amphipathic molecule used in cell membranes.

Section 2.8

1) What are the major properties that distinguish different amino acids from one another? What roles do these differences play in the structure and function of proteins?

Answer: The side chains or R groups. The properties of these groups are what lead to the specific structure and function of each protein.

2) What are the properties of glycine, proline, and cysteine that distinguish these amino acids?

Answer: Glycine: R group consists of H atom only.
Cysteine: R group is a sulfhydryl/thiol group, polar and uncharged, but forms a covalent bond with another cysteine to form a disulfide link.
Proline: R group is a hydrophobic imino ring, and causes kinks/hinges in polypeptide chains.

Section 2.9

1) How are the properties of an α helix different from a β strand? How are they similar?

Answer: Both are similar in that they are stabilized by hydrogen bonds between the peptide bonds of amino acids. α-helices are stabilized by H-bonds within a single polypeptide chain and form a coiled, cylindrical helix. β strand has a pleated or folded conformation with hydrogen bonds oriented perpendicular to the long axis of the chain and project across one part of the chain to another.

Section 2.10

1) What are some of the differences between X-ray crystallography and NMR for determining protein structure? What can NMR tell us that X-ray crystallography cannot?

Answer: The two methods differ in the type of physical radiation used to probe protein structure: X-ray crystallography uses X-rays, while NMR uses radio waves. X-ray crystallography requires that the proteins be formed into solid crystals, while NMR works on proteins in solution. Because of this latter difference, NMR can be used to observe conformation changes in proteins much more easily than by X-ray crystallography.

Section 2.11

1) Describe the difference between primary, secondary, tertiary, and quaternary structure.

Answer: Primary structure refers to the amino acid sequence, without regard to the geometrical arrangement of the amino acids in space. Secondary structure refers to which of several standard patterns the amino acid chain is arranged in, namely alpha helix or beta sheet. Tertiary structure refers to the completely three dimensional arrangement of the amino acids within a given protein. Quaternary structure refers to the spatial arrangement of multiple subunits in a multi-protein complex.

Section 2.12

1) Given that proteins act as molecular machines, explain why conformational changes are so important in protein function.

Answer: Function is directly related to protein structure. Conformational changes lead to proteins that are folded so that specific residues are ideally situated to perform unique functions. Each amino acid in a peptide can lend a specific charge, hydrophobicity, or structure. Thus, different conformations affect the function of a protein, sometimes by stopping the protein from functioning and other times by changing its function.

Section 2.15

1) Which of the two methods discussed in this section, proteomics or interactomics, gives us information about the primary structure of proteins? Which gives information about quaternary structure of protein complexes?

Answer: Proteomics generally relies on protein sequencing methods to determine which proteins are present in a mixture. It is thus a method based on primary structure. Interactomics tell us about protein complexes, thus providing information about quaternary structure.

Section 2.16

1) What are some of the ways that knowledge of protein structure can help with developing new drugs?

Answer: Protein structure can be used to predict chemicals that might bind to the protein, using computer programs that can predict molecular interacxtions. Knowledge of protein structure can also be used to design novel peptides or proteins that can themselves be used as drugs.

Section 2.17

1) Can a single amino acid change alter the secondary or tertiary structure of a protein?

Answer: Sometimes this can happen. An example is shown in Figure 2.53. But it is also possible to have single amino acid changes that leave secondary and tertiary structures intact.

Section 2.18

1) Describe the structure of nucleotides and the manner in which these monomers are joined to form a polynucleotide strand. Why would it be overly simplistic to describe RNA as a single-stranded nucleic acid?

Answer: Nucleotides consist of: (1) 5-carbon sugar—ribose or deoxyribose (2) a nitrogenous base (3) a phosphate group. Nucleotides are linked with a 3'-5' phosphodiester bond along a phosphate-ribose backbone. This would be simplistic since RNA strands can fold back on themselves to have double-stranded and complex 3-dimensional structures.

2) Which macromolecules are polymers? What is the basic structure of each type of monomer? How do the various monomers of each type of macromolecule vary among themselves?

Answer: Proteins, nucleic acids, polysaccharides, and some lipids. Monomers are low molecular weight building blocks. The sequence of the various monomers gives rise to the diversity of individual macromolecules.

Section 2.19

1) What type of evidence suggests that bacterial ribosomal subunits are capable of self-assembly, but eukaryotic subunits are not?

Answer: Complete, functional bacterial ribosomal subunits can be formed by mixing the purified proteins of the ribosomal subunit with purified small-subunit ribosomal RNA *in vitro*. Eukaryotic subunits cannot be reconstituted this way.

2) What evidence would indicate that a particular ribosomal protein had a role in ribosome function but not assembly?

Answer: Deletion of a particular ribosomal protein that does not alter the rate of correct assembly, so complete mRNA-ribosome complexes could be observed. However, there would be an absence or decrease in translation.

CHAPTER 2

THE CHEMICAL BASIS OF LIFE

OBJECTIVES

- Describe the role of electrons in the formation of covalent bonds.
- Explain the chemical basis of the use of radionuclides in imaging and treatment.
- Describe the role of noncovalent bonds in the structure of molecules such as water.
- Explain the characteristics of acids, bases, and buffers.
- Describe the general structure and functions of biological molecules.
- Identify the chemicals in fertilizer that are crucial to plant growth.
- Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.
- Analyze the evidence supporting the idea that bacterial ribosomal subunits are capable of self-assembly.

LECTURE OUTLINE

(2.0) The Chemical Origin of Life

I. Around 4.6 billion years ago, our solar system was formed from an enormous rotating cloud of gas and dust.
   A. For its first billion years, Earth was a tumultuous place, with violent volcanic eruptions and near constant collisions with asteroids.
      1. During this very time period, scientists believe that life, in the form of primitive cells, first appeared on the planet.
      2. Ancient microbes are thought to have formed stratified rock formations, called stromatolites, suggesting that life may have proliferated as early as 3.5 billion years ago.
   B. Researchers have hypothesized that the earliest cells (protocells) were very simple, made up of just nucleic acids (DNA and RNA) surrounded by a membrane.
      1. These cells may have formed in warm pools of water or in the ocean near deep-sea vents.
      2. A central dilemma, however, has been in understanding how nucleic acid and membrane molecules formed spontaneously on the young planet – how, essentially, biology was born from chemistry.

II. In 1952, Harold Urey and his graduate student, Stanley Miller, designed an experiment to test whether the conditions on the early Earth favored the spontaneous synthesis of biological molecules.
   A. They simulated Earth's early atmosphere by circulating water, methane, ammonia and hydrogen in a sealed glass apparatus and introduced energy in the form of heat and electricity, which mimicked lightning's effect.

B. Over 2 weeks, the glass became coated with with organic compounds that included a variety of amino acids and sugars.
   1. This supported the idea that conditions on early Earth may have been ideal for creating organic compounds that were eventually incorporated into early cells.
   2. In recent years, researchers have found prebiotically feasible pathways to create a number of additional molecules, like ribonucleotides and fatty acids, the building blocks of RNA and membranes.

C. More mysteries remain – molecules made by the Miller-Urey experiment tend to be an even mixture of left- and right-handed isomers (which are mirror images of each other).
   1. But all life on Earth use only left-handed amino acids and right-handed sugars.
   2. One wonders how and why this selection came about.

(2.1) Covalent Bonds

I. Properties of cells and their organelles derive directly from activities of the molecules of which they are composed.
   A. It is impossible to begin to understand cell function without a reasonable knowledge of the structure and properties of the major types of biological molecules.
   B. An understanding of the chemistry that allows biological events to occur is essential to understanding how life arose.

II. Molecular atoms are joined together by **covalent bonds** (electron pairs are *shared* between pairs of atoms).
   A. Formation of a covalent bond is governed by the basic principle that an atom is most stable when its outermost electron shell is full.
      1. The number of bonds an atom forms is determined by how many electrons are needed to fill its outer shell.
      2. The outer (and only) shell of hydrogen and helium atoms is filled when it contains 2 electrons; the outer shells of other atoms commonly found in living cells are filled when they contain 8 electrons.
      3. Oxygen with 6 outer-shell electrons can fill its outer shell by combining with 2 H atoms, forming a water molecule; the oxygen atom is linked to each H atom by a single covalent bond (H:O or H– O).
   B. Formation of a covalent bond is accompanied by the release of energy, which must be reabsorbed at a later time if the bond is to be broken.
      1. The energy required to cleave C—C, C—H or C—O covalent bonds is quite large (typically 80 - 100 kilocalories per mole (kcal/mole) of molecules.
         a. By comparison, the thermal energy of a molecule is only 0.6 kcal/mole.
         b. The thermal vibrations acting on a molecule are thus far too weak to break a covalent bond, making these bonds stable under most conditions.
      2. When we speak of bonds being strong, we mean that the energy required to break the bond is much greater that the thermal energy of the molecule.
      3. Conversely, when we talk about a bond being weak, we mean that the energy required to break the bond is of the same magnitude or smaller than the thermal energy.
   C. In many cases, 2 atoms can be joined by bonds in which >1 pair of electrons are shared:
      1. If 2 pairs of electrons are shared, it's a double bond (O_2).
      2. If 3 pairs of electrons are shared, it's a triple bond (N_2).
      3. No quadruple bonds are known.

D. The type of bond between atoms is important in determining the shape of molecules.
   1. Atoms joined by single bond can rotate relative to one another.
   2. Atoms of double and triple bonds cannot rotate relative to one another.

E. Also, double bonds can function as energy-capturing centers, driving such vital processes as respiration and photosynthesis.

III. Equal vs. unequal sharing in covalent bonds.

A. When atoms of the same element bond to one another (H_2) the electron pairs of the outer shell are shared equally between the 2 atoms.

B. When 2 unlike atoms are covalently bonded, however, the positively charged nucleus of one atom exerts a greater attractive force on the outer electrons than the other.
   1. Consequently, the shared electrons tend to be located more closely to the atom with the greater attractive force; this atom is the more **electronegative atom**.
   2. Among atoms most often seen in biological molecules, nitrogen and oxygen are strongly electonegative.
   3. The O–H bonds in H_2O are polarized; its single O atom attracts electrons much more forcefully than do either of its H atoms; O atom has partial negative charge; the other H has a partial positive charge.
   4. Such a molecule is a **polar** molecule; it has an asymmetric charge distribution or *dipole.*
   5. Biologically important polar molecules have 1 or more electronegative atoms - usually O, N and/or S.

C. Molecules without electronegative atoms and strongly polarized bonds (those made entirely of C and H atoms) are called **nonpolar**.
   1. The presence of strongly polarized bonds is of utmost importance in determining the reactivity of molecules.
   2. Large nonpolar molecules (like waxes and fats) are relatively inert.
   3. Some of the more interesting biological molecules, including proteins and phospholipids, contain both polar and nonpolar regions, which behave very differently.

IV. Ionization - some atoms are so strongly electronegative that they can capture electrons from other atoms during a chemical reaction.

A. Mix sodium (Na; a silver-colored metal) and chlorine (Cl; a toxic gas) together they form table salt.
   1. The single electron in each Na outer shell migrates to the electron-deficient chlorine atom.
   2. Each atom thus becomes transformed into charged atoms (**ions**).
   3. Because the chloride ion has an extra electron relative to the number of protons in its nucleus, it has a negative charge (Cl^-) and is called an **anion**.
   4. The sodium atom has lost an electron and has an extra positive charge (Na^+); it is called a **cation**.
   5. When present in crystals, these 2 ions form sodium chloride or table salt.

B. Ions like Na^+ and Cl^- are relatively stable because they have filled outer shells

C. A different electron arrangement within an atom can produce a highly reactive species (*free radical*).

THE HUMAN PERSPECTIVE: DO FREE RADICALS CAUSE AGING?

I. One biological factor that has long been imagined to drive aging is the gradual accumulation of damage to bodily tissues – appealing because it introduces a natural element of time.
   A. As long as damage occurs at some low, constant rate, the longer you live, the more damage you accumulate.
      1. Although it is appealing, testing the idea requires knowing the source of the damage and then asking whether changing the rate of damage actually alters aging.
      2. The most destructive damage probably occurs to DNA.
      3. DNA alterations lead to the production of faulty genetic messages that promote gradual cellular deterioration.
   B. How does cellular damage occur and why is it more rapid in a shorter-lived animal, like a chimpanzee than a human? The answer may reside at the atomic level.

II. Atoms are stabilized when their shells are filled with electrons.
   A. Electron shells consist of orbitals, each of which can hold a maximum of 2 electrons.
      1. Atoms or molecules that have orbitals containing a single unpaired electron tend to be highly unstable; they are called *free radicals*.
      2. Free radicals may be formed when a covalent bond is broken, such that each atom from the bond keeps one-half of the shared electrons.
      3. They may be formed when an atom or molecule accepts a single electron transferred during an oxidation-reduction reaction (example: $H_2O \rightarrow HO^{\cdot}$ [hydroxyl radical] + $H^{\cdot}$).
         a. Water can be converted into free radicals when exposed to radiation from the sun.
   B. Free radicals are extremely reactive and can chemically alter many types of molecules (proteins, nucleic acids, lipids).
      1. Certain cells of the immune system generate free radicals within their cytoplasm as a means to kill bacteria that these immune cells have injested.
      2. Formation of OH radicals may be major reason that sunlight is so damaging to the skin.

III. Denham Harman (Univ. of Nebraska, 1956) proposed that aging results from tissue damage caused by free radicals; no significant interest because biologists/doctors were not familiar with free radicals.

IV. Joe McCord and Irwin Fridovich (Duke Univ., 1969) discovered the enzyme superoxide dismutase (SOD), whose sole function was the destruction of the superoxide radical ($O_2^{\cdot -}$).
   A. Superoxide free radical is formed when molecular oxygen picks up an extra electron.
      1. SOD catalyzes the following reaction: $O_2^{\cdot -} + O_2^{\cdot -} + 2H^+ \rightarrow H_2O_2$ (hydrogen peroxide) + O_2
      2. H_2O_2 is also a potentially reactive oxidizing agent (often used as disinfectant and bleaching agent); it is normally destroyed in the cell by the enzymes catalase or glutathione peroxidase.
      3. If it is not rapidly destroyed, hydrogen peroxide can break down to form hydroxyl radicals that attack the cell's macromolecules.
   B. Subsequent research has revealed that superoxide radicals are formed in cells during normal oxidative metabolism and that SOD is present in cells of diverse organisms, from bacteria to humans.
      1. In fact, animals possess 3 different versions (isoforms) of SOD: a cytosolic, mitochondrial, and extracellular isoform.
      2. It is estimated that 1 – 2% of the oxygen taken into human mitochondria is converted to H_2O_2 rather than to water, the normal end product of respiration.

C. The importance of SOD is most clearly revealed in studies of mutant bacteria and yeast lacking SOD; these cells are unable to grow in the presence of oxygen.
   1. Similarly, mice lacking the mitochondrial SOD (SOD2) cannot survive more than a week or so after birth.
   2. Conversely, in 2005, mice genetically engineered so that their mitochondria have elevated levels of the H_2O_2-destroying enzyme, catalase, live 20% longer, on average, than untreated controls.
      a. This was the first demonstration that enhanced antioxidant defenses can increase mammalian life span.

V. Despite the unquestioned destructive potential of free radicals (superoxide and hydroxyl radicals), SOD's importance as a factor in aging is controversial.
   A. In some cases, perturbations that increase oxygen radicals were found to raise lifespan rather than decrease it.
   B. Moreover, the whole idea that aging involves accumulation of random damage has been challenged by the discovery of specific genes in model organisms, like yeast and nematode worms.
      1. These genes, when mutated, allow the organism to live much longer.
      2. The existence of such genes has led to a competing theory known as "programmed aging."
      3. This theory posits that organisms have evolved mechanisms to induce their own decline after they have passed reproductive age.
   C. The degree to which free radical damage determines aging is thus very much an open question.

VI. A related area of study concerns the study of substances called antioxidants that are able to destroy free radicals in the test tube.
   A. The sale of these substances provides a major source of revenue for the vitamin/supplements industry.
   B. Common antioxidants found in the body are glutathione, vitamins E and C, and beta-carotene (the orange pigment in carrots and other vegetables).
      1. These substances may prove beneficial in the diet because of their ability to destroy free radicals.
      2. But studies on rats and mice have not provided convincing evidence that they retard the aging process or increase maximum life span.

(2.2) Engineering Linkage: Radionucleotides for Imaging and Treatment

I. An atom’s identity and chemical properties are based on the number of protons in its nucleus.
   A. Even though hydrogen has 1 proton and helium has 2 protons, each of these atoms can vary in the number of neutrons they contain in their nucleus to form isotopes.
      1. The general term *nuclide* refers to atoms that differ in the number of protons and neutrons, while isotopes have the same number of protons.
      2. Hydrogen has three isotopes, all with 1 proton and either 0, 1 or 2 neutrons.
      3. Tritium (^3H) an isotope of hydrogen with 2 neutrons, has an unstable combination of neutrons and protons and is radioactive.
   B. Unstable atoms can disintegrate to become more stable, and in the process releases particles or electromagnetic radiation that can be detected.

1. Radioactive isotopes are naturally occurring, but also can be made from nonradioactive elements.
2. Biological molecules can be produced with radioactive atoms and purchased for research or medicine.
3. Radioactive atoms are referred to as *radioisotopes* in cell biology, or *radionuclides* in biomedical engineering and medicine.

C. There are three main forms of radiation released when radioactive atoms disintegrate.
1. An *alpha particle* has 2 protons and 2 neutrons, equivalent to the nucleus of a helium atom.
2. A *beta particle* is equivalent to an electron.
3. *Gamma radiation* consists of electromagnetic radiation or photons and is similar to X-rays.
   a. Gamma radiation is generated by a radioactive isotope, while X-rays are generated by a machine.

II. Radionuclides/radioisotopes have many applications in medicine and in basic research; medical applications, fall into two categories: imaging and radiation therapy.

A. Medical imaging uses radiation produced outside the body with X-rays or MRI, but also internally by injecting radionuclides into the patient.
1. When these radionuclides, or tracers, undergo decay, the radiation they emit can be picked up by scanners outside the body, allowing the location of the tracer to be visualized.
2. Technetium-99 (^{99}Tc) is a breakdown product of molybdenum-99 (^{99}Mo) and a common tracer, which releases gamma radiation with energy similar to standard medical imaging X-rays, allowing the position of the ^{99}Tc to be imaged.
3. ^{99}Tc has a very short half-life which presents minimal danger to the patient, yet presents a shipping issue as it can decay completely before getting to the hosptial.
   a. Suppliers instead ship ^{99}Mo enclosed in a shielded unit known as a generator, where it is continually breaking down to produce ^{99}Tc.
   b. ^{99}Tc is separated from ^{99}Mo and then be injected directly for use in bone scans or chemically modified to allow it to pass to other parts of the body such as blood vessels in the brain.
4. Indium and gallium are used as tracers and can attached to an antibody that targets tumor cells, allowing the cancer to be visualized in the patient.

III. Radionuclides can be used as therapy by emitting radiation to kill tumor cells, or outside the body as a source of radiation aimed at a tumor.

A. Cobalt-60 (^{60}Co) is a strong gamma emitter used to target brain tumors in the gamma knife technique.
1. Patients are exposed to a large number of small gamma ray beams emitted by chunks of ^{60}Co arranged in a hemisphere around the patient's head.
2. Beams are aimed to converge on a single point in three dimensions, allowing precise targeting of a tumor deep inside the brain.
3. Each chunk of ^{60}Co emits a small quantity of gamma rays, so tissue damage is confined to the point of overlap where their intensities combine to produce a large effect.
   a. Suppliers instead ship ^{99}Mo enclosed in a shielded unit known as a generator, where it is continually breaking down to produce ^{99}Tc.
   b. ^{99}Tc is separated from ^{99}Mo and then be injected directly for use in bone scans or chemically modified to allow it to pass to other parts of the body such as blood vessels in the brain.

B. Internal radiation therapy relies on implantation of radionuclides in or near a tumor.
1. Small pellets with iodine-125 (^{125}I) can be implanted near the prostate as a treatment for prostate cancer.
2. Both ^{60}Co and ^{125}I work by creating gamma radiation, which then damages the DNA in cancer cells.

C. Most of the damage created by gamma rays is not caused by the radiation itself. When gamma rays interact with atoms, they cause fast-moving electrons to be expelled.
   1. These secondary electrons can break bonds in DNA on impact and can also ionize water to create hydroxyl radicals that are highly reactive.

D. Radiation therapy may sound frightening, but is important in the fight against cancer and other diseases.

(2.3) 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 (**noncovalent bonds**), while covalent bonds are stronger.
   A. Noncovalent bonds depend on attractive forces between atoms with opposite charges, not shared electrons.
      1. Individual noncovalent bonds are often weak (~1 - 5 kcal/mole) and thus they readily break and reform.
      2. This feature allows noncovalent bonds to mediate dynamic interactions among molecules in the cell.
   B. Even though they are individually weak, when many of them act in concert (2 DNA strands, different parts of a large protein), attractive forces add up and together provide structure with considerable stability.

II. Types of noncovalent bonds: ionic bonds (or *salt bridges*).
   A. Crystal of table salt is held together by an electrostatic attraction between positively charged Na^+ and negatively charged Cl^- ions; this type of attraction between fully charged components is an **ionic bond**.
      1. Ionic bonds within a salt crystal may be quite strong.
      2. However, if salt crystal is dissolved in water, each of individual ions is surrounded by water molecules, which inhibit oppositely charged ions from approaching one another closely enough to form ionic bonds.
      3. Because cells are made primarily of water, bonds between *free* ions are of little importance.
   C. In contrast, weak ionic bonds between oppositely charged groups of large biological molecules are of considerable importance; ex.: negatively charged phosphate atoms in DNA molecule.
      1. When negatively charged phosphate atoms in DNA molecule are closely associated with positively charged groups on a protein's surface, ionic bonds between them help hold the complex together.
      2. Ionic bond strength in a cell is generally weak (~3 kcal/mole) due to the presence of water.
      3. However, deep in a protein core where water is often excluded, such bonds can be much stronger.

III. Types of noncovalent bonds: hydrogen (H) bonds.
   A. If H is bonded to electronegative atom (especially O or N), the single pair of shared electrons is greatly displaced toward the electronegative atom nucleus so that the H atom is partially positive.
      1. Due to charge shift, the bare positively charged nucleus of H can approach close to unshared pair of outer electrons of second electronegative atom.
      2. This creates an attractive, weak electrostatic interaction called a **hydrogen bond**.
      3. H bonds between or within molecules in solution have an energy of roughly 1 kcal/mole, which is similar in magnitude to the thermal energy; thus, such H bonds are easily broken.
      4. Occur between most polar molecules; they are important in determining structure and properties of water, also form between polar groups present in large biological molecules (like DNA double helix).
   B. Strong collectively because their strength is additive, e.g., DNA double helix is very stable.

1. However, because individual H bonds are weak, the 2 DNA strands can be partially separated to allow enzymes access to DNA's individual strands.

IV. Types of noncovalent bonds: hydrophobic (water-fearing) interactions.
   A. Polar molecules like amino acids and sugars are called **hydrophilic** (water-loving) because they can interact with water; nonpolar molecules (fat molecules or steroids; water-fearing) are essentially insoluble in water.
      1. Molecules with nonpolar covalent bonds lack charged regions that can attract them to poles of water molecules and are thus insoluble in water.
      2. When mixed with water, hydrophobic molecules are forced into aggregates, minimizing their exposure to the polar surroundings (fat on surface of chicken or beef soup even after stirring).
         a. This type of interaction or association is called **hydrophobic interaction**.
   B. Hydrophobic interactions are not classified as true bonds since they do not result from an attraction between hydrophobic molecules.
      1. Rather they result from an energetic drive to exclude water away from hydrophobic surfaces.
      2. This reflects the accepted hypothesis that hydrophobic interactions are driven by increased entropy (disorder).
         a. When a hydrophobic group projects into an aqueous solvent, the water molecules become ordered in a cage around the hydrophobic group.
         b. These solvent molecules become disordered when the hydrophobic group withdraws from the surrounding solvent.

V. Types of noncovalent bonds: van der Waals forces.
   A. Hydrophobic groups can form weak bonds with one another based on electrostatic interactions; due to slight perturbations of electron distributions.
      1. Polar molecules associate because they contain permanent asymmetric charge distributions within their structure.
      2. Closer examination shows that electron distributions of nonpolar covalent bonds of nonpolar molecules (like those in CH_4 or H_2) are not always symmetric.
         a. The electron distribution around an atom at any given instant is a statistical matter and, therefore, varies from one instant to the next.
      3. Thus, at any given time, electron density may be larger on one side of atom even though the electrons are shared equally with another atom.
         a. These transient charge asymmetries in electron distribution result in momentary charge separations (*dipoles*).
   B. If 2 such molecules with transitory dipoles are very close together and appropriately oriented, they experience a weak attractive force bonding them together (**van der Waals force**).
      1. Moreover, the formation of a temporary charge separation in one molecule can induce a similar separation in an adjacent molecule, generating additional attractive forces between nonpolar molecules.
      2. Single van der Waals force is very weak (0.1-0.3 kcal/mole) and very sensitive to the distance separating 2 atoms.
      3. Molecules must be close together and interacting portions usually have complementary shapes that allow close approach; many atoms of both interactants can approach each other closely.

4. Van der Waals interactions are important in biological interactions (ex.: between antibodies and viral antigens).

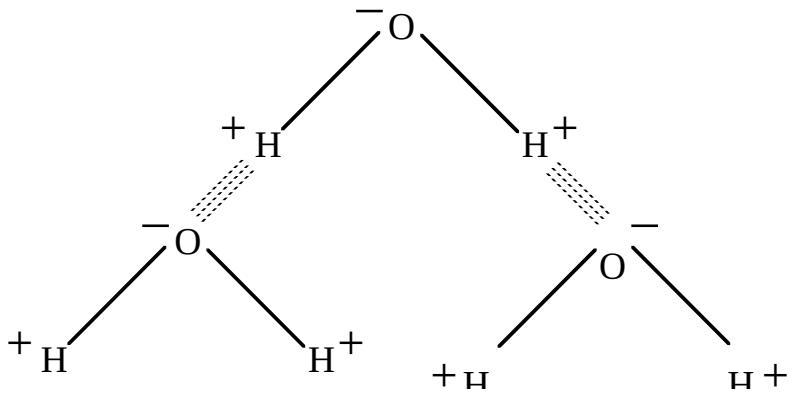
VI. Life is totally dependent on water as the unique water structure (although only 3 atoms) is responsible for its extraordinary life-supporting properties and attributes.
   A. Water is a highly asymmetric with an O atom at one end and 2 H atoms at the opposite end.
   B. Each of the 2 covalent bonds in the molecule is highly polarized.
   C. All 3 atoms in a water molecule are very adept at forming H bonds.

VII. What unique abilities do the properties of water confer upon it?
   A. Each H_2O molecule can H bond with up to 4 water molecules and can form a highly interconnected molecular network.
      1. Each H bond is formed when partially negative-charged O at one end of molecule aligns with partially positive-charged H of another one.
      2. H_2O molecules have an unusually strong tendency to adhere to each other due to their extensive H bonds.
      3. This feature is most evident in the thermal properties of water.
   B. Comparison of water structure with that of H_2S (hydrogen sulfide) helps one appreciate water's structure.
      1. Like oxygen, sulfur has 6 outer-shell electrons and forms single bonds with 2 hydrogen atoms.
      2. But because sulfur is a larger atom, it is less electronegative than oxygen, and its ability to form H bonds is greatly reduced.
      3. At room temperature, H_2S is a gas, not a liquid; in fact, temperature must drop to –86°C before it freezes into a solid.
   C. The plentiful H bonds of water lead to its properties that relate to its importance to life.

VIII. Tendency of water molecules to adhere to each other is evident in water's thermal properties.
   A. When water is heated, most of the thermal energy is consumed in disrupting H bonds rather than contributing to molecular motion (which is measured as an increased temperature).
      1. Thus, temperature does not rise too fast.
   B. Similarly, evaporation from the liquid to the gaseous state requires that water molecules break the H bonds holding them to their neighbors; this explains why it takes so much energy to convert water to steam.
      1. Mammals take advantage of this property when they sweat, because the heat required to evaporate the water in sweat is absorbed from the body, which thus becomes cooler.

IX. The small volume of aqueous fluid in a cell contains a remarkably complex mixture of dissolved substances or *solutes*; in fact, water is able to dissolve more types of substances than any other solvent.
   A. Water is more than a solvent as it determines the structure of biological molecules, the types of interactions in which they can engage and the complexes they form (like membranes).
      1. Water is the fluid matrix around which the insoluble fabric of the cell is constructed.
      2. It is also the medium through which materials move from one compartment of the cell to another.
      3. It is a reactant or product in many cellular reactions.
      4. It also protects cell in many ways: from excessive heat, cold or damaging radiation.
   B. Water is such an important factor in a cell because it can form weak interactions with so many types of chemical groups; it solubilizes ions and organic molecules.

1. With its strongly polarized O—H bonds, water forms a shell around ions separating them from one another.
2. H bonds with organic molecules containing polar groups (e. g., amino acids, sugars) and larger macromolecules; this ensures their solubility within the cell.
   a. Water stabilizes the structure of macromolecules, promoting their function.
   b. Water molecules are H bonded to each other and to specific amino acids of a protein.



(2.4) Acids, Bases and Buffers

I. Protons are not only found within atomic nuclei, but they are also released into medium whenever a hydrogen atom loses a shared electron (ex.: acetic acid).
   A. **Acid** - a molecule that is capable of releasing donating a H^+ ion to the medium through *dissociation*; a proton dissociates and is released into the medium whenever a hydrogen atom loses an electron.
   B. Acetic acid from vinegar can undergo dissociation into a proton (H^+ ion) and a negatively-charged acetate ion.
      1. The proton released by acetic acid does not remain in a free state; instead, once dissociated, it combines with another molecule, forming H_3O^+, H_2O, NH_3^+, etc.
      2. Possible reactions involving a proton include:
         a. Combination with a water molecule to form an hydronium ion (H_3O^+).
         b. Combination with a hydroxyl ion (OH^-) to form a molecule of water.
         c. Combination with an amino group ($-NH_2$) in a protein to form a charged amine ($-NH_3^+$).
   C. **Base** - any molecule capable of accepting a hydrogen ion (proton); acids and bases exist in pairs (*couples*).
      1. When an acid loses a proton (as when acetic acid gives up a H^+ ion), it becomes a base (an acetate ion), which is the *conjugate base* of the acid.
      2. Similarly, when a base (like an $-NH_2$ group) accepts a proton, it forms an acid (in this case, $-NH_3^+$), it is termed the *conjugate acid* of that base.
      3. Thus, the acid always contains one more positive charge than its conjugate base.
      4. Water is an example of an *amphoteric* molecule, i.e., one that can serve both as an acid and a base; usually has both a positive and a negative charge; amino acids, in addition to water, are examples.

II. Acids vary greatly in the ease with which they give up a proton.
   A. The more readily a proton lost, the less strong is the attraction of conjugate base for its proton and the stronger is the acid (ex.: HCl); it readily transfers its proton to water.

1. Strong acid's conjugate base (ex.: Cl) is weak base; H^+ dissociates, since H_2O is a stronger base.

B. In contrast, a relatively weak acid (ex.: acetic acid) is mostly undissociated when dissolved in H_2O; the acetate ion is a stronger base than H_2O.

C. One can consider the degree of dissociation of an acid in terms of the competition for protons among the components of a solution.
   1. Water is a better competitor (a stronger base) than a chloride ion, so HCl completely dissociates.
   2. In contrast, the acetate ion is a stronger base than water, so it remains largely as undissociated acetic acid.

III. The acidity of a solution is measured by the concentration of H^+ ions and is expressed in terms of **pH** (a measure of H^+ concentration) = - $\log_{10}[H^+]$, where $[H^+]$ is the molar concentration of protons.

A. A solution with a pH of 5 contains a H^+ ion concentration of 10^{-5} M.

B. Logarithmic scale - increase of 1 pH unit means 10-fold increase in $[OH^-]$ and a 10-fold decrease in $[H^+]$.
   1. Stomach juice (pH 1.8) has nearly one million times the $[H^+]$ of blood.

C. Formula for dissociation of water into a hydroxyl ion and a proton: $H_2O \leftrightarrow H^+ + OH^-$ or more accurately 2 $H_2O \leftrightarrow H_3O^+ + OH^-$

D. In pure water, $[H^+] = [OH^-] = \sim 10^{-7}$ M; low dissociation indicates water is very weak acid.
   1. Pure water thus has a pH of 7; since water is viewed as being the standard solvent into which other molecules are dissolved, it is common to refer to pH 7 as "neutral."

E. In aqueous solutions, protons do not exist in the free state, but rather as H_3O^+ or $H_5O_2^+$ ions, but for simplicity one can refer to them as hydrogen ions or protons.

IV. Most biological processes are acutely sensitive to pH, since $[H^+]$ changes affect biological molecule ionic states.

A. Amino acid R groups can acquire charge (-COOH → $-COO^-$; $-NH_2$ → $-NH_3^+$).
   1. As $[H^+]$ increases, the $-NH_2$ group of the amino acid arginine becomes protonated to form $-NH_3^+$, which can disrupt the activity of the entire protein.

B. Even slight pH changes altering these groups can disrupt shape and activity of entire protein and impede biological reactions.

C. Organisms, and the cells of which they are comprised, are protected from pH fluctuations by **buffers**.

V. Buffers are compounds that react with free H^+ or OH^- ions, thereby resisting changes in pH; minimizes pH fluctuations.

A. Buffer solutions usually contain a weak acid together with its conjugate base.

B. Blood is buffered by carbonic acid (H_2CO_3) and bicarbonate (HCO_3^-) ions.
   1. If $[H^+]$ rises (as occurs during exercise), HCO_3^- ions combine with excess protons derived from carbonic acid, removing them from solution.
   2. Conversely, excess OH^- ions (which rise during hyperventilation) are neutralized by protons derived from carbonic acid; blood normally held at pH 7.4.
   3. The pH of the fluid within the cell is regulated in a similar manner by a phosphate buffer system consisting of $H_2PO_4^-$ and HPO_4^{-2}.

(2.5) The Nature of Biological Molecules

I. Most of an organism is water; if you evaporate water away, most of remaining dry weight consists of molecules containing carbon atoms.
   A. When first discovered, carbon-containing molecules were thought to be present only in living organisms and referred to as *organic molecules* to distinguish them from *inorganic molecules* found in inanimate world.
   B. Chemists learned to make more and more of them, so some of mystique about organic molecules was dispelled.
   C. Current use of the term organic refers to a molecule containing a carbon-carbon bond, regardless of the source of the molecule.
   D. Compounds made by living organisms are now called **biochemicals**.

II. Organic chemistry centers around carbon atom – both its size and electronic structure allow carbon to form an incredible number of molecules (several 100,000 of which are known).
   A. Carbon binds to up to 4 other atoms, since it has only 4 outer-shell electrons (8 needed to fill shell).
   B. Most importantly, each carbon atom is able to bond with other carbon atoms to construct carbon-containing backbones with long chains, which may be linear, branched or cyclic.
      1. In contrast, silicon (4 outer-shell electrons) is too large for its +-charged nucleus to attract neighboring atom outer-shell electrons enough to hold such large molecules together.
         a. Silicon is just below carbon in the periodic table.

III. *Hydrocarbons* contain only hydrogen and carbon atoms and are the simplest group of organic molecules.
   A. Ethane (C_2H_6) is a simple hydrocarbon consisting of 2 atoms of carbon in which each carbon is bonded to the other carbon as well as 3 hydrogen atoms.
   B. As more carbons added, organic molecule skeletons increase in length and structure gets more complex.

IV. Functional groups are particular atom groupings that often behave as a unit to give organic molecules their physical properties, chemical reactivity and solubility in aqueous solutions.
   A. Hydrocarbons do not occur in significant amounts in living cells, although they form the bulk of fossil fuels formed from the remains of ancient plants and animals.
      1. Many of the organic molecules important in biology contain chains of carbons like those in hydrocarbons, but certain hydrogens are replaced by various **functional groups**.
      2. Two of the most common linkages between functional groups are **ester bonds** (form between carboxylic acids and alcohols) and **amide bonds** (form between carboxylic acids and amines).
   B. Some major functional groups.
      1. Hydroxyl group -OH
      2. Carboxyl group -COOH; acquires charge $-COO^-$
      3. Sulfhydryl group -SH; react to form disulfide bonds in polypeptides.
      4. Amino group $-NH_2$; acquires charge $-NH_3^+$
      5. Phosphate group $-PO_3H_2^{2-}$
      6. Carbonyl group -C=O
      7. Methyl group $-CH_3$
   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) and thus make organic molecules more polar, more water soluble and more reactive.
2. Several of them are capable of ionization and may become positively or negatively charged.

D. Example of functional group importance (ethane → ethanol → acetic acid → ethyl mercaptan); effect on molecules by the substitution of various functional groups is readily demonstrated.
1. Ethane (CH_3CH_3) is a toxic, flammable gas; if replace one H with hydroxyl (-OH) to get ethanol.
2. Ethyl alcohol (ethanol; CH_3CH_2OH) which is palatable; if replace -CH_2OH with -COOH to get acetic acid.
3. Acetic acid (CH_3COOH), strong-tasting vinegar ingredient; if replace -COOH with -CH_2SH to get ethyl mercaptan.
4. Ethyl mercaptan (CH_3CH_2SH) is a strong, foul-smelling agent used to study enzyme reactions.

V. Organic molecules commonly found within living cells can be divided into several categories based on their role in metabolism.

A. **Macromolecules** form the structure and carry out the activities of cells; huge and highly organized molecules; contain from dozens to millions of carbon atoms.
1. Because of their size and the intricate shapes they can assume, some can perform complex tasks with great precision and efficiency.
2. Presence of macromolecules, more than any other trait, endows organisms with the properties of life and sets them apart chemically from the inanimate world.
3. Can be divided into 4 major categories: proteins, nucleic acids, polysaccharides, certain lipids - first 3 types are *polymers*; made of large number of low MW building blocks or *monomers*.
   a. These macromolecules are constructed from monomers by a process called *polymerization* that resembles coupling railroad cars onto a train.
4. Basic structure and function of each type of macromolecule are similar in all organisms.
   a. If one looks at the specific sequences of monomers making up these various individual macromolecules, the diversity among organisms becomes apparent.

B. Macromolecule building blocks – most macromolecules in a cell have a short lifetime compared with cell itself (with exception of DNA); they are continually broken down and replaced by new macromolecules.
1. Most cells contain a pool of low-MW precursors that are ready to be incorporated into macromolecules.
   a. These include sugars, the precursors of polysaccharides; amino acids, the precursors of proteins; nucleotides, the precursors of nucleic acids and fatty acids, incorporated into lipids.

C. Metabolic intermediates (metabolites) – molecules in cell have complex chemical structures and must be synthesized in a step-by-step sequence beginning with specific starting materials.
1. In the cell, each series of chemical reactions is called a **metabolic pathway**.
   a. Pathway starts with a compound and converts it to other ones sequentially until some functional end product that can be used in other reactions (like an amino acid building block of protein) is made.
   b. A → B → C → D → E → ……..→ P
2. 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 and are called **metabolic intermediates**.

D. Molecules of miscellaneous function – the vast bulk of cell dry weight is made up of macromolecules and their direct precursors, which is a broad category of molecules not as large as one might think.
1. Vitamins function primarily as adjuncts to proteins.
2. Certain steroid or amino acid hormones.

3. Molecules involved in energy storage (ATP, creatine phosphate).
4. Regulatory molecules like cyclic AMP.
5. Metabolic waste products like urea.

(2.6) Green Cells: Chemical Fertilizers

I. Plants and algae are important to the ecosystem since they can use solar energy to drive photosynthesis.
   A. Photosynthesis allows plants to make sugar molecules for their energy needs without having to consume other organisms.
   B. For growth, plants also need chemical elements to make biochemicals such as carbohydrates, proteins, and nucleic acids.
      1. Water can provide a source of oxygen and water, but cells also need carbon, nitrogen, and phosphorus to make the biochemicals needed for growth.
      2. Plants use carbon fixation to harvest carbon from the CO_2 in the air.

II. Since plant cells cannot fix nitrogen from the air, plants like the legumes have a symbiotic relationship with rhizobia, bacteria that fix nitrogen similar to the cyanobacteria.
   A. Rhizobia grow on the roots of legumes in structures called root nodules.
   B. When root nodules are forming, rhizobia invade and colonize the plant cells in specialized membrane compartments called bacteroids, essentially becoming a type of organelle.
   C. It is here where nitrogen is converted into ammonia, which plants can use to synthesize amino acids and other biochemicals.
   D. Amino acids left behind in the dead plant parts after harvesting end up in the soil and are degraded to nitrate, used by other plants as a nitrogen source.

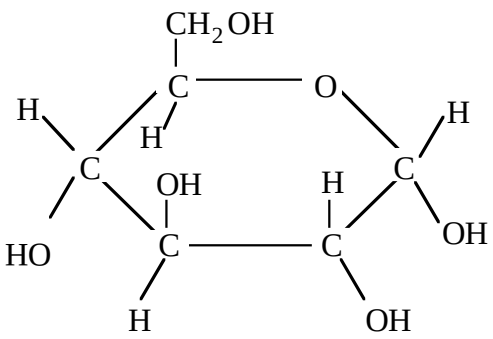
III. This residual nitrogen source is the basis of "crop rotation," a farming practice in which legumes are grown for one season and then some other plant that cannot fix nitrogen is grown in the next season.
   A. Without crop rotation, non-leguminous plants will use up the nitrogen making it harder for them to grow.
   B. Farmers and gardeners can supplement the soil with nitrogen-containing fertilizers prior to planting.
      1. Ammonia-derived chemicals can be used from animal products like chicken manure and chemically synthesized fertilizers that contain ammonium nitrate (NH_4NO_3).
      2. The ammonia in nitrogen fertilizers is made in factories via reactions of natural gas with atmospheric nitrogen.

IV. Phosphorus is included in fertilizers as a chemical source for the production of ATP, DNA, and RNA.
   A. Phosphorus is not found in air but exists in the ground in rocks and minerals such as hydroxyapatite.
      1. Phosphorus leaches out of the rock and into groundwater where it is available for absorption by plants. However, it is also carried away in runoff.
   B. Phosphorus can become depleted since it is consumed by plants and not replaced by fixation.
   C. Fertilizers supplement phosphorous levels using phosphorus-rich animal products such as bone meal or minerals that have been acid-treated to release the phosphate in water-soluble form.
   D. There are global economic implications since phosphorus-rich minerals are abundant in countries such as China, the United States, and Morocco, while other countries need to import these minerals.

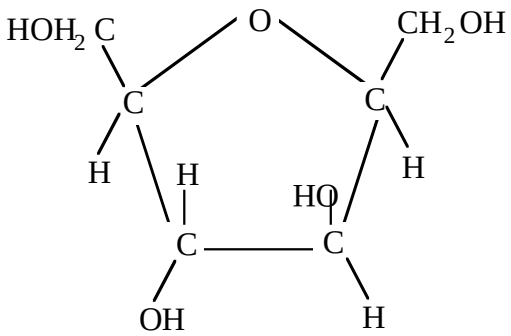
(2.7) Four Types of Biological Molecules

I. **Carbohydrates** (often called **glycans**) comprise a group of substances, including simple sugars (*monosaccharides*) and all larger molecules constructed from sugar building blocks.
   A. Function primarily as chemical energy storehouse and durable building materials for biological construction.
   B. Most have general formula $(CH_2O)_n$.
      1. Important ones in cell metabolism have from 3 to 7 carbons (n = 3 - 7).
      2. *Trioses*, *tetroses*, *pentoses*, *hexoses*, and *heptoses* - 3, 4, 5, 6, and 7 carbons, respectively.

II. In the structure of simple sugars, each sugar molecule consists of carbon atom backbone linked together in a linear array by single bonds.
   A. Each carbon of backbone is linked to single OH group except for one bearing a *carbonyl* (C=O) group.
      1. *Ketose* - carbonyl group located at an internal chain position; forms ketone group (e.g., fructose).
      2. *Aldose* - carbonyl group is located at one end of sugar; forms an aldehyde group (e.g., glucose).
   B. Because of their large numbers of hydroxyl groups, sugars tend to be highly water soluble.
   C. Sugars with ≥5 carbons spontaneously convert by self-reaction into a closed, ring-containing molecule with Hs and OHs above or below ring; ring not planar, but in 3D-conformation resembling a chair.
      1. Only a tiny fraction of sugar molecules in solution are found in the open-chain linear form; the rest are in the ring form.
   D. The linear form is biochemically important because the aldehyde group at the end of the chain is reactive and can react with proteins, notably hemoglobin.
      1. Patients with diabetes have higher levels of sugar in their blood, and this sugar in its open chain form reacts with hemoglobin to produce a modified hemoglobin called Hemoglobin A1c.
      2. Hemoglobin A1c is often used in blood tests to track the progress of diabetes.
   E. Similar reactions of linear-form sugar with proteins involved in cholesterol metabolism are one of the reasons that diabetes causes heart disease.
      1. The open chain form of sugars is thus quite important in medicine.
      2. But the vast majority of sugars is found in ring form; it is in this form that they are used as building blocks to build other types of carbohydrates.
   F. The ring forms of sugars are usually depicted as flat (*planar*) structures lying perpendicular to the plane of the paper with the thickened line situated closest to the reader.
      1. The H and OH groups lie parallel to the plane of the paper, projecting either above or below the ring of the sugar.
      2. In actual fact, the sugar ring is not a planar structure, but usually exists in a three-dimensional conformation resembling a chair.



GLUCOSE



FRUCTOSE

III. Stereoisomerism - arrangement of groups around a carbon atom is depicted with the carbon in the center of a tetrahedron and the bonded groups projecting into its 4 corners; carbon can bond with 4 other atoms.

A. Glyceraldehyde (the only aldotriose) – the second carbon atom of glyceraldehyde is linked to 4 different groups (-H, -OH, -CHO, and -CH_2OH).

1. If the 4 groups bonded to a carbon atom are all different, as in glyceraldehyde, then 2 possible configurations exist that cannot be superimposed on one another.
2. These 2 molecules, termed *stereoisomers* or *enantiomers*, have essentially the same chemical reactivities, but their structures are mirror images (not unlike a pair of right and left human hands).
3. By convention, the molecule is called D-glyceraldehyde if the OH group of carbon 2 projects to the right and L-glyceraldehyde if it projects to the left.
4. Because it acts as a site of stereoisomerism, carbon 2 is referred to as an *asymmetric* carbon atom.

B. As the backbone of sugar molecules increases in length, so too does the number of asymmetric carbon atoms and, consequently, the number of stereoisomers.

1. Aldotetroses have 2 asymmetric carbons and thus can exist in 4 different configurations.
2. Similarly, there are 8 different aldopentoses and 16 different aldohexoses.
3. D or L designation of these sugars is based by convention on the arrangement of groups attached to the asymmetric carbon farthest from aldehyde (the carbon associated with the aldehyde is designated as C1).
   a. If the OH group on the carbon farthest from the aldehyde projects to right, the aldose is a D-sugar
   b. If it projects to the left, it is an L-sugar.
4. Enzymes present in living cells can distinguish between D- and L-forms; usually, cells use only one of the stereoisomers (such as D-glucose and L-fucose).

C. Straight-chain glucose converts by self-reaction into 6-membered *pyranose* ring with carbon 1 being asymmetric.

1. Unlike its open chain precursor, C1 of the ring form bears 4 different groups and thus becomes a new center of asymmetry within sugar molecule.
2. Because of this extra asymmetric carbon atom, each type of pyranose exists as α and β stereoisomers.
3. By convention, if the OH group of carbon 1 is below the plane of the ring, it is an α-pyranose; if the OH group is above the plane of the ring, it is a β–pyranose.
4. The difference between the 2 forms has important consequences; results in the compact shape of glycogen and starch (α) molecules and the extended conformation of cellulose (β).

IV. Linking sugars together to make larger molecules – covalent bond joining sugars together is called **glycosidic bond** (–C—O—C–); forms by reaction between C1 of one sugar and OH of another sugar.

A. Sugars can be joined by quite 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) – present in the milk of most mammals; fuel for early growth and development of newborn mammals.
   a. The enzyme lactase that hydrolyzes lactose is found in plasma membranes of cells lining intestines.
   b. Many people lose this enzyme after childhood if they eat dairy products, it causes digestive discomfort.

C. **Oligosaccharides** are linked small chains of sugars (*oligo* - few) usually covalently attached to lipids and proteins, converting them to glycolipids and glycoproteins, respectively.

1. They are particularly important on the glycoproteins and glycolipids of the plasma membrane at the cell surface from which they project.
2. They may be composed of many different sugar unit combinations and thus play an informational role.
3. They can serve to distinguish one cell type from another and help mediate specific interactions of a cell with its surroundings.

V. Claude Bernard and diabetes – by mid-19th century, it was known that the blood of diabetics had sweet taste due to elevated levels of glucose (key sugar in energy metabolism).

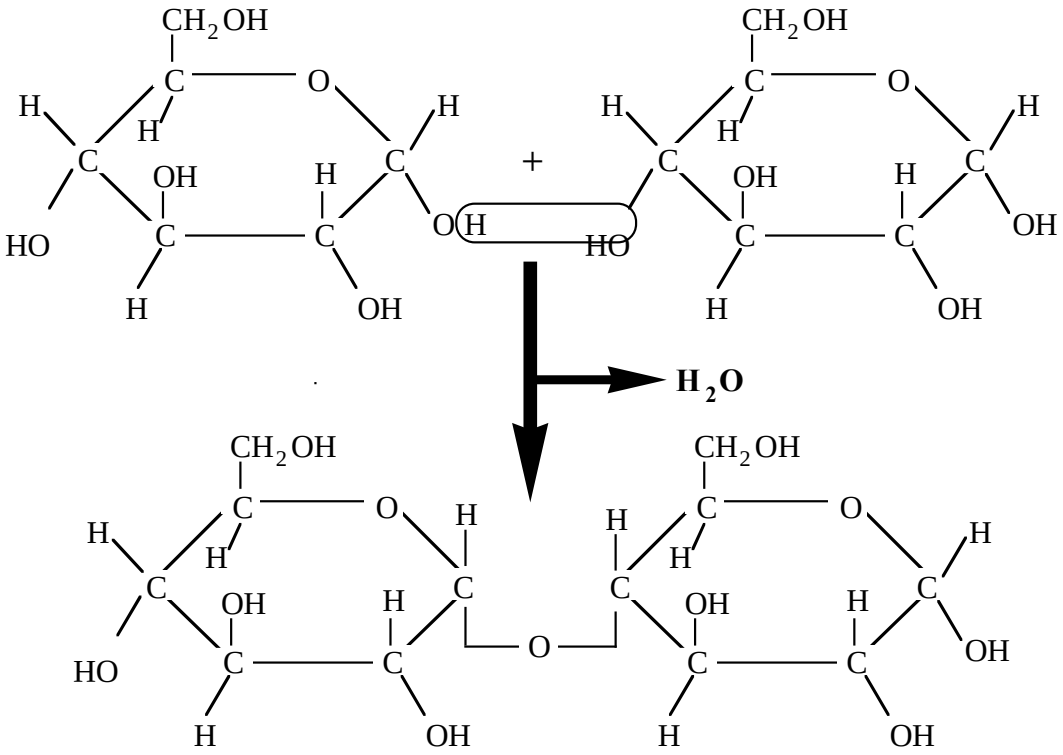
A. Claude Bernard, a prominent French physiologist of the time, tried to find diabetes cause by investigating source of blood sugar (at first, thought any sugar present in human or animal had to come from diet).

B. Worked with dogs and found that even if the animals were placed on a diet totally lacking carbohydrates, their blood still contained a normal amount of glucose.

1. It was thus clear that glucose could be formed in the body from other types of compounds.
2. Eventually found liver releases glucose to blood by hydrolyzing **glycogen** (insoluble glucose polymer).
3. Concluded that various food materials like proteins were carried to the liver where they were chemically converted to glucose and stored as glycogen, released from liver if needed.
4. Then, as the body needed sugar for fuel, the glycogen in the liver was transformed to glucose, which was released into the bloodstream to satisfy glucose-depleted tissues.
5. Felt that the balance between glycogen formation and glycogen breakdown in liver is the prime determinant in maintaining the relatively constant (*homeostatic*) level of glucose in the blood.
6. Bernard was, of course, correct.

D. The molecule he named glycogen is a type of **polysaccharide**, a polymer of sugar units joined by glycosidic bonds.

Condensation of Glucose Monosaccharides



VI. Glycogen and starch are nutritional polysaccharides.

A. Glycogen is a branched glucose polymer mostly joined by α(1→ 4) bonds (only one kind of monomer).

1. Sugar at branch point is joined to 3 neighboring units instead of 2, as in unbranched segments of polymer; bond at branch is α(1→ 6) bond linkage.
2. Glycogen is surplus chemical energy storehouse in most animals.
3. Human skeletal muscles have enough glycogen to fuel about 30 min of moderate activity.
4. Glycogen typically ranges in MW from ~1 - 4 million daltons.
5. When stored in cells, it is highly concentrated in what appears as dark-staining, irregular granules in EM.

B. **Starch** is a glucose polymer used by most plants to bank their surplus chemical energy (potatoes and cereals are primarily starch).

1. Starch is actually a mixture of 2 different polymers, amylose and amylopectin.
   a. Amylose is an unbranched, helical molecule, whose sugars are joined by α(1→ 4) linkages.
   b. Amylopectin is branched (it differs from glycogen in being much less branched and has an irregular branching pattern); α(1→ 6) bonds at branch.
2. Starch is stored as densely packed granules (*starch grains*), which are enclosed in membrane-bound organelles (*plastids*) within the plant cell.
3. Animals possess enzyme (*amylase*) to hydrolyze starch, even though they don't synthesize it.

VII. Cellulose, chitin and glycosaminoglycans are structural polysaccharides that form tough, durable structural materials instead of easily digested energy stores.

A. Cotton and linen are largely comprised of **cellulose**, which is the major plant cell wall component.
   1. Cotton textiles owe their durability to the long, unbranched cellulose molecules of which they are made.
      a. They are ordered into side-by-side aggregates to form molecular cables that are ideally constructed to resist pulling (tensile) forces.
   2. Like glycogen and starch, cellulose contains solely glucose monomers.
      a. Its properties differ dramatically from the above polysaccharides because of the difference in bonds joining the glucose units, which are joined by β(1→ 4) bonds rather than α(1→ 4) bonds.
      b. Ironically, multicellular animals (with rare exception) lack enzyme needed to degrade cellulose, even though it is the most abundant organic material on Earth and rich in chemical energy.
   3. Amimals that make a living by digesting cellulose (termites, sheep) do so by harboring bacteria and protozoa that make the needed enzyme cellulase.
   4. Cellulose is an important component of *dietary fiber*, a broad term that includes all the polysaccharides we eat that cannot be digested by human enzymes.

B. *Chitin* is an unbranched polymer of the sugar N-acetylglucosamine (similar in structure to glucose, but with an acetyl amino group, instead of an OH, bonded to second carbon atom of glucose ring).
   1. Occurs widely as a structural material among invertebrates, particularly the outer covering of insects, spiders, crustaceans.
   2. It is a tough, resilient, yet flexible material, not unlike certain plastics; insects owe much of their success to this highly adaptive polysaccharide covering.

C. **Glycosaminoglycans** (**GAGs**) are group of polysaccharides with a more complex structure.
   1. Unlike other polysaccharides they have the structure of repeating disaccharides (2 different sugars in a repeating pattern; —A—B—A—B—); most occur in spaces surrounding cells.
   2. The best studied GAG is heparin, which is secreted by cells in the lungs and other tissues in response to tissue injury.
      a. Heparin inhibits blood coagulation thereby preventing the formation of clots that can block the flow of blood to the heart or lungs.
      b. Heparin does this by activating an inhibitor (antithrombin) of a key enzyme (thrombin) that is required for blood coagulation.
      c. Heparin is usually extracted from pig tissue and has been use for decades to prevent blood clots in patients following major surgery.
   3. Unlike heparin, most GAGs are found in the spaces that surround cells.

D. The most complex polysaccharides are found in plant cell walls (discussed in Ch. 7).

Polysaccharide	Location	Component Sugar	Type of Bonds	Function
Glycogen	Animal tissues	D-glucose	α(1→4); highly branched every 8-10 residues via α(1→6) linkages	Principal animal energy storage product (especially in liver and muscle)
Starch				
α-Amylose	Plants	D - glucose	α(1 → 4); unbranched, forms helical coil	Principal higher plant energy storage product with amylopectin

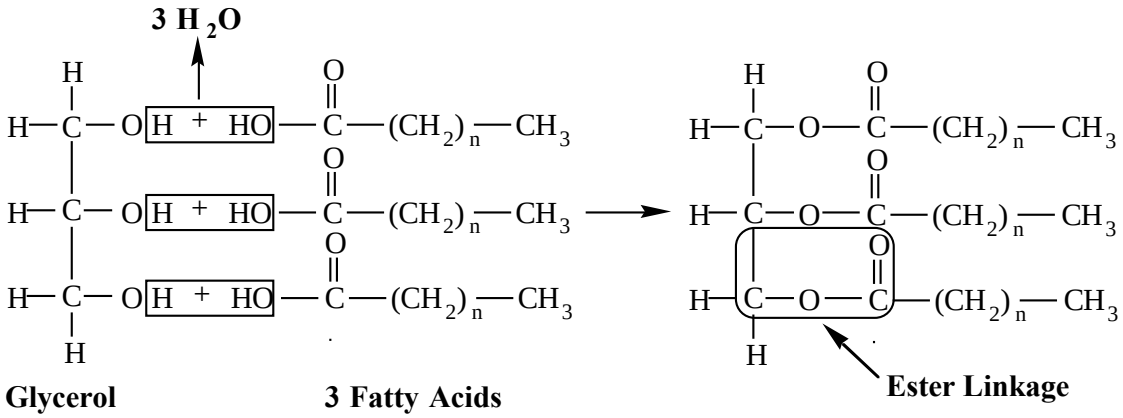
Amylopectin	Plants	D - glucose	α(1→4); branched every ~12-25 residues along backbone via α(1→6) bonds. Branches ~12 residues long	Principal higher plant energy storage product with α–amylose
Cellulose	Some lower invertebrates and plants; usually extracellular (ex. pure cotton and linen)	D-glucose (disaccharide - cellobiose)	β (1→4); no branching	Mainly structural; nutrient if can break it down (mammals who use as food lack enzyme to digest cellulose but get it from bacteria and protozoa in rumen); highly insoluble
Peptidoglycans	Bacterial cell wall	NAG-NAM (N-acetyl glucosamine and N-acetyl-muramic acid)	β(1→4); no branching; cross-linked by peptide bonds between attached aminos	Structural, possibly some minor barrier functions
Chitin	Exoskeletons of insects and crustaceans	NAG (N-acetyl-D-glucosamine)	β(1→4); no branching; close relative of cellulose	Structural component of exoskeleton (tough, resilient, flexible)
Glycosaminoglycans (GAGs)	Extracellular material (in spaces around cells) and connective tissue	Two alternating sugars (usually one is an amino sugar)	β(1→4) and β(1→4) or β(1→3); no branching	Structural; extremely important in development

VIII. Lipids are a diverse group of nonpolar biological molecules whose common properties are their solubility in organic solvents (benzene, chloroform) and insolubility in H_2O (explains many of their varied biological functions).

A. Lipids of importance in cell function include fats, oils, steroids and phospholipids; composed principally of C, H and O - not macromolecules, but aggregate to form large complexes.

B. **Fats** consist of a glycerol molecule linked by ester bonds to 3 fatty acids; the composite molecule is called triglyceride or **triacylglycerol**.

1. Formed by 3 condensation reactions, which form ester linkages (—C—O—C—) between glycerol (a polar molecule) and 3 fatty acids.



IX. **Fatty acids** are long, unbranched hydrocarbon chains with a single carboxyl group at one end; the 2 ends of the molecule have very different structures and also very different properties.

A. They are both hydrophobic (long hydrocarbon chain) and hydrophilic (carboxyl [—COOH]; negative charge at physiological pH) in character.
   1. A molecule that has both hydrophobic and hydrophilic regions is said to be **amphipathic**; such molecules have unusual and biologically important properties.
   2. Soap properties are good examples since it consists of fatty acids.
      a. In past centuries, soap was made by heating animal fat in strong alkali (NaOH, KOH) to break bonds between fatty acid and glycerol; most are now made synthetically.
      b. They owe their grease-dissolving capability to the fact that the hydrophobic end of each fatty acid can embed itself in grease, while the hydrophilic end can interact with the surrounding water.
      c. As a result, greasy materials are converted into complexes that can be dispersed by water (*micelles*).

B. Fatty acids can differ from each other in length of their hydrocarbon chain (usually even and vary from 14 to 20 carbons) and degree of saturation (presence or absence of double bonds).
   1. **Saturated** chains lack double bonds as every C is attached to maximum number of Hs; chains straight, pack tightly together and are solid above room temperature (e.g. stearic acid, animal fats).
   2. **Unsaturated** chains have 1 or a few double bonds that causes bend/kink and if in *cis* configuration prevents tight packing; if prevalent, liquid at room temperature (e.g. plant fats called oils).
      a. Multiple double bonds leads to oils being called “polyunsaturated.”

C. Naturally occurring fatty acids have double bonds in the *cis* configuration, which produces kinks in a fatty acid chain.
   1. Thus, the more double bonds that fatty acid chains possess, the less effectively these long chains can be packed together; this lowers the temperature at which a fatty-acid-containing lipid melts.
   2. Tristearate, whose fatty acids lack double bonds, is a common component of animal fats and remains in a solid state well above room temperature.
   3. In contrast, the profusion of double bonds in vegetable fats accounts for their liquid state (both in plant cell and grocery store) and for them being labeled as polyunsaturated.
      a. Fats that are liquid at room temperature are described as **oils**.
      b. Linseed oil is a highly volatile lipid extracted from flax seeds, remains a liquid at much lower temperatures than does tristearate.

4. Solid shortenings, like margarine, are formed from unsaturated vegetable oils by chemically reducing the double bonds with H atoms (a process termed *hydrogenation*).
   a. The hydrogenation process also converts some of the *cis* double bonds into *trans* double bonds, which are straight rather than kinked.
   b. This process generates partially hydrogenated or trans-fats; eating trans-fat raises the heart disease risk, and artificial trans-fat is now banned in several countries, with others considering similar measures.

D. A molecule of fat can contain 3 identical fatty acids or it can be a *mixed fat*, containing >1 fatty acid species; most natural fats like olive oil and butterfat are mixtures of molecules with different fatty acids.

E. Fats are very rich in chemical energy (a gram of fat contains more [>2X] energy than a gram of carbohydrates) and serve as structural components and as fuel molecules.
   1. Carbohydrates serve primarily as short-term, rapidly available energy source.
   2. Fat reserves store energy more efficiently on a long-term basis.

F. It is estimated that a person of average size contains ~0.5 kg of carbohydrate, primarily in the form of glycogen (~2000 kcal of total energy) and ~16 kg of fat (144,000 kcal of energy).
   1. During a strenuous day's exercise, a person can virtually deplete his or her body's entire carbohydrate store; fat stores take a long time to deplete (weight loss is difficult).

G. Since they lack polar groups and are extremely water-insoluble, they are stored in the form of dry lipid droplets in cells (extremely concentrated fuel storage since they contain no water, unlike glycogen granules).
   1. In many animals, fats are stored in special cells, *adipocytes*, whose cytoplasm is filled with one or a few large lipid droplets.
   2. Adipocytes show a remarkable ability to change their volume to accommodate varying fat quantities.

X. Steroids are complex and built around a characteristic 4-ringed hydrocarbon skeleton, where the joined rings differ in numbers and positions of double bonds and functional groups.

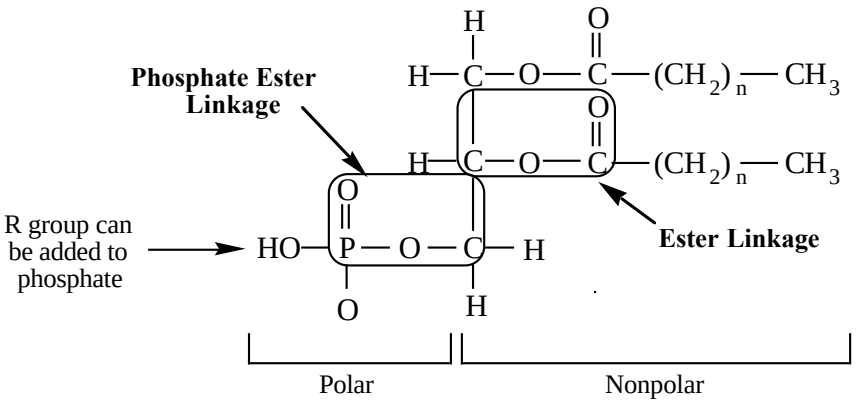
A. Most common and one of most important is *cholesterol*, a component of animal cell membrane, but not in internal membranes or in plants.
   1. A precursor for the synthesis of many steroid hormones (testosterone, progesterone, estrogen).

B. While largely absent from plant cells as vegetable oils are considered to be cholesterol-free, plant cells may contain large quantities of related compounds.

XI. Phospholipids (phosphoglyceride, diacylglycerol) resemble a fat (triacylglycerol), but has only 2 fatty acids and is a *diacylglycerol*, where the third OH of glycerol is covalently bonded to a phosphate group.

A. Structure: glycerol + 2 fatty acids + phosphate group on third hydroxyl, and the phosphate group is covalently bonded to a small polar group such as choline.
   1. Unlike fat molecules, phospholipids contain 2 ends that have very different properties.
   a. The end containing the phosphate group has a distinctly hydrophilic character.
   b. The other end has a distinctly hydrophobic character.

B. Major cellular function - presence in membranes (properties of which depend on phospholipids).



XII. Amino acids are the building blocks of proteins, which are very large macromolecules that act as the molecular tools and machines that make things happen.

A. Traits and functions have a more varied role than other organism's molecules (enzymes, structural or both) and execute almost all of the cell's activities.

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

B. The wide variety of protein functions comes from the virtually unlimited molecular structures they, *as a group*, can assume.

1. They can exhibit a great variety of structures and thus a great variety of activities.
2. Each protein has a unique and defined structure enabling it to carry out a particular function.
3. Their shapes and surfaces allow them to interact *selectively* with other molecules, so they have a high degree of **specificity**.
   a. A certain DNA-cutting enzyme recognizes DNA segment containing a specific 8-nucleotide sequence.
   b. At the same time, it ignores all of the other 65,535 possible sequences composed of 8 nucleotides.

C. Proteins are polymers of amino acid monomers and their polymer sequences give them unique properties.

1. Many protein capabilities can be understood by examining the chemical properties of its constituent amino acids.
2. Twenty different amino acids are used to construct proteins, whether from a virus or a human.
3. Two aspects of amino acid structure: that which is common to all and that which is unique to each.

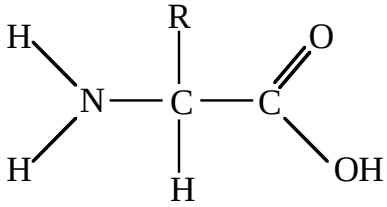
XIII. Basic monomer building block is the amino acid and central C with 4 attached groups, shared properties.

A. Amino acid backbone is α-carbon between NH_2 and COOH groups.

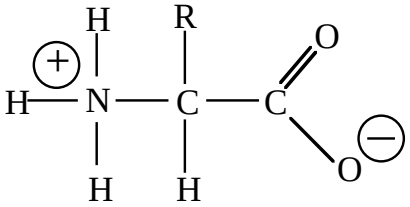
B. In aqueous environment, the α-carboxyl group (-COOH) ionizes to $-COO^-$ and the α-amino group ($-NH_2$) ionizes to $-NH_3^+$

C. Carbon atoms bonded to 4 different groups can exist in 2 configurations (*stereoisomers*) that cannot be superimposed on one another; amino acids have such asymmetric carbon atoms.
   1. Except for glycine, the α-carbon of amino acids bonds to 4 different groups so that each amino acid can exist in either a D or an L stereoisomer form; only L-form used in proteins made on ribosomes.
   2. The "selection" of L-amino acids must have occurred early in cellular evolution and has been conserved for billions of years.
   3. Microorganisms, however, use D-amino acids to synthesize certain small peptides, including those of the cell wall and several antibiotics (e.g., gramicidin A).

Generalized Amino Acid



Generalized Amino Acid at Physiological pH

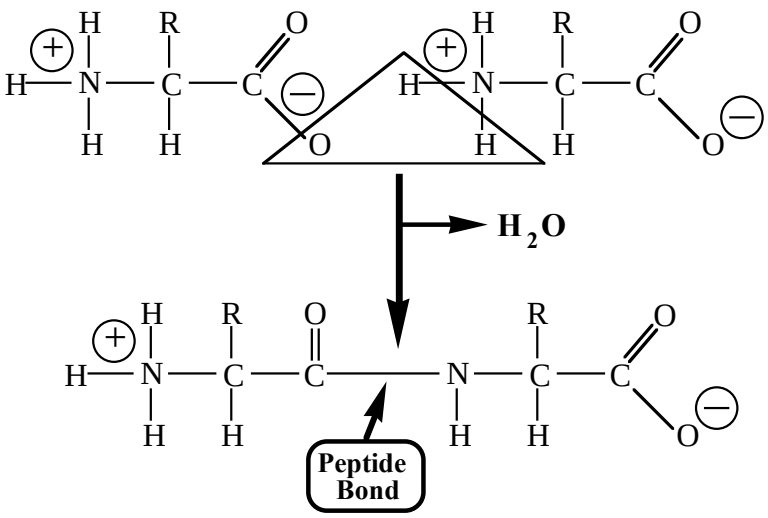


R = Radical or Side Group

D. During protein synthesis, each amino is joined to 2 other amino acids forming a long, continuous, unbranched polymer (**polypeptide chain**).
   1. Peptide bonds form by condensation reactions through the elimination of water molecules, attaching α-carboxyl of one amino acid to α-amino group of another; form polypeptides by repeating the process.
   2. Average polypeptide chain has ~450 amino acid residues; longest known protein is the muscle protein titin (>30,000 amino acids).
   3. Once incorporated into polypeptide chain amino acids are termed *residues*; residue on the *N-terminus* contains free α-amino group while the *C-terminus* has a free α-carboxyl group.

E. In addition to amino acids, many proteins contain other types of components that are added after the polypeptide is synthesized.
   1. Carbohydrates to form glycoproteins.
   2. Metal-containing groups to form metalloproteins.
   3. Organic groups, e.g., flavoproteins.

A Dipeptide Formed By Condensation-Dehydration



XIV. Properties of **R group** or **side chains** determine *inter*- and *intra*molecular interactions that stabilize molecular structure and protein activities, respectively and give amino acids their variability.

A. Polypeptide backbone is made of that portion of each amino acid that is common to all of them.
   1. Common amino acid backbones allow polymerization to form by the same reaction no matter which two amino acids are being joined.

B. Side chain (R group) bonded to α-carbon is highly variable among the 20 building blocks; this ultimately gives proteins their diverse structures and activities.
   1. The various side chains, considered together, exhibit a large variety of structural features, ranging from fully charged to hydrophobic.
   2. They can participate in wide variety of covalent and noncovalent bonds.

C. Side chains of enzyme "active sites" can facilitate (catalyze) many different organic reactions; assorted characteristics of amino acid side chains are important in:
   1. *Intra*molecular interactions determine molecule's structure and activity; and
   2. *Inter*molecular interactions determine polypeptide relationship with other molecules, including other polypeptides.

XV. Four amino acid and 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. ***Polar charged*** contain R groups that are relatively strong organic acids and bases; can form ionic bonds.
   1. Almost always present in fully charged state (lysine, arginine, aspartic acid, glutamic acid) at pH 7 (physiological pH).
   2. Consequently, they can form ionic bonds due to charges with other charged species in cell; histones with arginine (+-charge) bind to negatively charged phosphate groups of DNA.
   3. Histidine is also usually considered to be a polar, charged amino acid.
      a. It is usually only partially charged at pH 7.
      b. Because of its ability to gain or lose a proton in physiologic pH ranges, histidine is a particularly important residue in the active site of many proteins.

B. ***Polar uncharged*** have R groups that are weakly acidic or basic with a partial negative or positive charge, so can form H bonds with other molecules including water.
   1. Asparagine and glutamine [amides of aspartic and glutamic acid, respectively], threonine, serine, and tyrosine.
   2. They are often quite reactive.

C. ***Nonpolar*** have hydrophobic R groups, so they generally lack O and N and cannot interact with water or form electrostatic bonds.
   1. Alanine, valine, leucine, isoleucine, tryptophan, phenylalanine, and methionine.
   2. They vary primarily in size and shape, which allows one or another of them to pack tightly into a particular space within the core of a protein.
   3. In protein core, they associate with one another via hydrophobic interactions and van der Waals forces.

D. ***The other 3 amino acids*** are glycine, proline, cysteine, with unique properties separating them from the others.
   1. Glycine (R = H) is a very important amino acid.
      a. Small R group makes backbone flexible and able to move so it is useful in protein hinges.
      b. Owing to lack of a side chain, glycine residues provide a site where backbones of 2 polypeptides (or 2 segments of the same polypeptide) can approach one another very closely.
      c. Also, glycine is more flexible than other amino acids and allows part of backbone to move or form hinge.
   2. Proline is unique in with its α-amino group as part of a ring (making it an imino acid) with its R group.
      a. Hydrophobic amino acid that doesn't readily fit into orderly secondary structure like the α-helix.
      b. It often produces kinks in backbone or hinges.
   3. Cysteine has a reactive –SH for an R group, and forms covalent **disulfide** (—S—S—) **bridge** with other cysteines often at some distance away in the polypeptide backbone or even in 2 separate chains.
      a. Stabilizes the intricate shapes of proteins, particularly those outside of cells where they are subjected to added chemical and physical stress.
      b. Hair is made of a large number of filaments composed of the cysteine-rich protein keratin, cross-linked together by disulfide bridges.
         (1) When someone gets a "perm" to make hair curlier, hairdresser puts the hair into curlers and then adds reducing agent that breaks disulfide bridges, letting the keratin filaments slide past each other.
         (2) When the reducing agent is washed out, disulfide bridges re-form, locking the keratin in the new positions and causing the hair to be permanently locked into the shape imparted by the curlers.
         (3) The same principle is used in hair-straightening treatments, except that in this case the hairs are pulled out straight before the reducing agent is applied.

E. Not all amino acids are seen in all proteins, nor are the various amino acids distributed in an equivalent manner.

XVI. A number of other amino acids are also seen in polypeptide chains.

A. They arise from R group alteration of the 20 basic amino acids *after* their incorporation into the polypeptide, so are called **posttranslational modifications** (**PTMs**).

B. Dozens of different types of PTMs have been documented, the most widespread of which is the covalent addition of a phosphate group to a serine, threonine or tyrosine residue.

1. PTMs can generate dramatic changes in the properties and function of a protein, most notably by modifying its interactions with other molecules or shortening its lifespan in cell.
   a. Dozens of different types of PTMs have been documented.
   b. The most widespread and important PTM is the reversible addition of a phosphate group to a serine, threonine or tyrosine residue.
   c. Lysine acetylation is another PTM affecting thousands of proteins in a mammalian cell.
2. PTMs can generate dramatic changes in protein properties and functions, most notably by modifying:
   a. Its 3D structure.
   b. Level of activity.
   c. Localization within the cell.
   d. Life span.
   e. Its interactions with other molecules.
3. The presence or absence of a single phosphate group on a key regulatory protein has the potential to determine whether a cell will behave as a cancer cell or a normal cell.
4. Because of PTMs, a single polypeptide can exist as a number of distinct biological molecules.

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

A. Most soluble (i.e., nonmembrane) proteins are constructed so that polar residues are on the molecule surface.
   1. There they can associate with surrounding H_2O and contribute to protein's solubility in aqueous solution.

B. In contrast, nonpolar residues are situated predominantly in the core of a protein.
   1. These hydrophobic residues are usually tightly packed together, creating a type of 3D puzzle excluding H_2O.
   2. Contribute substantially to overall stability of protein and are the driving force during protein folding.
   3. Reactive polar groups can project into the nonpolar interior, giving a protein its catalytic activity.
      a. Nonpolar environment can enhance ionic interactions between charged groups that would be lessened by competition with water in aqueous environment.
      b. Reactions that might proceed at an imperceptibly slow rate in H_2O occur in millionths of a second within protein.

XVIII. Proteins are a good illustration of intimate relationship between form and function.

A. The structure of most proteins is completely defined and predictable.

B. Each amino acid in one of these giant macromolecules is located at a specific site within the structure, giving the protein the precise shape and reactivity required for the job at hand.

C. Protein structure can be described at several levels of organization, each emphasizes a different aspect and each is dependent on different types of interactions.
   1. Customarily, 4 such levels are described: *primary*, *secondary*, *tertiary* and *quaternary*.
   2. The first, primary structure, concerns the amino acid sequence of a protein; the latter 3 levels concern the organization of the molecules in space.
   3. To understand the mechanism of action and biological function of a protein, it is essential to know how that protein is constructed.

XIX. Primary (1°) structure is the specific linear sequence of amino acids that constitute the chain; all levels of structure are ultimately determined by the primary level.

A. Number of different chains that can be made = 20^n, where n = number of amino acids in chain.
   1. Most polypeptides have >100 amino acids.
   2. The variety of possible sequences is essentially unlimited.

B. The information for the precise order of amino acids in every protein that an organism can produce is encoded within that organism's genome.

C. Amino acid sequence contains most, if not all, information needed to determine protein's 3D shape and thus its function.
   1. The amino acid sequence, therefore, is all-important and changes in the sequence arising from mutation may not be readily tolerated.
   2. The earliest and best-studied example is the change in the hemoglobin amino acid sequence that causes *sickle cell anemia*.
      a. This serious, inherited anemia results from a single change in the amino acid sequence in the hemoglobin molecule.
      b. A nonpolar valine residue is present where a polar, charged glutamic acid is normally located.
      c. The change in hemoglobin structure can have a dramatic effect on red blood cell shape, converting them from disk-shaped cells to sickle-shaped cells.
      d. The sickle shaped cells tend to clog small blood vessels, causing pain and life-threatening crises.
   3. Not all amino acid changes have as dramatic an effect.
      a. This is evidenced by differences in amino acid sequence in the same protein among related organisms.
      b. The degree to which changes in the primary sequence are tolerated depends on the degree to which the protein shape or the critical functional residues are disturbed.

D. The first protein sequenced was the protein hormone insulin determined by Frederick Sanger and co-workers (Cambridge University, early 1950s), a momentous feat in new molecular biology field.
   1. Beef insulin was chosen for its availability and small size (2 polypeptide chains of 21 and 30 aminos each).
   2. Insulin *sequencing* showed that proteins have a definable substructure that is neither regular nor repeating, unlike those of polysaccharides.
      a. Each particular protein has a precise amino acid sequence that doesn't vary from 1 molecule to another.
   3. With the advent of techniques for rapid DNA sequencing, the primary structure of a polypeptide can be deduced from the nucleotide sequence of the encoding gene.
   4. In the past few years, complete sequences of genomes of hundreds of organisms, including humans, have been determined.
      a. This will eventually allow researchers to learn about every protein an organism can make.
      b. But translating information about primary sequence into knowledge of higher protein structure levels remains a formidable challenge.

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

A. All matter exists in space and therefore has a 3D expression, and since proteins are formed by linkages among vast numbers of atoms, consequently their shape is complex.

1. The term **conformation** refers to 3D arrangement of a molecule's atoms, that is, to their spatial organization.

B. Linus Pauling and Robert Corey (Cal Tech) studied the structure of simple peptides consisting of a few amino acids linked together.
   1. They concluded that polypeptide chains exist in preferred conformations that provide the maximum possible number of H bonds between neighboring amino acids.
   2. Two conformations were proposed: the **alpha (α) helix** and **beta (β) sheet**.

C. In the α-helix, the backbone of the polypeptide assumes the form of cylindrical, twisting spiral; the backbone lies on the inside of the helix, and the R groups project outward.
   1. The helical structure is stabilized by H bonds between atoms of one peptide bond and those situated above and below it along the spiral; H bonds are parallel to molecular axis.
   2. Seen in X-ray diffraction patterns of actual proteins in the 1950s; found in the protein keratin from hair and later in various oxygen-binding proteins like myoglobin and hemoglobin.
   3. Surfaces on opposite sides of α-helix may have contrasting properties – in water-soluble proteins, often polar amino acids are on outside of helix in contact with solvent with nonpolar R groups facing inward.

D. β-pleated sheets consist of several polypeptide segments lying side-by-side; the backbone of each segment of polypeptide (or β-strand) in a β-sheet assumes a folded or pleated conformation, unlike α-helix.
   1. Like α-helix, the β-sheet is also characterized by a large number of H bonds, but these are oriented perpendicular to polypeptide chain long axis and project across from one part of the chain to another.
      a. The strands of a β-sheet can be arranged either parallel to one another, with neighboring strands running in the same direction.
      b. They can be antiparallel, with neighboring strands running in opposite directions.
   2. Like α-helix, the β-sheet has also been found in many different proteins.
   3. Because β-strands are highly extended; the β-sheet resists pulling (tensile) forces; very strong (ex.: silk fibroin contains a lot of β-sheet).
      a. Silk fibers are thought to owe their strength to this architectural feature.
      b. Remarkably, 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.

E. Polypeptide hain portions not organized into α-helix or β-pleated sheet may consist of hinges, turns, loops, or finger-like extensions.
   1. Often these are the most flexible portions of chain and sites of greatest biological activity.
   2. For example, antibody molecules are known for their specific interactions with other molecules (antigens); these interactions mediated by a series of loops at one end of the antibody molecule.

F. In drawings of protein structure showing secondary structure, α-helices are most simply represented by helical ribbons and β-strands are represented as flattened arrows.

XXI. Tertiary (3°) structure, the level above secondary structure, is the conformation of the entire protein.

A. It results from an intramolecular noncovalent interaction array between the diverse R groups in same chain, while 2° structure is stabilized primarily by H bonds between atoms forming the backbone peptide bonds.
   1. Tertiary structures are virtually unlimited in number of structures unlike secondary structure that is largely limited to only a few conformations.

B. **X-ray crystallography** can be used to determine detailed tertiary structure.
   1. In X-ray crystallography, a protein crystal is bombarded by a thin X-ray beam.

2. The radiation scattered (diffracted) by the electrons of the protein's atoms is allowed to strike radiation-sensitive detector.
3. This forms an image of spots, the patterns of which do not directly show the protein structure.
   a. Computer programs based on the mathematics of diffraction can be used to derive the structure responsible for producing the pattern.

C. Tertiary structure can also be determined by **nuclear magnetic resonance (NMR) spectroscopy**, a method in which proteins in solution are placed in a powerful magnetic field.
1. They are then probed with radio waves to produce a spectrum, providing information about the distances between atoms.
2. As with X-ray crystallography, the spectrum itself does not directly show protein structure, but computer programs can derive the structures most likely to give the observed spectrum.

D. X-ray crystallography and NMR have different strengths and weaknesses.
1. X-ray crystallography can provide higher resolution structures for larger proteins but is limited by the ability to get any given protein to form pure crystals.
2. NMR does not require crystallization, can provide information about dynamic changes in protein structure and can rapidly reveal drug-binding sites on a protein.
   a. It becomes increasingly difficult to apply as the size of a protein increases.

E. A newer methodology is based on the use of electron microscopy to elucidate the 3° structure of proteins.
1. Electron microscopy is traditionally a method to visualize the structure of cellular organelles..
2. To prevent the electron beam from destroying the proteins before the image can be collected, proteins are embedded in ice and imaged at low temperatures.
3. Imaging a single proteins gives a fuzzy image, but imaging upwards of millions of proteins by computer gives the protein structure with high precision.
4. The technique, called *cryoelectron microscopy* (*cryo-EM*), can reach atomic resolution without needing to form protein crystals.
5. Cryo-EM may be the method of choice in the future for detemining protein structure.

XXII. For many years, it was presumed that all proteins had a fixed 3D structure, which gave each protein its unique properties and specific functions.

A. Over the past decade or so, it has been discovered surprisingly that many proteins of higher organisms contain sizable segments that lack a defined conformation.
1. Examples of proteins containing these types of unstructured or *disordered* segments can be seen in models of the PrP protein and the histone tails.
2. The disordered regions in these proteins are depicted as dashed lines conveying the fact that these polypeptide segments (like pieces of spaghetti) can occupy many different positions.
   a. Since they occupy many different positions, these segments can't be studied by X-ray crystallography.
3. Disordered segments tend to have a predictable amino acid composition, being enriched in charged and polar residues and deficient in hydrophobic residues.
4. Disordered regions play key roles in vital cell processes (often binding DNA or other proteins) even though it may seem strange that regions lacking fully defined structure can engage in useful function.
5. Remarkably, these segments often undergo a physical transformation once they have bound to an appropriate partner and are then seen to possess a defined, folded structure.

XXIII. Most proteins can be categorized on the basis of their overall conformation as being either fibrous or globular.

A. **Fibrous proteins** have a highly elongated shape, with long strands or flattened sheets that resist pulling or shearing forces to which they are exposed.
   1. Most proteins that act as structural materials outside living cells are usually fibrous proteins.
   2. Collagens and elastins of connective tissue, keratins of hair, skin, fingernails, silk are examples of fibrous proteins.

B. **Globular proteins** comprise most proteins in the cell and have a compact shape with chains folded and twisted into complex shapes.
   1. Distant points on the linear sequence of amino acids are brought next to each other and linked by various types of bonds.
   2. The first glimpse of at the tertiary structure of a globular protein was obtained by John Kendrew et al., Cambridge Univ. (1957) who used X-ray diffraction patterns of myoglobin.

C. Myoglobin structure as a globular protein example, it functions as a storage site for O_2 in muscle tissue.
   1. O_2 binds to Fe atom at heme group center which gives muscle its reddish color, heme is a prosthetic group (part of protein not made of amino acids added after translation on ribosome.
   2. First report provided lower resolution – molecule is compact (globular), folds back on itself in complex arrangement.
      a. There was no evidence of regularity or symmetry within the molecule like that reported for DNA.
      b. This early crude profile revealed 8 rodlike α-helical stretches (7-24 amino acids long; ~75% of chain, an unusually high percentage compared with other proteins examined); there was no β-pleated sheet.
   3. Later X-ray diffraction data provided a more detailed picture.
      a. Heme group is situated within a pocket of hydrophobic R groups that promotes O_2 binding without Fe atom oxidation (electron loss).
      b. Myoglobin has no disulfide bonds; its 3° structure is totally held together by noncovalent interactions.
      c. All of the noncovalent bonds thought to occur between R groups within proteins (H bonds, ionic bonds, hydrophobic interactions) have been found.
      d. Unlike myoglobin, most proteins have both α-helix and β-pleated sheet.
      e. Most importantly, these early landmark studies revealed that each protein has a unique tertiary structure that can be correlated with its amino acid sequence and its biological function.

XXIV. Similarity in primary sequence is often used to decide whether 2 proteins have similar structure and function; sometimes proteins that appeared unrelated at 1° sequence level have been seen to have similar 3° structure.

A. Because 3° structure determines the interactions and enzymatic activity of a protein, such structural similarity indicates that these proteins may have similar functions.

B. Actin from eukaryotic cells and MreB from bacteria have 1° sequences that show no similarity, but their 3° structures are clearly related and both proteins form cytoskeletal filaments.

XXV. Unlike myoglobin, most eukaryotic proteins are made of ≥2 spatially distinct modules (**domains**) that fold independent of one another; often represent parts that function in semi-independent manner.

A. Phospholipase C consists of 4 distinct domains; the different domains of a polypeptide often represent parts of the protein that function in a semi-independent manner.

1. Domains might bind different factors, like a coenzyme and a substrate or a DNA strand and another protein, or they might move relatively independent of one another.
2. Protein domains are often identified with a specific function, e.g. PH domains and chromodomains.
   a. PH domains bind to membranes containing a specific phospholipid, whereas proteins containing a chromodomain bind to a methylated lysine residue in another protein.
   b. The functions of a newly identified protein can usually be predicted by the domains of which it is made.

B. Proteins with >1 domain may have arisen during evolution by the fusion of genes that encoded different ancestral proteins; each domain represents a part that was once a separate molecule.
   1. Mammalian phospholipase C - each domain has been identified as a homologous unit in another protein.

C. Some domains have been found only in one or a few proteins; other domains have been shuffled widely about during evolution.
   1. These widely shuffled domains appear in variety of proteins whose other regions show little or no evidence of an evolutionary relationship.
   2. Shuffling of domains creates proteins with unique combinations of activities.
   3. On average, mammalian proteins tend to be larger and contain more domains than proteins of less complex organisms (fruit flies, yeast).

XXVI. Dynamic changes within proteins – X-ray crystallographic structures possess exquisite detail, but they are static images frozen in time.

A. Proteins, in contrast, are not rigid and inflexible, but capable of considerable internal movements; proteins have "moving parts."
   1. Since they are so tiny, they can be greatly influenced by the energy of their environment.
   2. Random, small-scale fluctuations in protein bond arrangement create an incessant thermal motion within the molecule.

B. Spectroscopic techniques like NMR can be used to monitor such dynamic movements within proteins.
   1. They reveal H bond shifts, waving movements in of external side chains, and the full rotation of the aromatic rings phenylalanine and tyrosine residues about one of the single bonds.

C. The important role that such movements can play in a protein's function is illustrated by acetylcholineaterase studies; it degrades acetylcholine left behind after transmission of impulse from one nerve cell to another.
   1. When its 3° structure was first revealed by X-ray crystallography, there was no obvious pathway for acetylcholine to enter the enzyme's catalytic site, situated at the bottom of a deep gorge in molecule.
   2. In fact, active site's narrow entrance was completely blocked by a number of bulky amino acid side chains.
   3. High-speed computer analysis was used to simulate the random movements of thousands of atoms in the enzyme; can't be done using experimental techniques.
      a. These molecular dynamic (MD) simulations showed that movements of the side chains in the protein would lead to rapid opening and closing of a "gate," allowing acetylcholine to diffuse into catalytic site.

D. The X-ray crystallographic structure of a protein (i.e., its *crystal structure*) can be considered to be an average structure or "ground state."
   1. A protein can undergo dynamic excursions from the ground state and assume alternat conformations that are accessible based on the amount of energy that the protein contains.

XXVII. **Conformational changes** are predictable (nonrandom) movements within a protein that are triggered by binding of a specific molecule; they accompany virtually every activity in which a protein takes part.

A. Conformational changes typically involve the coordinated movements of various parts of the molecule.

B. Examples of protein molecules that have been shown to undergo conformational changes.

1. Dramatic conformational change occurs in the bacterial protein GroEL as it interacts with another protein, GroES.
2. Muscle myosin binds actin and a small (20°) rotation of the myosin head moves adjacent actin filament 50 - 100Å during muscle contraction.
   a. This dynamic event is very important and can be appreciated when one considers that the movement of one's body results from the additive effect of millions of conformational changes.
   b. These changes occur within the muscle's contractile proteins.

XXVIII. Most proteins are made up of more than one chain, or **subunit** and the **quaternary** (4°) **structure** is the linking of polypeptide chains to form multisubunit functional protein via intermolecular R group interactions.

A. Subunits may be linked by covalent disulfide bonds, but most often held together by noncovalent bonds as occurs typically between hydrophobic patches on neighboring polypeptides' complementary surfaces.

B. Proteins composed of ≥ 1 subunit are said to have quaternary structure and depending on the protein, the polypeptide chains may be identical or nonidentical.

1. Protein complex composed of 2 identical subunits is a *homodimer*.
2. Protein complex composed of 2 nonidentical subunits is a *heterodimer*.

XXIX. The structure of hemoglobin, the best studied multisubunit protein, was first studied by Max Perutz at Cambridge (1959) as an early molecular biology landmark. Made from 2 α-globin and 2 β-globin subunits, each binds one molecule of O_2.

A. Each subunit has 3° structure similar to that of myoglobin which suggested that the 2 proteins had evolved from a common ancestral polypeptide with a common O_2-binding mechanism.

B. Perutz also compared the structure of the oxygenated and deoxygenated versions of hemoglobin.

1. He found that O_2-binding was accompanied by movement of the bound iron atom closer to the plane of heme group.
2. This seemingly inconsequential shift in position of a single atom pulled on an α-helix to which the iron atom is connected.
   a. This led to a series of increasingly larger movements within and between the subunits.
3. This finding revealed for the first time that the complex functions of proteins may be carried out by means of small changes in their conformation.

C. Hemoglobin consists of 4 subunits, but is still considered to be a single protein with a single function.

XXX. There are many known different proteins, each with a specific function, that physically associate to form a much larger complex called a **multiprotein complex**.

A. One of the first to be discovered and studied was the *E. coli* pyruvate dehydrogenase complex consisting of 60 polypeptide chains constituting 3 different enzymes in a stable complex.

1. Its enzymes catalyze a reaction series connecting 2 metabolic pathways: glycolysis and the TCA cycle.
2. Because the enzymes are so closely associated, the product of one enzyme can be channeled directly to the next enzyme in sequence without becoming diluted in the cell's aqueous medium.

B. Some associations are stable while some are transient, driven by complementary surfaces as part of one fits in pocket on other, and stabilized by noncovalent bonds.
   1. Most proteins interact with others in highly dynamic patterns, associating and dissociating depending on conditions within the cell at any given time.
   2. Interacting proteins tend to have complementary surfaces; projecting portion of one protein fits into a pocket within its partner; once in close contact, their interaction is stabilized by noncovalent bonds.

C. The SH3 domain is found in >200 different proteins involved in molecular signaling.
   1. The SH3 domain surface contains shallow hydrophobic pockets that become filled by complementary knobs projecting from another protein.

D. A large number of different structural domains have been identified that, like SH3, act as adaptors to mediate interactions between proteins.
   1. Often protein-protein interactions are regulated by modifications like phosphate group addition to a key amino acid, which may act as a switch to turn on or off the protein's ability to bind a protein partner.
   2. As more and more complex molecular activities have been discovered, the importance of protein interactions has become increasingly apparent.
      a. Transient protein interactions are important in DNA synthesis, ATP formation, and RNA processing.
      b. All of these are done by "molecular machines" made of many interacting proteins, some of which form stable relationships, while others form transient liaisons.
      c. Several hundred different protein complexes have been purified in large-scale yeast studies.

XXXI. 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 and 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, protein 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	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

		interactions	functions. Proteins flexible and 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 and usually bound together by noncovalent bonds (electrostatic or hydrophobic). Homodimers - 2 identical subunits; heterodimers - at least 2 nonidentical subunits (Ex.: hemoglobin)

XXXII. How does such a complex, folded, asymmetric organization, like that of a protein, arise in a cell?

XXXIII. Christian Anfinsen (NIH, 1956) gained the first insight into protein folding; it was serendipitous.

A. Anfinsen studied the properties of ribonuclease A (RNase), a small enzyme consisting of 1 chain of 124 aa with 4 disulfide bonds linking various parts of the chain.

B. To denature RNase, typically break disulfide bonds with mercaptoethanol, a reducing agent that converts a disulfide bridge to a pair of sulfhydryl (-SH) groups.
   1. To make the disulfide links accessible to mercaptoethanol, the protein was first partially unfold.
   2. **Denaturation**, or the unfolding or disorganization of a protein, is done with a variety of agents: detergents, organic solvents, radiation, heat, and compounds like urea and guanidine chloride to break interactions that stabilize 3° structure and denature protein.
   3. Anfinsen treated RNase with concentrated urea and mercaptoethanol to eliminate enzyme activity by unfolding RNase.

C. When he removed urea and mercaptoethanol, normal enzyme activity returned.
   1. The reformed molecules were indistinguishable structurally and functionally from the correctly folded (**native**) ones present at the beginning of the experiment.
   2. He concluded that the linear amino acid sequence contained all of the information needed for formation of 3D-conformation; RNase is capable of **self-assembly**.
   3. Events tend to progress toward states of lower energy, so the 3° structure that a polypeptide chain assumes after folding is the accessible structure with lowest energy.
      a. It is the most thermodynamically stable structure that can be formed by that chain.
      b. It appears that evolution selects for those amino acid sequences that generate a polypeptide capable of spontaneously arriving at a meaningful native state in a biologically reasonable time period.

XXXIV. In the field of protein folding dynamics, there have been many controversies.

A. Many stem from fact that the field is characterized by highly sophisticated experimental, spectroscopic, and computational procedures needed to study complex molecular events usually occurring in microseconds.
   1. These efforts have yielded conflicting results and generated data open to more than one interpretation.
   2. For simplicity, it is good to concentrate on simple proteins like RNase that consist of a single domain.

B. A hotly debated issue is whether all members of a population of unfolded proteins of a single species fold along a similar pathway or fold by means of a diverse set of routes that converge on the same native state.
   1. Recent studies with high-speed computers that have simulated folding suggested that both views are right and proteins initially explore a wide range of different combinations when they first start to fold.
   2. Eventually they funnel down into an increasingly restricted set of possible configurations.

C. Another debated issue concerns the types of events that occur at various stages during the folding process.
   1. Secondary structure may form before tertiary structure forms.
      a. Protein folding is initiated by interactions among neighboring residues that lead to the formation of much of the secondary structure of the molecule.
      b. Once α helices and β sheets are formed, subsequent folding is driven by hydrophobic interactions that bury nonpolar residues together in the central core of the protein.
   2. A loose tertiary structure may form before secondary structure is set.
      a. The first major event in protein folding is the collapse of polypeptide to form a compact structure in which the backbone adopts a native-like topology.
      b. Only then, after the collapse, does significant secondary structure develop.

D. Recent studies suggest that the two pathways above lie at opposite extremes with most proteins probably folding by a middle-of-the-road scheme.
   1. Secondary structure formation and compaction occur simultaneously in this scheme.
   2. These early folding events lead to formation of a partially folded, transient structure that resembles native protein.
   3. The transient structure formed lacks many of the specific interactions between amino acid side chains that are present in the fully folded molecule.

XXXV. If the information that governs folding is embedded in a protein's amino acid sequence, then it should be possible to predict the 3° structure of the protein just from its sequence.

A. Such *de novo* structural prediction has been a holy grail in protein science and is still an unsolved problem, although structures on the order of 100 amino acids can be solved by de novo prediction.

B. However, if a protein of interest is closely related at the 1° sequence level with another protein whose 3° structure is known, then one can make a reasonable guess about the unknown protein's 3° structure.
   1. This is done by aligning the amino acids of the unknown protein onto the corresponding amino acids in the protein whose structure is known; this process is known as **threading**.

XXXVI. The fact that 1° sequence determines the folding of a protein means that alterations in this sequence have the potential to change the way a protein folds, leading to an an abnormal 3° structure.

A. Many mutations responsible for inherited disorders have been found to alter a protein's 3D structure.

B. Sometimes the consequences of protein misfolding can be fatal.

THE HUMAN PERSPECTIVE: PROTEIN MISFOLDING CAN HAVE DEADLY CONSEQUENCES

I. Paper published in medical journal *Lancet* (April 1996) caused widespread alarm in the populations of Europe; it described a study of 10 persons who had been afflicted with Creutzfeld-Jakob disease (CJD).

A. It is a rare, fatal disorder that attacks the brain, causing a loss of motor coordination and dementia.
   1. Like other diseases, it can occur as an inherited disease that runs in certain families or as a sporadic form that appears in individuals who have no family history of the disease.

B. Unlike almost every other heritable disease, however, CJD can also be *acquired*.

1. Until recently, persons who had acquired CJD had been recipients of organs or organ products donated by a person with undiagnosed CJD.
2. The *Lancet* article reported on CJD that had apparently been acquired from contaminated beef that infected individuals had eaten years before, instead from another person.
3. The contaminated beef was derived from cattle raised in England that had contracted a neurodegenerative disease that caused animals to lose motor coordination and develop demented behavior.
4. The disease has come to be commonly known as "mad cow disease."

C. Patients who got CJD from contaminated beef can be distinguished by several criteria from those who suffer from the classical forms of the disease.

D. To date, ~200 people have died of CJD acquired from contaminated beef, and the numbers of such deaths are declining.
1. On the surface, this would suggest that the epidemic has run its course, but there are several reasons for public health officials to remain concerned.
2. For one, studies of tissues that had been removed during surgeries in England indicate that thousands of people are likely to be infected with the disease without exhibiting symptoms.
   a. Even if these individuals never develop clinical disease, they remain potential carriers who could pass CJD on to others through blood transfusions.
   b. In fact, at least 2 individuals are believed to have contracted CJD after receiving blood from a donor harboring the disease.
   c. These findings underscore the need to test blood for the presence of the responsible agent.

II. A disease that runs in families can invariably be traced to a faulty gene, while diseases acquired from a contaminated source can invariably be traced to an infectious agent – how can a disease arise both ways?

A. The answer has emerged gradually over the past several decades.

B. Understanding started with observations by D. Carleton Gajdusek in 1960s on strange malady that once afflicted the native population of Papua, New Guinea.
1. Gajdusek showed that islanders were contracting a fatal neurodegenerative disease (kuru) during a funeral ritual in which they ate brain tissue of a recently deceased relative.
2. Brain autopsies of kuru patients who died had distinct pathology (*spongiform encephalopathy*) in which certain brain regions were riddled with microscopic holes (vacuolations); tissue looked like a sponge.

C. It was soon shown that the brains of kuru-suffering islanders were strikingly similar in microscopic appearance to the brains of those afflicted with CJD.
1. Raised a question: Does the brain of a CJD sufferer with the inherited form contain an infectious agent?
2. Gajdusek (1968) injected extracts prepared from the brain biopsy of a dead CJD victim into a suitable lab animal and the animal developed spongiform encephalopathy similar to kuru or CJD.
3. Concluded that the extracts contained an infectious agent presumed, at the time, to be a virus.

D. Stanley Prusiner (Univ. of Cal., SF, 1982) reported that, unlike viruses, the CJD-causing agent lacked nucleic acid and instead was composed solely of protein; called the protein a *prion* (meaning protein only).
1. This "protein only" hypothesis was originally met with considerable skepticism, but subsequent studies (Prusiner and others) have provided overwhelming support.
2. Prion protein was first thought to be an external agent, a virus-like particle lacking nucleic acid.

E. However, the prion protein was soon shown to be encoded by a gene (*PRNP*) within cell's own chromosomes.

1. The gene is expressed in normal brain tissue and encodes a protein designated PrP^C (prion protein cellular) that resides at the surface of nerve cells; its precise function is a mystery.
2. A modified version of protein called PrP^{Sc} (prion protein scrapie) is present in brains of humans with CJD; unlike normal PrP^C, this protein accumulates within nerve cells (forms aggregates that kill cells).

F. In their purified states, PrP^C and PrP^{Sc} have very different physical properties.
1. PrP^C remains as monomeric molecule that is soluble in salt solutions and readily destroyed by protein-digesting enzymes.
2. In contrast, PrP^{Sc} molecules interact with one another to form insoluble fibrils that are resistant to enzymatic digestion.
3. Based on these differences, one might expect the two forms of the PrP protein to be composed of distinctly different amino acid sequences, but this is not the case.
4. Instead, they can have identical amino acid sequences, but differ in the way the polypeptide chains fold to form the 3D protein molecule.
5. PrP^C molecule consists largely of α-helical segments and interconnecting coils; the core of a PrP^{Sc} molecule consists largely of β sheet.
6. PrP can be converted from soluble, protease-sensitive conformation into insoluble, protease-insensitive aggregates *in vitro* by simply changing the conditions in test tube.

III. It's not hard to understand how a mutant polypeptide might be less stable and more likely to fold into abnormal PrP^{Sc} conformation, but how does it act as an infectious agent?

A. It is presently thought that abnormal prion molecule PrP^{Sc} can bind to a normal PrP^C protein and cause the normal protein to fold into the abnormal form.
1. This conversion can be shown to occur in test tube as the addition of PrP^{Sc} to preparation of PrP^C can convert PrP^C molecules into the PrP^{Sc} conformation.
2. According to this hypothesis, the appearance of the abnormal protein in body, whether due to a rare misfolding event as with sporadic disease or by exposure to contaminated beef, starts a chain reaction.
3. The chain reaction involves the gradual conversion of normal cell proteins to the misshapen prion form as they are recruited into growing insoluble fibrils.

B. The precise mechanism by which prions lead to neurodegeneration remains unclear.

IV. CJD is rare disease caused by a protein with unique infective properties, but Alzheimer's disease (AD), on the other hand, is common disorder that strikes up to 10% of people ≥65 and 40% who are ≥80.

A. Persons with AD exhibit memory loss, confusion and loss of reasoning ability.

B. CJD and AD share a number of important features.
1. Both are fatal neurodegenerative diseases that can occur in either inherited or sporadic form.
2. Like CJD, AD sufferer's brain contains fibrillar deposits of insoluble material called *amyloid.*
   a. It should be noted that the term *amyloid* is not restricted to the abnormal protein found in AD.
   b. Many different proteins are capable of assuming an abnormal conformation rich in β-sheet, which causes the protein monomers to aggregate into characteristic anyloid fibrils that bind certain dyes.
   c. Amyloid fibrils are defined by their molecular structure in which the β-strands are oriented perpendicular to the long axis of the fibrils.
   d. The PrP^{Sc}-forming fibrils of prion diseases are also described as amyloid.
3. In both CJD and AD, toxic fibrillar deposits result from the self-association of a polypeptide composed predominantly of β sheet.

C. There are also many basic differences between the 2 diseases.
   1. The proteins that form the disease-causing aggregates are totally unrelated.
   2. The parts of the brain that are affected are distinct.
   3. The protein responsible for AD does not act like an infectious agent (not contagious; it is nontransmissable from person-to-person); however, it may spread from cell ot cell within the brain.

V. Over the past 2 decades, AD research has been dominated by the *amyloid hypothesis*, which contends that the disease is caused by the production of a molecule, the *amyloid β-peptide* (*Aβ*).
   A. Aβ is originally part of a larger protein called *amyloid precursor protein* (*APP*), which spans the nerve cell membrane.
      1. Aβ peptide is released from APP molecule after cleavage by 2 specific enzymes, β-secretase and γ-secretase.
      2. The length of Aβ peptide is somewhat variable; the predominant species has length of 40 amino acids (Aβ40), but a minor species with 2 additional hydrophobic residues (Aβ42) is also produced.
   B. Both of these peptides can exist in a soluble form consisting predominantly of α helices, but Aβ42 has a tendency to spontaneously refold into a different conformation containing considerable β-pleated sheet.
      1. The Aβ42 version of the protein has the greatest potential to cause AD; it tends to self-associate to form small complexes (oligomers), as well as large aggregates that are visible as fibrils in EM.
      2. These amyloid fibrils are deposited outside of nerve cells in form of extracellular *amyloid plaques.*
   C. The issue is not settled, but a body of evidence suggests that it is the soluble oligomers that are most toxic to nerve cells, rather than the insoluble aggregates.
      1. Cultured nerve cells are much more likely to be damaged by the presence of soluble intracellular Aβ oligomers than by either Aβ monomers or extracellular fibrillar aggregates.
      2. In the brain, the Aβ oligomers appear to attack synapses that connect one nerve cell to another and eventually lead to the death of the nerve cells.
   D. People who suffer from an inherited form of AD carry a mutation that leads to an increased production of the Aβ42 peptide.
      1. Aβ42 overproduction can be caused by possession of extra copies (duplications) of *APP* gene, by mutations in *APP* gene or by mutations in genes (*PSEN1, PSEN2*) that encode subunits of γ-secretase.
      2. Individuals with such mutations exhibit symptoms of the disease at early age, typically in their 50s.
   E. Arguments for and against the amyloid hypothesis.
      1. The fact that all mutations associated with these inherited, early-onset forms of AD lead to increased Aβ42 production is strongest argument favoring amyloid formation as the underlying basis of AD.
      2. The strongest argument against the amyloid hypothesis is the weak correlation that can exist between the number and size of amyloid plaques in the brain and the severity of the disease.
         a. Elderly persons with little or no sign of memory loss or dementia can have relatively high levels of amyloid deposits in their brain and those with severe disease can have little or no amyloid deposition.

VI. All of the drugs currently on the market for the treatment of AD are aimed only at symptom management; none has any effect on stopping disease progression.
   A. With the amyloid hypothesis as the guiding influence, researchers have followed 3 basic strategies in the pursuit of new drugs for prevention and/or reversal of AD-associated mental decline; these strategies are:

1. To prevent the formation of the Aβ42 peptide in the first place.
2. To remove the Aβ42 peptide (or the amyloid deposits it produces) once it has been formed.
3. To prevent the interaction between Aβ molecules, thereby preventing the formation of both oligomers and fibrillar aggregates.

B. How do investigators determine what type of drugs might be successful in AD treatment or prevention?

VII. One of the best approaches to develop human disease treatments is to find lab animals, particularly mice, that develop similar diseases and use these animals to test the effectiveness of potential therapies.

A. Such animals that exhibit a disease that mimics a human disease are called *animal models.*

B. Brains of aging mice show no evidence of the amyloid deposits found in humans and up until 1995, there was no AD animal model.

1. In 1995, researchers found they could create mouse strain that developed amyloid plaques in the brain and performed poorly at tasks that require memory.
2. They created this strain by genetically engineering the mice to carry a mutant human *APP* gene, one responsible for causing AD in families.
3. These genetically engineered (*transgenic*) mice have been invaluable for testing potential AD therapies.

VIII. The greatest excitement in the AD therapeutics field has centered on the second strategy above; one can use these investigations to illustrate some of the steps required in developing a new drug.

A. Dale Schenk, et al. (Elan Pharmaceuticals, 1999) published an extraordinary finding, found that the formation of amyloid plaques in mice carrying the mutant human *APP* gene could be blocked.

1. This could be done by repeatedly injecting animals with the very same molecule that causes problem, the aggregated Aβ42 protein.
2. In effect, the researchers had immunized (i.e., vaccinated) the mice against the disease.
3. When young (6-week-old) mice were immunized with Aβ42, they failed to develop the amyloid brain deposits as they grew older.
4. When older (13-month-old) mice whose brains already contained extensive amyloid deposits were immunized with Aβ42, a significant fraction of fibrillar deposits was cleared out of the nervous system.
5. Even more important, the immunized mice performed better than their non-immunized littermates on memory-based tests.

B. These success of these experiments combined with the fact that the animals suffered no ill effects led government regulators to approve Phase I clinical trials of the Aβ42 vaccine.

1. This is the first step in testing a new drug or procedure in humans and usually comes after years of preclinical testing on cultured cells and animal models.
2. A Phase I clinical trial is the first step in testing a new drug or procedure in humans and usually comes after years of preclinical testing on cultured cells and animal models.
3. Phase I tests are carried out on a small number of subjects and are designed to monitor the safety of the therapy and the optimal drug dose, rather than its effectiveness against the disease.
4. None of subjects in 2 separate Phase I trials of the Aβ vaccine showed any ill-effects from injection of amyloid peptide.

C. Their success allowed them to go on to a Phase II clinical trials with a larger group of subjects; designed to get a measure of procedure/drug effectiveness (a randomized, double-blind, placebo-controlled study).

1. Patients randomly divided into 2 groups treated similarly except that one group given curative factor being tested and other group given placebo (inactive substance with no therapeutic value).

a. Patients are *randomly* divided into 2 groups that are treated similarly, except one is given the curative factor being tested and the other is given a placebo (inactive substance with no therapeutic value).
b. The study is double-blinded (means that neither researchers nor patients know who is receiving treatment and who is receiving placebo.

2. Phase II trial for Aβ vaccine enrolled 350 people (US, Europe) diagnosed with mild to moderate AD and began in 2001.
   a. After 2 injections of synthetic β-amyloid (or placebo), 6% of subjects experienced a potentially life-threatening brain inflammation; most were successfully treated with steroids but study was stopped.
   b. More recently, vaccination trials have been conducted using fragments of Aβ protein that do not induce inflammation, but the results of these trials have been disappointing.
      i. The smaller fragments did not cause inflammation, but had little effect on the disease.
   c. Current vaccination trials are geared towards prevention in a genetically predisposed patient population.

D. Once it was apparent that vaccination of patients with Aβ42 had inherent risks, it was decided to pursue a safer form of immunization therapy.
   1. The safer therapy is to administer antibodies directed against Aβ that were produced outside of body.
   2. This type of approach is called *passive immunization* since the patient does not produce the therapeutic antibodies.
   3. Passive immunization with an anti-Aβ42 antibody (bapineuzumab) had already proven capable of restoring memory function in transgenic mice.
      a. It was quickly shown to be safe and apparently effective, in Phase I and II clinical trials.

E. The last step before government approval is Phase III trial; it typically employs large numbers of subjects (>1000 at several research centers) and compares effectiveness of new treatment against standard approaches.
   1. The first results of the Phase III trials on bapineuzumab were reported in 2008 and were disappointing.
   2. There was little or no evidence that the antibody provided benefits in preventing disease progression.
   3. Another antibody known as Solanezumab, which recognizes a different part of the Aβ protein, was advanced to Phase III trials despite almost no positive effects from Phase II trials.
      a. All anti-amyloid antibody clinical trials to date have failed, but surprisingly failed Phase II trials have been advanced to Phase III trials.

F. With AD's human health impact and the large amount of money that could be earned from such a drug, drug companies are willing to take a risk that one of these immunologic strategies shows some therapeutic value even when Phase II trials do not show a beneficial effect.

IX. Meanwhile a comprehensive analysis of some of the patients who were vaccinated with Aβ42 in the original 2001 immunization trial was also reported in 2008.

A. Analysis of this patient group showed that Aβ42 vaccination had had no effect on preventing disease progression.
   1. It was particularly striking that in several of these patients who had died of severe dementia, there were virtually no amyloid plaques left in their brains.
   2. This suggested strongly that removal of amyloid deposits in a patient already suffering the symptoms of mild-to-moderate dementia does not stop disease progression.

B. These results can be interpreted in more than one way.
   1. The amyloid deposits are not the cause of the dementia symptoms.

2. The irreversible toxic effects of deposits had already occurred by the time immunization had begun and it was too late to reverse the disease course using treatments that remove existing amyloid deposits.
   a. It is important to note that amyloid deposit formation in the brain starts ≥10 before any clinical symptoms of AD are reported.
   b. It is possible that if these treatments had started earlier, the disease symptoms might never have appeared.

C. The first clinical trial of this type (a preventive trial) was begun in 2012.
1. Several hundred individuals who would normally have been destined to develop early-onset AD (due to mutations in the *PSEN1* gene) were treated with an anti-Aβ antibody.
   a. The hope was that the vaccine would block the future buildup of amyloid and prevent disease.
2. Additional trials using larger numbers of patients who do not carry any known AD predisposing mutations have also been started, using the new generation antibody Solanezumab.
3. Preventive trials that test the ability of treatment to slow the onset of a disease that normally can take decades to manifest itself naturally take a long time to carry out and so the results are not yet known.

D. Clearly the preventive treatments make the most sense for patients who might be on the verge of developing disease.
1. Recent advances in brain-imaging procedures now allow clinicians to observe amyloid deposits in the brains of individuals long before any symptoms of AD have developed.
2. Based on these studies, it may be possible to begin preventive treatments in persons who are at very high risk of developing AD before they develop symptoms.

X. Drugs have also been developed that follow the other 2 strategies outlined previously.

A. Alzhemed and scyllo-inositol are 2 small molecules that bind to Aβ peptides and block molecular aggregation and fibril formation.
1. Clinical trials have failed to demonstrate that either drug is effective in stopping disease progression in patients with mild to moderate AD.

B. The third strategy outlined above is to stop production of Aβ peptides.
1. This can be done by inhibiting either β- or γ-secretase, because both enzymes are required in pathway that cleaves the APP precursor to release the internal peptide.
2. Drug companies have had great difficulty developing a β-secretase inhibitor that is both potent and small enough to enter the brain from the bloodstram.
   a. However, one such inhibitor, known as MK-8931, failed to show benefits in Phase III clinical trials.
3. A number of potent γ-secretase inhibitors have been developed that block the formation of all Aβ peptides, both in cultured nerve cells and in transgenic AD mice.
   a. But there is a biological problem that has to be overcome with this class of inhibitor.
   b. In addition to cleaving APP, γ-secretase activity is also required in a key signaling pathway involving a protein called Notch.
   c. Two of the most promising γ-secretase, flurizan and semagacestat, have both failed to show any benefit in stopping AD progression.
   d. In addition to its lack of efficacy, semagacestat caused adverse side effects that were probably a result of blockade of the Notch pathway.
   e. The goal of drug designers is to develop a compound (e.g., begacestat) that blocks APP cleavage, but does not interfere with cleavage of Notch.

C. Taken collectively, apparent failure of all these drugs, aimed at various steps in formation of Aβ-containing aggregates and amyloid deposition, has left the field of AD therapeutics without a clear plan for the future.
   1. Some drug companies are continuing to develop new drugs aimed at blocking amyloid aggregate formation, whereas others are moving in different directions.
   2. These findings also raise the basic question: Is the Aβ peptide even part of the underlying mechanism that leads to AD?

XI. Aβ is not the only misfolded protein protein found in the brains of persons with AD.
   A. Another protein called tau, which functions as part of a nerve cell's cytoskeleton can develop into bundles of tangled cellular filaments called neurofibrillary tangles (or NFTs).
      1. NFTs interfere with the movement of substances down the length of the nerve cell.
      2. NFTs form when the tau molecules in nerve cells become excessively phosphorylated.
      3. Mutations in the gene that encodes tau have been found to cause a rare form of dementia (called FTD), which is characterized by the formation of NFTs.
   B. Thus, NFTs have been linked to dementia, but they have been largely ignored as a causative factor in AD pathogenesis, due primarily to the fact that the transgenic AD mouse models do not develop NFTs.
   C. If one extrapolates the results of these mouse studies to humans, they suggest that NFTs are not required for the cognitive decline that occurs in AD patients.
   D. However, brain autopsies of humans who died of AD suggest that NFT burden correlates much better with cognitive dysfunction and neuronal loss than does the concentration of amyloid plaques.
   E. Given that mutations in genes in the Aβ pathway are clearly a cause of AD, and yet it is the NFT burden that correlates with cognitive decline, it appears that both Aβ and NFTs must be involved in AD etiology.
      1. Many researchers believe that Aβ deposition somehow leads to NFT formation, an idea known as the "amyloid cascade hypothesis," but the mechanism by which this might occur remains unknown.

XII. It is evident that a great deal of work on AD has been based on transgenic mice carrying human AD genes.
   A. These animals have served as the primary preclinical subjects for testing AD drugs, and they have been used extensively in basic research that aims to understand the disease mechanisms leading to AD development.
   B. Many questions have been raised as to how accurately these animal models mimic the disease in humans.
      1. This is particularly true of the sporadic human cases in which affected individuals lack the mutant genes that cause the animals to develop the corresponding disorder.
      2. In fact, one of the most promising new drugs at the time of this writing is one that acts on NFTs rather than β-amyloid.
         a. In this case, the drug methylthioninium chloride, which dissolves NFTs, was tested on a group of >300 patients with mild to moderate AD in a Phase II trial.
         b. The drug was found to reduce mental decline over a period of 1 year by an average of 81% compared to patients receiving a placebo.
      3. A modified version of the drug, known as leuco methylioninium, was tested in larger Phase III studies for both AD and FTD, but results were not convincing.
         a. A problem arose due to the fact that methylioninium drugs turn urine blue, so placebo controls needed to also have the same effect.
         b. However, in the trials a lower dose of the methylioninium drugs were used, a questionable control that clouds the validity of the studies.

C. Other compounds that inhibit one of the enzymes (GSK-3) that adds phosphate groups to the tau protein are also being investigated as AD therapeutics.
   1. Clinical trials of one GSK-3 inhibitor, valproate, have been stopped due to adverse effects.

XXXVII. **Molecular chaperones** help proteins to fold into final 3D conformation, and are non-specific in their activity.
A. Not all proteins can assume their final 3° structure by a simple process of self-assembly.
   1. This does not occur because the 1° structure lacks the required information for proper folding.
   2. Proteins undergoing folding must be prevented from interacting nonselectively with other molecules in the crowded compartments of the cell.
B. Several protein families have evolved whose function is to help unfolded or misfolded proteins achieve proper 3-D conformation; these "helper proteins" are called molecular chaperones.
   1. They selectively bind to short stretches of hydrophobic amino acids that tend to be exposed in non-native proteins, but buried in proteins having a native conformation.
C. Molecular chaperones are involved in a multitude of activities within cells, ranging from the import of proteins into organelles to the prevention and reversal of protein aggregation in addition to protein folding.
D. Polypeptides are made on ribosomes by adding amino acids, one at a time, starting at the chain's N-terminus.
   1. Chaperones of the Hsp70 family bind to elongating polypeptide chains as they emerge from the exit channel within the large subunit of the ribosome.
   2. Hsp70 chaperones are thought to prevent these partially formed (*nascent*) polypeptides from binding to other proteins in the cytosol that would cause them either to aggregate or misfold.
   3. Once their synthesis has been completed, many of these proteins are simply released by the chaperones into the cytosol where they spontaneously fold into their native state.
      a. Other proteins are repeatedly bound and released by chaperones until they finally reach their fully folded state.
E. Many of the larger polypeptides are transferred from Hsp70 proteins to a different type of chaperone called a *chaperonin.*
   1. Chaperonins are cylindrical protein complexes that contain chambers in which newly synthesized polypeptides can fold without interference from other macromolecules in the cell.
   2. TRiC is a chaperonin thought to assist in the folding of up to 15% of the polypeptides synthesized in mammalian cells.

EXPERIMENTAL PATHWAYS:
CHAPERONES – HELPING PROTEINS REACH THEIR PROPER FOLDED STATE

I. F. M. Ritossa (1962) an Italian biologist studying development of *Drosophila* reported a curious finding.
A. When the temperature at which fruit fly larvae were developing was raised from normal 25°C to 32°C, a number of new sites on the giant chromosomes of the larval cells became activated.
   1. The giant chromosomes of these insect larvae provide a visual exhibit of gene expression.

2. Results suggested that elevated temperature induced new gene expression; confirmed a decade later with the characterization of several proteins that appeared in larvae after temperature elevation.

B. It was soon found that this response, the **heat shock response** was not confined to fruit flies, but can be initiated in many different cells from virtually every type of organism (bacteria to plants and mammals).
   1. More study showed that the *heat-shock proteins* (*hsps*) produced during response were found not only in heat-shocked cells, but also at lower concentration in cells under normal conditions.
   2. Their function was at first unknown, but was gradually revealed by a series of seemingly unrelated studies.

II. Multisubunit structures (bacterial ribosome, tobacco mosaic virus particle) can self-assemble from purified subunits.
   A. 1960s – it was demonstrated that proteins that make up bacteriophage particles also possess a remarkable ability to self-assemble, but can't form complete, functional virus particles by themselves *in vitro.*
      1. Phage assembly experiments in bacteria cells showed that phages require bacterial help.
   B. 1973 – found that a certain mutant bacteria strain (*GroE*) did not support assembly of normal phages.
      1. Depending on the type of phage, the phage particle head or tail was assembled incorrectly.
      2. Studies suggested that a protein encoded by the bacterial chromosome participated in viral assembly, even though this host protein was not a component of the final virus particles.
      3. Since it did not evolve as an aid for viral assembly, the bacterial protein needed for phage assembly had to play some role in the bacterial cell's normal activities, but its precise role was obscure.
   C. Later studies – the *GroE* site on the bacterial chromosome actually contains 2 separate genes, *GroEL* and *GroES* that encode 2 separate proteins, GroEL and GroES.
      1. Under EM, the purified GroEL protein appeared as a cylindrical assembly consisting of 2 disks.
      2. Each disk was composed of 7 subunits arranged symmetrically around the central axis.
   D. Several years later, a study on pea plants hinted at the existence of a similar assembly-promoting protein in the plant's chloroplasts.
      1. Rubisco is a large chloroplast protein that catalyzes the reaction in which CO_2 molecules taken up from the atmosphere are covalently linked to organic molecules during photosynthesis.
      2. Rubisco comprises 16 subunits: 8 small subunits (molecular mass of 14,000 daltons) and 8 large subunits (55,000 daltons).
      3. It was found that large Rubisco subunits, synthesized in chloroplast, are not present in an independent state, but associated with a huge protein assembly consisting of identical 60,000 dalton (60 kDa) subunits.
         a. Researchers considered the possibility that the complex formed by the large Rubisco subunits and the 60-kDa-polypeptides was an intermediate in the assembly of a complete Rubisco molecule.
   E. A separate study on mammalian cells revealed the existence of proteins that appeared to assist assembly of multisubunit proteins.
      1. Like Rubisco, antibody molecules consist of a complex of 2 different types of subunits, smaller light chains and larger heavy chains.
      2. Like large Rubisco subunits, heavy chains of the antibody complex become associated with another protein not found in the final complex.
         a. This protein, which associates with newly synthesized heavy chains, but not with heavy chains that are already bound to light chains, was named *binding protein* (BiP).
         b. BiP was subsequently found to have a molecular mass of 70,000 daltons (70 kDa).

F. In 1986, the 2 lines of study (heat-shock response and proteins that promote protein assembly) came together.
   1. It was shown that a very prominent heat shock response protein, *heat-shock protein 70* (*hsp70*) because of its molecular mass, was identical to BiP, the protein implicated in antibody molecule assembly.

III. Even before the heat-shock response was known, protein structure was known to be sensitive to temperature
   A. It was known that a small rise in temperature could cause these delicate proteins to begin to unfold.
      1. Unfolding exposes hydrophobic residues previously buried in the protein's core.
      2. Hydrophobic residue patches on protein surfaces attract each other like fat droplets in a bowl of soup.
      3. Thus, when a cell is heat shocked, soluble proteins become denatured and form aggregates.
   B. A 1985 report showed that, after temperature elevation, newly synthesized hsp70 molecules enter cell nuclei and bind to nuclear protein aggregates.
      1. They then act like molecular crowbars to promote disaggregation.
      2. Due to their role in assisting protein assembly by preventing undesirable interactions, hsp70 and related proteins were named **molecular chaperones**.

IV. Soon it was demonstrated that the bacterial heat-shock protein GroEL and the Rubisco assembly proteins in plants were homologous proteins.
   A. The 2 proteins share the same amino acids at nearly half of >500 residues in their respective molecules.
      1. The fact that the 2 proteins (both members of *Hsp60 chaperone family*) have retained so many of the same amino acids reflects their similar and essential function in the 2 cell types.
   B. At this point, it was thought that their primary function was to mediate the assembly of multisubunit complexes like Rubisco.
   C. This view was changed in 1989 by experiments studying molecular chaperones in mitochondria – Arthur Horwich (Yale U.) and F.–Ulrich Hartl, Walter Neupert et al. (Univ. of Munich).
      1. Newly made mitochondrial proteins made in the cytosol have to cross the outer mitochondrial membranes in an unfolded, extended, monomeric form.
      2. A mutant was found that altered the activity of another member of the Hsp60 chaperone family that resided inside mitochondria.
      3. In cells containing this mutant chaperone, proteins that were transported into mitochondria failed to fold into their active forms.
      4. Even proteins consisting of a single polypeptide chain failed to fold into their native conformation.
      5. This finding changed the perception of chaperone function from a notion that they assist assembly of already-folded subunits into larger complexes.
         a. The idea that they assist polypeptide chain folding within the crowded confines of the cell.

V. Such results indicated the presence in cells of at least 2 major molecular chaperone families: the Hsp70 chaperones (BiP) and Hsp60 chaperones (*chaperonins*; Hsp60, GroEL [best understood], Rubisco assembly protein).
   A. 1979 – GroEL was shown to be huge molecular complex of 14 polypeptide subunits arranged in 2 stacked rings, resembling a double doughnut.
   B. 15 years later – the 3D structure of the GroEL complex was determined by X-ray crystallography; it revealed the presence of a central cavity within GroEL cylinder.
      1. Later studies showed that the cavity was divided into 2 separate chambers.

2. Each chamber was situated within the center of one of the GroEL complex rings and was large enough to enclose a polypeptide undergoing the fold process.

C. EM studies also provided information about the structure and function of a second protein, GroES, which acts in conjunction with GroEL.
   1. GroES, like GroEL, is a ring-like protein with 7 subunits arrayed symmetrically around a central axis.
   2. GroES, however, consists of only one ring and its subunits are much smaller (10.000 daltons) than those of GroEL (60,000 daltons).
   3. GroES is seen as a cap or dome that fits on top of either end of a GroEL cylinder.
   4. Attachment of GroES to one end of GroEL causes a dramatic conformational change in the GroEL protein that markedly increases the volume of the enclosed chamber at that end of the complex.

E. The importance of this conformational change has been revealed in remarkable detail in X-ray crystallographic studies by Arthur Horwich and Paul Sigler (Yale U.).
   1. Binding of the GroES cap is accompanied by a 60° rotation of the apical domain of the subunits that make up the GroEL ring at that end of the GroEL cylinder.
   2. GroES attachment does more than trigger a conformational change that enlarges the GroEL chamber.
      a. Before GroES attachment, the inner wall of the GroEL chamber has exposed hydrophobic residues that give the lining a hydrophobic character.
      b. Nonnative polypeptides also have exposed hydrophobic residues that become buried in the interior of the native polypeptide.
   3. Since hydrophobic surfaces tend to interact, the hydrophobic lining of the GroEl cavity binds to the surface of nonnative polypeptides.
   4. Binding of GroES to GroEL buries the hydrophobic residues of the GroEL wall and exposes a number of polar residues, thereby changing the chamber wall character.
   5. Thus, a nonnative polypeptide that had been bound to the GroEL wall by hydrophobic interactions is displaced into the space within the chamber.
   6. Once freed from its attachment to the chamber wall, the polypeptide is given the opportunity to continue its folding in a protected environment.
   7. After ~15 seconds, the GroES cap dissociates from the GroEL ring, and the polypeptide is ejected from the chamber.
   8. If the polypeptide has not reached its native conformation by the time it is ejected, it can rebind to the same or another GroEL, and the process is repeated.

F. Approximately 250 of the roughly 2400 proteins present in the cytosol of an *E. coli* cell normally interact with GroEL, so how does GroEL bind to so many different polypeptides?
   1. GroEL's binding site consists of a hydrophobic surface formed largely by 2 α-helices of the apical domain.
   2. This surface is capable of binding virtually any sequence of hydrophobic residues that might be accessible in a partially folded or misfolded polypeptide.
   3. A crystal structure comparison of the unbound GroEL molecule with that of GroEL bound to several different peptides was performed.
      a. It revealed that the binding site on a GroEL subunit apical domain can locally adjust its positioning when bound to different partners.
      b. This finding indicates that the binding site has structural flexibility that allows it to adjust its shape to fit the shape of the particular polypeptide with which it has to interact.

VI. A number of studies have also suggested that GroEL does more than simply provide a passive chamber in which proteins can fold without outside interference.
   A. In one study, site-directed mutagenesis was utilized to modify a key residue, Tyr71 of GroES, whose side chain hangs from the ceiling of the folding chamber.
      1. Due to its aromatic ring, tyrosine is a modestly hydrophobic residue.
      2. If Tyr71 replaced by a "+" or "-" charged amino acid, the resulting GroEL-GroES variant exhibited a higher ability to help the folding of a specific foreign polypeptide, the green fluorescent protein (GFP).
      3. But substitutions for Tyr71 that improved GroES-GroEL's ability to increase GFP folding made the chaperonin less competent to help its natural substrates fold.
      4. Thus, as the chaperonin became more and more specialized to interact with GFP, it lost its general ability to assist the folding of proteins having an unrelated structure.
      5. This finding suggests that individual amino acids in the folding chamber wall may participate somehow in the folding reaction.
      6. Thus, chaperonins may do more than simply provide a passive chamber in which proteins can fold without outside interference.
   B. Data from another study has suggested that binding of a nonnative protein to GroEL is followed by a forced unfolding of the substrate protein.
      1. FRET (fluorescence resonance energy transfer) is a technique that allows researchers to determine the distance between different parts of a protein molecule at different times during a given process.
      2. In this study, investigators found that the protein undergoing folding, in this case Rubisco, bound to the apical domain of the GroEL ring in a relatively compact state.
      3. The compact nature of the bound protein was revealed by the close proximity to one another of the FRET tags, which were attached to amino acids located at opposite ends of the Rubisco chain.
      4. Then, during the conformational change that enlarges the GroEL cavity volume, the bound Rubisco protein was forcibly unfolded, as evidenced by the increased distance between Rubisco's 2 tagged ends.
      5. This study suggests that the Rubisco polypeptide is taken completely back to the unfolded state, where it is given the opportunity to refold from scratch.
         a. This action should help prevent the nonnative protein from becoming trapped permanently in a misfolded state.
         b. In other words, each individual visit to a GroEL-GroES chamber provides an all-or-none attempt to reach the native state.
         c. Each visit to the chamber does not provide a chance to finish just one stage in a series of steps in which the protein moves closer to the native state with each round of folding.
   C. Molecular chaperones do not convey information for the folding process, but instead prevent proteins from veering off their correct folding pathway and finding themselves in misfolded or aggregated states.
      1. As Anfinsen discovered decades ago, a protein's 3D structure is determined by its amino acid sequence.

XXXVIII. Genome sequencing has gotten a lot of attention in recent years, and it is easy to forget that genes are primarily information storage units, while proteins actually orchestrate cellular activities.
   A. Genome sequences provides a kind of "parts list."
   B. Human genome probably contains between 20,000 and 22,000 genes, each of which can potentially give rise to a variety of different proteins, only a fraction of which have been characterized to date.

1. One gene can give rise to >1 polypeptide in several ways, but there are two prominent mechanisms.
   a. Alternative splicing.
   b. Posttranslational modification.
2. It should also be noted that many proteins have more than one distinct function.
   a. Even myoglobin, long studied as an oxygen-storage protein, has recently been shown to be involved in the conversion of nitric oxide (NO) to nitrate (NO_3^-).

XXXIX. The **proteome** is the entire inventory of proteins produced by an organism.

A. The term *proteome* is also applied to the inventory of all proteins that are present in a particular tissue, cell or cellular organelle.
B. Because the number of proteins currently being studied, researchers have sought to develop techniques that allow them to determine the properties/activities of a large number of proteins in a single experiment.
C. The term **proteomics** was coined to describe the expanding field of protein biochemistry.
   1. This term carries with it the concept that advanced technologies and high speed computers are being used to perform large-scale studies on diverse arrays of protein, the same approach used lately on genomes.
   2. The study of proteomics is inherently more difficult than the study of genomics because proteins are more difficult to work with than DNA.
      a. In physical terms, one gene is pretty much the same as all other genes, while each protein has unique chemical properties and handling requirements.
      b. Small quantities of a particular DNA segment can be expanded greatly using readily available enzymes, while protein quantities cannot be increased.
      c. This is particularly troublesome since many proteins regulating important cell processes are present in only a handful of copies per cell.
D. Traditionally, protein biochemists have sought to answer a number of questions about particular proteins.
   1. What specific activity does the protein demonstrate *in vitro*, and how does this activity help a cell carry out a particular function such as cell locomotion or DNA replication?
   2. What is the protein's 3D structure?
   3. When does the protein appear in the development of the organism and in which types of cells?
   4. Where in the cell is it localized?
   5. Is the protein modified after synthesis by the addition of chemical groups (e.g., phosphates or sugars) and, if so, how does this modify its activity?
   6. How much of the protein is present, and how long does it survive before being degraded?
   7. Does the level of the protein change during physiologic activities or as the result of disease?
   8. Which other proteins in the cell does it interact with?
E. Biologists have been trying to answer these questions for decades but, mostly, one protein at a time.
   1. Proteomics researchers try to answer similar questions on a more comprehensive scale using large-scale or *high-throughput*, techniques to catalog the vast array of proteins produced by a particular cell.
   2. Trying to separate and identify the large numbers of proteins made in cells and tissues of complex tissues and answer a host of difficult questions about each protein.

XXXX. The key technology in proteomics is mass spectrometry, a method to probe the chemical structure of an unknown sample.

A. Mass spectrometry is a technique to determine the precise mass of a molecule or fragment of a molecule, which can then be used to identify that molecule.
B. Suppose that you want to identify an unknown protein you have in a test tube.
    1. The protein is first digested into peptides with the enzyme trypsin.
    2. When these peptides are introduced into a mass spectrometer, they are converted into gaseous ions and separated according to their mass/charge (m/z) ratio.
    3. The results are displayed as a series of peaks of known m/z ratio.
    4. The pattern of peaks constitutes a highly characteristic *peptide mass fingerprint* of that protein, and then genomics can then be used to identify the protein based on the peptide mass fingerprint.
    5. Once a genome has been sequenced, the amino acid sequences of encoded proteins can be predicted.
    6. The list of "virtual proteins" can then be subjected to a theoretical trypsin digestion and the masses of the resulting virtual peptides calculated and entered into a database.
    7. The actual peptide masses of a purified protein obtained by the mass spectrometer can be compared using a computer to masses predicted by theoretical digests of all polypeptides encoded by the genome.
    8. Usually, the isolated protein subjected to mass spectrometry can be directly identified from the database.
C. Mass spectrometers are also capable of analyzing proteins present in complex mixtures.
    1. By analyzing the total protein extracted from a cell by mass spectrometry, it is possible to determine a list of all proteins present in the cell, known as the proteome of the cell.
D. Proteomic studies are particularly useful when 2 different samples are compared to see how the protein composition changes over time.
    1. For example, the proteome can be analyzed before and after the secretion of a hormone within the body, after taking a drug or during a particular disease.
    2. The changes in the abundance of different proteins during such transitions can give valuable clues about how cell function is changing in a given situation.

XXXXI. Proteomics is playing an increasingly important role in advancing the practice of medicine.
A. It is thought that most human diseases leave telltale patterns (*biomarkers)* among the thousands of proteins present in the blood or other bodily fluids.
    1. The simplest way to find whether a particular biomarker protein characteristic of a disease is present in a blood or urine sample and how much is present is to measure its interaction with a specific antibody.
    2. This is the basis of the PSA test used in routine screening of men for prostate cancer.
        a. PSA is a protein that is found in the blood of normal men, but is present at elevated levels in individuals with prostate cancer.
        b. PSA levels are determined by measuring the amount of protein in the blood that binds to anti-PSA antibodies.
B. The most challenging hurdle in developing this type of diagnostic test is knowing what protein will act as the most reliable biomarker, and this is where proteomics comes into play.
    1. Many efforts have been made to compare the proteins present in the blood of healthy individuals with those present in the blood of persons suffering from various diseases, especially cancer.
    2. Initially, the results of these biomarker searches were unreliable in that the findings of one research group could not be duplicated by the efforts of other groups.
    3. The primary difficulty stems from the fact that human blood serum is such a complex solution containing thousands of proteins that range in abundance over 9 – 10 orders of magnitude.

4. But as proteomic and mass spectrometry technology improves, the information about complex samples becomes richer and richer.

C. Already, proteomics has been used to reveal important disease biomarkers.
1. The OVA1 blood test for ovarian cancer, which detects a collection of biomarkers using antibody-based tests, was invented using data from proteomic analysis of a large number of patient samples.
2. The OVA1 test is mainly used for cancers that have already been detected, in order to provide more information about the tumor before surgery.

D. It is hoped that, one day, it will be possible to use a single blood test to reveal the existence of early-stage heart, liver or kidney disease that can be treated before it becomes a life-threatening condition.

XXXXII. Protein separation and mass spectrometric methods tell us nothing about a protein's function, so researchers have been working to devise methods, allowing protein function to be determined on a large scale, not one at a time.

A. Several new technologies have been developed to accomplish this goal, especially the use of genome-wide RNA interference (RNAi) screens.
1. RNAi is a cellular process by which cells produce small RNAs (siRNAs) that bind to specific mRNAs and inhibit the translation of these mRNAs into proteins.
2. Researchers can synthesize a collection (library) of siRNAs that are capable of inhibiting the translation of virtually any mRNA that is produced by a genome.
a. Each mRNA represents the expression of a specific gene encoding a particular protein.
b. Therefore, one can find out which proteins are involved in a particular cellular process by determining which siRNAs interfere with that process.

B. RNAi can be used in large-scale screens of the whole genome and can also be used to test the function of individual genes or sets of genes.

C. The combination of RNAi with proteomics is extremely powerful.
1. Assume that proteomic analysis suggests that the level of a particular protein changes during a process of interest.
2. The expression of that protein can then be blocked by an appropriate RNAi to ask whether its function is relevant for that process.

D. A new method for testing gene function has been developed that uses the CRISPR pathway.
1. Conceptually similar to RNAi since an RNA template targets the sequence of interest.
2. The CRISPR target is the gene within the DNA genome in contrast to the RNA transcript.

XXXXIII. Those who study protein-protein interactions want to know whether the protein they are working with (X) interacts physically with another protein (Y); several techniques can be used.

A. Can study protein-protein interactions on global scale, e.g., one might want to know all of the interactions that occur among the ~6000 proteins encoded by the budding yeast *Saccharomyces cerevisiae* genome.
1. The entire genome of this yeast has been sequenced and virtually every gene within the genome is available as an individual DNA segment that can be cloned and used as desired.
2. Testing for potential interactions between all of these proteins requires methods that can be easily automated and performed by robots.
a. These methods must also be reliable and sensitive enough to detect most of the true interactions without producing too many false interactions in the process.

B. One such widely-adopted method is TAP-tag mass spectrometry.

1. DNA from a gene of interest is fused to DNA encoding a protein tag called a TAP tag that is easily purified using affinity chromatography.
2. This TAP-tagged gene is expressed in a cell, the cell is lysed and the TAP-tagged protein is purified, carrying any interacting proteins along with it.
3. The set of proteins that are co-purified with TAP-tagged protein is then identified by mass spectrometry.
4. This process is repeated for each gene in the genome, eventually producing a map that shows all of the proteins that co-purify with each other and therefore presumably interact inside the cell.
5. This complete set of interactions is called the *interactome* of the cell.

C. The results from large-scale protein-protein interaction studies can be presented in the form of a network.
1. One could display the potential binding partners of the various yeast proteins that contain an SH3 domain and illustrate the complexities of such interations at the level of an entire organism.
2. The proteins that have multiple binding partners are referred to as *hubs* of the protein interaction network.
3. Hub proteins are more likely than non-hub proteins to be essential proteins, that is, proteins that the organism cannot survive without.
4. Some hub proteins have several different binding interfaces and are capable of binding a number of different binding partners at the same time.
5. In contrast, other hubs have a single binding interface, which is capable of binding several different partners, but only one at a time.

D. Overall, it is estimated that, on average, each protein encoded in a eukaryotic organism genome interacts with ~5 different protein partners; thus, human proteinswould engage in ~100,000 different interactions.

E. One must remember when viewing these diagrams that any large-scale, automated experiment will have some errors.
1. For interactome data, some true interactions will be missed because the interaction is too weak to survive the purification process or because the TAP tag disrupts an interacting domain.
2. Likewise, some interactions reported in the analysis may not actually represent true interactions in cell.
   a. For example, if a protein is present as a contaminant when purifying a particular protein complex, it would be incorrectly reported as interacting with the TAP-tagged protein.
   b. It is therefore best to view interactions data sets as showing potential interactions that can serve as a guide for future experiments using more direct means.

F. Genome-sequencing projects have provided scientists with the amino acid sequences of a huge number of proteins whose very existence was previously unknown – what do they do?
1. One approach to determining a protein's function is to identify the proteins with which it associates.
2. If a known protein is shown to be involved in DNA replication, and an unknown protein is found to interact with that protein, then it is likely the unknown protein is part of DNA-replication machinery.

G. Regardless of their limitations, these large-scale protein interaction studies provide the starting point to infer the function of unknown genes based on their possible interaction partners.

XXXXIV. Scientists can now design and mass produce novel proteins that are different from those made by living organisms.

A. Scientists can make an artificial gene to be used in making a protein with any desired amino acid sequence; they can also make polypeptides from scratch in the lab using chemical techniques.
1. This latter strategy allows researchers to incorporate building blocks other than the 20 amino acids that normally occur in nature.

2. The problem with these engineering efforts is in knowing which of the virtually infinite variety of possible proteins one could make might have some useful function.
3. Suppose a pharmaceutical company wanted to manufacture a protein that would bind to the AIDS or influenza virus surface.
4. Assume computer simulations could predict the shape such a protein should have to bind the virus surface; this requires detailed insight into rules governing relationship between 1° and 3° structure.

B. Protein biochemists now have knowledge that allows them to construct a protein capable of binding to the surface of another protein.
1. This protein is the hemagglutinin (HA) protein present in the reconstructed 1918 influenza virus.
2. HA protein is used by the virus to gain entry to a human cell, so that inhibition of HA could, in theory, prevent viral infection.
3. Scientists created an engineered protein that is capable of binding to a hydrophobic patch on the surface of the HA protein with high affinity.
4. Side chains from the designed protein interact in highly specific ways with sites on the α-helix of HA.
5. This peptide was found to inhibit the function of the HA protein, potentially pointing the way to future efforts to developing HA-binding peptides as antiviral drugs.

XXXXV. One might think that designing an enzyme to catalyze a given chemical reaction would be beyond the capability of present-day biotechnology, but it is possible to produce novel proteins.

A. In 2008, researchers reported they had successfully designed and produced artificial proteins that were capable of catalyzing two different organic reactions, neither of which was catalyzed by any known natural enzyme.
1. One involved breaking a carbon-carbon bond and the other the transfer of a proton from a carbon atom.
2. They began by choosing a catalytic mechanism that might accelerate each chosen reaction.
3. They then used computer-based calculations to construct an idealized space in which amino acid side chains were positioned (forming an *active site*) to accomplish the task.
4. They then searched among known protein structures to find ones that might serve as a framework or scaffold that could hold the active site they has designed.
5. To transform the computer models into an actual protein, they used computational techniques to generate DNA sequences that had the potential to encode such a protein.
6. The proposed DNA molecules were synthesized and introduced into bacterial cells where the proteins were manufactured and the catalytic activities of the proteins were then tested.

B. Those proteins that showed the greatest promise were subjected to a process of test-tube evolution.
1. The proteins were mutated to create a new generation of altered proteins, which could, in turn, be screened for enhanced activity.
2. Eventually, they obtained proteins that could accelerate the rates of reaction as much as one million times that of the uncatalyzed reaction.
3. A natural enzyme could do better, but it is a remarkable accomplishment for a team of biochemists.
4. It suggests that scientists will ultimately be able to construct proteins from scratch that will be capable of catalyzing virtually any chemical reaction.

XXXXVI. Another approach to the production of novel proteins is to modify proteins already made by cells.

A. Isolate individual human gene from chromosomes, alter its sequence in precise way and make modified protein with altered amino acid sequence (**site-directed mutagenesis**); use it to study protein structure.

1. If want to know of a particular residue on protein folding or function, the gene can be mutated in a way that substitutes an amino acid with different charge, hydrophobic character or H-bonding properties.
   a. The effect of the substitution on structure and function of the modified protein can then be determined.
2. Site-directed mutagenesis has proven invaluable in analysis of specific functions of minute parts of virtually proteins of interest to biologists.

B. Site-directed mutagenesis is also used to modify structure of clinically useful proteins to bring about various physiological effects; an example:
1. The drug Somavert (approved by FDA in 2003) is a modified version of human growth hormone (GH) containing several alterations.
2. GH normally acts by binding to a target cell surface receptor to trigger a physiological response.
3. Somavert competes with GH in binding to the GH receptor, but interaction between the drug and receptor fails to trigger the cellular response.
4. Somavert is prescribed for the treatment of acromegaly, a disorder that results from excess production of growth hormone.

XXXXVII. Structure-based drug design is used to develop new drugs that act by binding known proteins and inhibiting their activity, and drug companies have access to vast chemical libraries that contain millions of different organic compounds.

A. Expose protein being targeted to combinations of organic compounds made over years by drug companies or isolated from plants or microorganisms (there are millions of them).
1. Determine which ones, if any, happen to bind to the target protein with reasonable affinity.
2. Alternative if protein 3° structure is known then use computers to design virtual drug molecules whose size and shape might allow them to fit into the protein's cracks and crevices, rendering it inactive.
3. Called *structure-based drug design.*

B. Can illustrate both technologies by a look at the drug Gleevec's development (generic name: Imatinib), which has revolutionized treatment of relatively rare cancers like chronic myelogenous leukemia (CML).
1. Tyrosine kinases (TKs) are often involved in the transformation of normal cells into cancer cells.
2. TKs catalyze the addition of phosphate groups to specific tyrosine residues within a target protein that may activate or inhibit the target protein.
3. CML development is driven almost single-handedly by the presence of an overactive TK called ABL.

C. During 1980s, researchers identified a compound, 2-phenylaminopyrimidine, that was able to inhibit TKs.
1. This compound was discovered by randomly screening a large chemical library for compounds that exhibited this particular activity.
2. As is usually the case in blind screening experiments, 2-phenylaminopyrimidine would not have made a very effective drug.
3. First, it was only a weak enzyme inhibitor, which meant it would need to be used in very large quantities.
4. 2-phenylaminopyrimidine is described as a *lead* compound, a starting point, from which usable drugs might be developed.
5. Starting with this lead molecule, compounds of greater potency and specificity were synthesized using structure-based drug design.
6. One compound to emerge from this process was Gleevec; it was found to bind tightly to the inactive form of the ABL TK and prevent it from being activated, which is necessary for cell to become cancerous.

D. Preclinical studies demonstrated that Gleevec strongly inhibited the growth in the lab of cells from CML patients and that the compound showed no harmful effects in animal tests.
   1. In the very first clinical trial of Gleevec, virtually all of the patients went into remission after taking once-daily doses of the compound.
   2. Gleevec has gone on to become the primary drug prescribed for CML treatment, but this is not the end of the story.
   3. Many patients taking Gleevec eventually experience a recurrence of their cancer when the ABL kinase mutates to become resistant to the drug.
   4. In such cases, the cancer can continue to be suppressed by treatment with more recently designed second-generation drugs that are capable of inhibiting Gleevec-resistant forms of the ABL kinase.

XXXXVIII. Adaptations are traits that improve the likelihood that an organism will survive in a particular environment.

A. Proteins are biochemical adaptations that are subject to natural selection and evolutionary change in the same way as other types of characteristics such as eyes and skeletons.
B. This is best revealed by comparing evolutionarily related (*homologous*) proteins in organisms living in very different environments.
   1. Proteins of halophilic (salt-loving) archaebacteria have amino acid substitutions allowing them to maintain their solubility and function at very high cytosolic salt concentrations (up to 4 M KCl).
   2. The surface of the halophilic version of the protein malate dehydrogenase is coated with aspartic and glutamic acid residues, whose carboxyl groups compete with salt for water molecules.
C. Homologous proteins isolated from different organisms can exhibit virtually identical shapes and folding patterns, but show strikingly divergent amino acid sequences.
   1. The greater the evolutionary distance between 2 organisms, the greater is the difference in the amino acid sequences of their proteins.
   2. Sometimes, only a few key amino acids located in a critical portion of the protein is present in all of the organisms from which that protein has been studied.
   3. In one comparison 226 globin sequences, only 2 residues were found to be absolutely conserved in all of these polypeptides (one is a histidine residue that plays a key role in O_2-binding and -release).
   4. All of these observations indicate that 2° and 3° structures of proteins change much more slowly during evolution than their 1° structures.
D. This does not mean that the conformation of a protein cannot be affected be affected in a major way by simple changes in 1° structure, as determined by experimental amino acid substitution in a protein.
   1. An amino acid substitution was experimentally introduced into malate dehydrogenase that completely altered the conformation of a small domain within a large protein molecule.
   2. The original version of protein had a leucine at position 45 and a conformation consisting of a bundle of 3 α-helices.
   3. Altered version has tyrosine at position 45 and a conformation with just 1 α-helix and a 4-stranded β-sheet.
   4. If a mutation having an effect of this magnitude occurred in nature, it might result in the formation of a protein with new functional properties.
      a. It could thus be responsible for generating the ancestral form of an entirely new protein family.

XXXIX. Not only does evolution produce different versions of proteins in different organisms, but it has also produced different versions of proteins in individual organisms.

A. Several different versions of proteins with a given function (e.g. globin, collagen) are encoded by human genome.
   1. Usually, different versions of a protein (known as **isoforms**) are adapted to function in different tissues or at different stages of development.
   2. Humans have 6 different genes encoding isoforms of the contractile protein actin; 2 are found in smooth muscle, 1 in heart muscle and 2 in virtually all other cell types.

B. Now that a large number of amino acid sequences and 3° structures have been reported, it is clear that most proteins are members of much larger **families** (or *superfamilies*) of related molecules.
   1. Genes encoding the various members of a protein family are thought to have arisen from a single ancestral gene that underwent a series of duplications during the course of evolution.
   2. Over long periods of time, the nucleotide sequences of the various copies diverge from one another to generate proteins with related structures.
   3. Many protein families contain a remarkable variety of proteins that have evolved diverse functions.
   4. The expansion of protein families is responsible for much of the protein diversity encoded in the genomes of today's complex plants and animals.

L. Nucleic acids are primarily involved in the storage and transmission of genetic information; may also play structural or catalytic roles.

A. They are macromolecules constructed as a long chain (strand) of monomers called **nucleotides**.

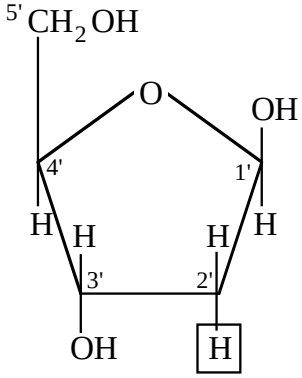
B. Two types found in living organisms: **deoxyribonucleic acid** (**DNA**) and **ribonucleic acid** (**RNA**).
   1. DNA serves as the genetic material of all cellular organisms; RNA plays that role in many viruses.
   2. Information stored in DNA is used to govern cell activities through the formation of RNA messages.

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

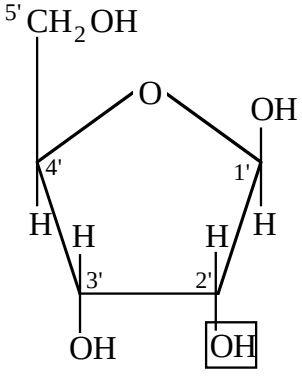
LI. Both DNA and RNA are composed of nucleotides connected to form polymers or polynucleotides, and the nucleotides consist of 3 parts (phosphate + sugar + base).

A. Phosphate group (PO_4^-) is linked to the 5'-carbon of sugar.

B. 5-carbon sugar is either ribose for RNA or deoxyribose for DNA.



Deoxyribose

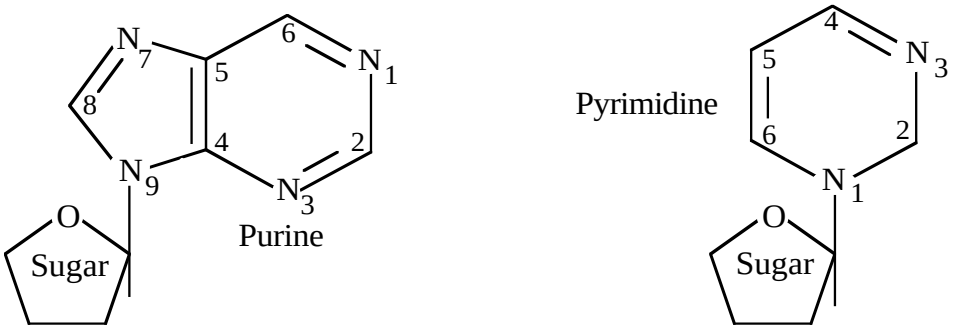


Ribose

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

1. The phosphate is normally linked to the 5'-carbon of the sugar.
2. The nitrogenous base is attached to the sugar's 1'-carbon.

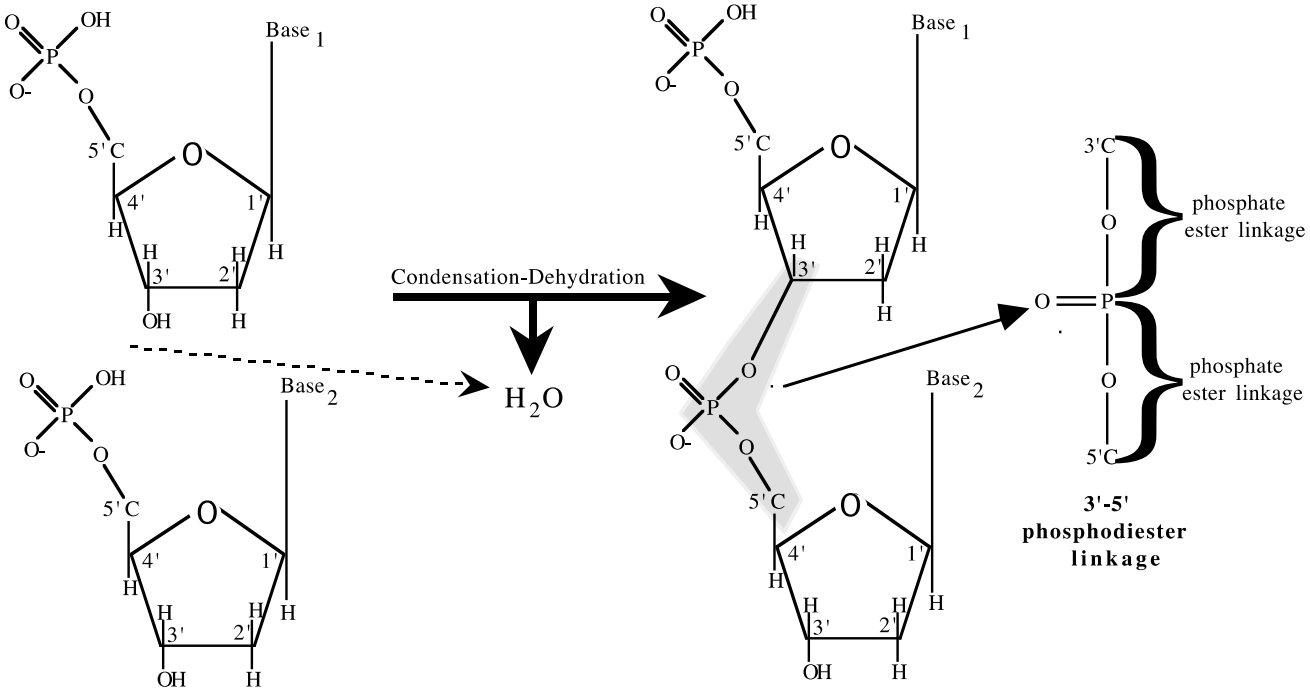
Generalized Stucture of Purines and Pyrimidines



LII. Monomers polymerize when one sugar's 3'-OH is linked by ester bond to 5'-phosphate of the next nucleotide in the chain; thus, nucleotides (RNA or DNA) are joined by sugar-phosphate linkages.

A. Phosphate groups in backbone are attached to 2 sugars by ester linkages.

1. The linkage is called a *3'-5' phosphodiester bond* because the phosphorus atom is esterified to 2 oxygen atoms, one from each of the 2 adjoining sugars.



LIII. A strand of RNA (or DNA) contains 4 different types of nucleotides distinguished by their nitrogenous base, divided into 2 types of nitrogenous bases: **pyrimidines** and **purines**.

A. **Purines** in RNA and DNA are **adenine** and **guanine**; they are larger structures, consisting of 2 rings.

B. **Pyrimidines** in RNA are **cytosine** and **uracil**; in DNA, **uracil** is replaced by **thymine**, a pyrimidine with an extra methyl group attached to the ring; they are smaller, single ring structures.

LIV. RNAs consist of a continuous single strand, but they often fold back on themselves to produce molecules with extensive double-stranded segments with local H-bond pairing (just like in DNA) and complex 3° structures.
- A. With DNA, it directs and carries out protein synthesis.
- B. Ribosomal RNAs (rRNAs) are not molecules that carry genetic information; rather, they serve as structural scaffolds on which the proteins of the ribosome can be attached.
  1. They also serve as elements that recognize and bind various soluble components required for protein synthesis.
  2. One of the rRNAs of the large ribosomal subunit acts as the catalyst for the reaction by which amino acids are covalently joined during protein synthesis.
- C. RNAs having a catalytic role (like rRNA of large ribosomal subunit) are called RNA enzymes or **ribozymes**; the hammerhead ribozyme is able to cleave its own RNA strand.
- D. The double-stranded regions of these RNAs are held together by H bonds between the bases; the same principle is responsible for holding together the 2 strands of the DNA molecule.

LV. Nucleotides are not only important as the building blocks of nucleic acids; they also have important functions in their own right.
- A. Most of the energy being put to use at any given moment in any living organism is derived from the nucleotide **adenosine triphosphate (ATP)**.
- B. **Guanosine triphosphate (GTP)** is another molecule of enormous importance in cellular activities.
  1. GTP binds to a variety of proteins (called G proteins) and acts as a switch to turn on their activities.

(2.8) The Formation of Complex Macromolecular Structures

I. The assembly of cellular organelles is poorly understood, but it is apparent that different types of subunits can self-assemble to form higher-order arrangements.

II. The assembly of tobacco mosaic virus (TMV) particles is an example of the most convincing trype of evidence that an assembly process can be self-directed.
- A. The best such evidence is the demonstration that assembly can occur outside of the cell (*in vitro*) under physiological conditions when the only macromolecules present are those that make up the final structure.
- B. Heinz Fraenkel-Conrat and Robley Williams, (University of California-Berkeley, 1955) demonstrated that TMV particles are capable of self-assembly.
  1. TMV consists of one long RNA molecule (~6600 nucleotides) wound within a helical capsule made of 2130 identical protein subunits.
  2. They purified TMV RNA and protein separately and mixed them together under suitable conditions, and they recovered mature, infective particles after a short period of incubation.
- C. The proteins and RNA of TMV contain all the information necessary for particle formation.

III. The assembly of ribosomal subunits, like TMV ribosomal subunits, are made of RNA and protein.
- A. Unlike the simpler TMV, ribosomes contain several different types of RNA and a considerable collection of different proteins; all ribosomes, regardless of their source, are composed of 2 subunits of different size.
- B. Although ribosomal subunits are often depicted as symmetric structures, in fact, they have a highly irregular shape.
  1. The large 50S subunit of bacteria contains 2 RNA molecules and ~32 different proteins.

2. The small 30S ribosomal contains 1 RNA molecule and 21 different proteins.

C. Masayasu Nomura et al., Wisconsin (mid-1960s) reconstituted complete, fully functional 30S subunits, he mixed 21 purified proteins of the small subunit with purified small-subunit rRNA from *E. coli.*
   1. The small subunit components contain all of the information necessary for assembly of the entire particle.
   2. Analysis of the intermediates that form at different stages during reconstitution *in vitro* indicates that subunit assembly occurs in a sequential step-by-step manner that closely parallels the process *in vivo.*

D. The step-by-step process of small ribosomal subunit assembly as deduced by Nomura et al. - incorporation of individual proteins alters the growing particle, making it more receptive to attachment of new proteins.
   1. At least one small subunit protein (S16) seems to function only in ribosome assembly; its deletion from the reconstitution mixture greatly slowed assembly, but did not block fully functional ribosome formation.
      a. Many other small subunit proteins function primarily to stabilize assembled structure.
   2. Ribosomal reconstitution process takes up to 2 hrs at 50°C *in vitro,* but a few minutes in a bacterium at temperatures as low as 10°C.
      a. It may be that the bacterium uses something that is not available to the investigator who begins with purified components.
      b. Assembly of the ribosome within the cell may include participation of accessory factors, like chaperones, that function in protein folding that the investigators did not have.
   3. In fact, the formation of ribosomes within a *eukaryotic* cell requires transient association with many proteins that do not end up in the final particle.
      a. Also about half of the nucleotides of the large ribosomal RNA precursor are removed.
      b. As a result, the components of the mature eukaryotic ribosome no longer possesses the information to reconstitute themselves *in vitro.*

E. The large ribosomal subunits of bacteria were also reconstituted in next decade.

IV. Larger-scale assemblies may rely on the principles of phase separation and generation of phase-separated compartments.

A. Liquid–liquid phase separation results in droplet formation when separating two mutually insoluble liquids, like a shaken salad dressing that has a mixture of vinegar droplets floating in oil.
   1. Phase separation occurs when one phase is hydrophobic (e.g. oil) and one is hydrophilic (e.g. water or vinegar).

B. Phase separation can occur between water soluble polymers dissolved in water, with high concentrations of one polymer but not the other, in a phenomenon called *aqueous phase separation*.

C. Aqueous phase separation occurs in cells, as RNA and interacting proteins form liquid phase-separated droplets in the cytoplasm or nucleus.
   1. Being liquids, these droplets can fuse and split apart.
   2. Proteins and regions forming phase-separated compartments include nucleoli; many RNA splicing factors; Whi3, involved in cell polarization; and FUS, an RNA-binding protein linked to neurodegeneration.

V. Many proteins that undergo liquid phase separation share two features: intrinsic disordering and RNA binding.

A. Protein domains can be disordered and lack a defined three-dimensional structure, and these *intrinsically disordered domains* are a common for proteins found in these droplets.

B. The flexibility of these domains make the proteins similar to soluble polymers known to undergo aqueous phase separation.
C. Weak interactions between proteins and RNA may produce a liquid-like state where molecules rearrange rapidly as a protein releases one RNA molecule then binds another.

VI. Liquid phase-separated compartments in the cell may drive interactions between RNA and proteins.
A. In liquid phase separation, polymers become more concentrated in their own phase after separation.
1. Proteins and RNA concentrations would be higher in the phase-separated droplet than predicted for a normal distribution throughout the whole cell, thus promoting molecular interactions.
B. Droplets may have a functional role, as their microscopic visualization helps predict where a cell will undergo polarized growth.
C. To show phase separation is necessary and sufficient for function requires elimination and then restoration of interactions, generated by mutation or deletion of disordered domains of a phase-separating proteins.
1. Perturbations can block phase separation and often lead to alterations in function.
2. A regulator of RNA splicing, Rbfox, assembles into droplets through an intrinsically disordered domain, the deletion of which blocks phase separation and prevents proper function in splicing.
3. However, the gold standard to restore phase separation using a different disordered domain has not yet been met, raising persistent questions about the biological relevance of phase separation.

VII. An alternative to liquid phase separation is gelation, in which the proteins and RNA form an elastic gel.
A. A liquid (e.g. maple syrup) doesn't compress when a force is applied, whereas a gel (e.g. gelatin) may deform then reform after the force is removed.
B. When interactions are strong enough, a phase-separated compartment will behave like a gel and is less likely to appear round, although the biological importance is whether a protein forms a liquid versus a gel phase is uncertain.

LECTURE HINTS

Basic Biological Chemistry

Some students of cell biology are distressed to discover that an understanding of cell biology requires a command of (at least) some basic chemical principles. This distress most often stems from a fear (oftentimes a long-standing one) of chemistry that can make these students refractory to the information you are trying to convey. Thus, it is important to present these concepts in a less threatening manner so that the students will have a better chance of understanding the principles. If the students begin to grasp the concepts, they will, in most cases, gain confidence and eventually lose at least some of the fear that is hampering their learning.

Chemical Bonding

The first of the principles that students of cell biology must master is that of chemical bonding. It has been useful to explain ionic bonds and covalent bonds in the following way. Chemical bonds arise as a way of filling the outer electron shells of atoms participating in the bond. Full outer shells confer enhanced stability on the atoms. Ionic bonds occur as a result of *valence electron transfer* from one atom to another, e.g., from Na to Cl (as in table salt), creating Na^+ and Cl^- ions that attract one another. On the other hand, covalent bonds involve a *sharing of valence electrons*. There are two types of covalent bonds: polar and nonpolar covalent bonds. Polar covalent bonds result from an unequal sharing of electrons that causes the more electronegative atom, like the oxygen atom in water, to acquire a partial negative charge since the electrons spend more time in its vicinity. At the same time, the other atom acquires a partial positive charge since this same electron is not spending as much time in its vicinity (the H atoms of water). Nonpolar covalent bonds involve an equal sharing of electrons as in molecular hydrogen and CH_4 (methane). Consequently, there is no separation of charge across this bond and the bond is nonpolar. This situation can be illustrated as follows:

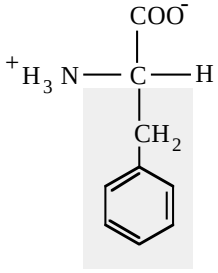
complete e^- transfer	unequal sharing of e^-	equal sharing of e^-
ionic bond	**polar covalent**	**nonpolar covalent**
full charge	partial charge	no charge

The formation of these bonds can be considered to lie along a continuum from complete electron transfer (ionic bonds) to an equal sharing of electrons (nonpolar covalent bonds). One might well think of unequal sharing of electrons as an incomplete transfer of electrons and equal sharing of electrons as essentially no transfer of electrons. It depends on your viewpoint.

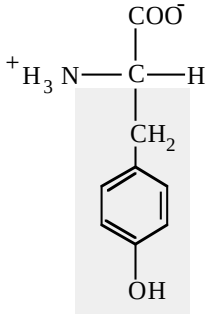
Hydrophobic vs. Hydrophilic

Perhaps the most difficult concept for students to grasp has been the difference between hydrophilic and hydrophobic molecules and how these molecules or portions of molecules may be recognized. I have attempted more complex chemical arguments that have met with limited success. However, for some students, simplifying the definition does relatively little to improve matters. You could tell students to consider straight carbon chains or ring structures without attached hydrophilic groups (hydroxyls, sulfhydryl groups, amino groups, carboxyl groups, etc.) as earmarks of hydrophobic molecules or sections of molecules. Using the amino acids phenylalanine and tyrosine as examples, emphasize the effect of one hydroxyl group on a largely hydrophobic

structure. Phenylalanine and tyrosine are considered to be, respectively, hydrophobic and polar, uncharged amino acids by virtue of their R group chemistry (the shaded area in the drawing below).



Phenylalanine
Nonpolar Amino Acid



Tyrosine
Polar, Uncharged
Amino Acid

The R groups of both amino acids contain the same typically hydrophobic carbon-containing ring. However, tyrosine contains, in addition, a single hydrophilic hydroxyl group that alters its properties enough to justify its classification as a polar, uncharged amino acid. By applying some simple rules, a student can determine whether a molecule has hydrophobic and/or hydrophilic properties. The presence of a carbon-containing ring or chain of carbons without characteristically hydrophilic groups indicates a hydrophobic molecule. The most common hydrophilic groups in living cells are amino, carboxyl, hydroxyl, and sulfhydryl groups. A molecule possessing both hydrophilic and hydrophobic substituents is said to be amphipathic (*amph* - meaning on both sides).

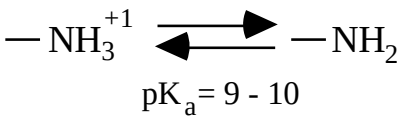
Demonstration

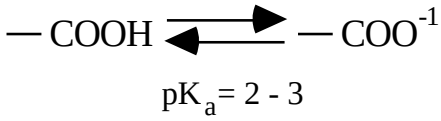
SOLUBILITY TOYS

I demonstrate the differential solubility of hydrophobic and hydrophilic materials by using souvenirs that invariably show up in gift shops at the beach. The simplest of these toys consists of a hollow plastic cube containing two liquids with a plastic surfer that manages to ride the waves at the interface between the two fluids. The lower fluid is blue due to the presence of a blue dye that is undoubtedly a hydrophobic molecule soluble in other hydrophobic materials, but insoluble in water. Consequently, no matter how hard one shakes the cube, the blue dye will never leech out into the water that makes up the upper liquid in the cube. You can shake the cube as vigorously as you wish, but the oil and the dye along with it will separate from the water above after a short time. There are more elaborate toys like this on the market, some of which appear to contain both hydrophilic and hydrophobic dyes. For example, the Colorfall by the Carlisle Company of Carson City, Nevada (702-246-7822; PO Box 21029, Carson City NV 89721) contains water colored with a hydrophilic yellow dye. Two chambers within the device contain mineral oil and the dyed water. The mineral oil in the two chambers contains dyes of different colors. The device clearly demonstrates the tendency of oil and water to remain separate and the solubility of hydrophobic substances in other hydrophobic substances. The dyes do not leech into materials of dissimilar chemistry. Students can also observe the fusion of drops of oil as they encounter one another in both toys, although the Colorfall is better suited to this purpose. This can aid the students in understanding how membranes can fuse during exocytosis. Either type of toy serves as an effective demonstration of the differential solubilities of hydrophobic and hydrophilic substances.

Acids, Bases, Buffers and pH

Another aspect of chemistry that is likely to cause problems is the comprehension of acids and bases. An acid is traditionally described as a molecule that can donate protons (H^+ ions), while a base is described as a molecule that can accept protons. These definitions, while certainly correct, can be troublesome for students when they are presented with circumstances in which a base can donate electrons and an acid can accept them. For instance, at pHs below 9 - 10, amino groups, the major biological bases, act like classic bases as described above. However, at pH values above 10, an amino group actually tends to lose its extra proton; that is, under these extreme conditions of pH, a base behaves like an acid. Conversely, the carboxyl group, the major biological acid, acts like an acid at pH values in excess of 2 - 3. At more acidic pHs, this chemical group tends to acquire and retain the proton it usually loses in solutions at pHs above 2 - 3; at these lower pHs, this normally acidic group acts like a base. It is perhaps best to remind students that the definition for acids and bases was developed within a frame of reference at pH 7 where an acid behaves like an acid and a base like a base. It is wise to stress that an acid loses its protons at pH's above 2 - 3, while bases lose their protons at pH's above 9 - 10 as illustrated below:





In other words, both acids and bases can gain and lose protons. The difference between them is the pH at which this happens. The ionization of amino and carboxyl groups under different pH conditions is of paramount importance to protein structure since charges on such groups, when altered, can affect the three-dimensional shape of proteins.

In my experience, students do not have much difficulty comprehending buffers. I do, however, try whenever possible to get students to apply what they are learning to common everyday experiences. If I do it enough, I hope that they may get the idea and make the same sorts of connections on their own. A prime example of this refers to a well-known pain reliever. Most students are aware that because of its acidity, aspirin can exacerbate stomach and intestinal troubles in certain individuals. I ask the students to think back and come up with a product that is advertised as a remedy for this particular problem. One of them usually comes up with Bufferin; I then ask them if they know the origin of the name. Eventually, students (perhaps with some prodding) realize that buffer is added to the aspirin to neutralize its acidity and prevent acid build-up in the stomach. This also serves as an example of another practice I employ as much as possible. I urge my students to think about names. Why is Bufferin called Bufferin? What do the prefixes and roots in "exothermic" and "endergonic" mean? Instead of memorizing terms and their definitions, students can learn to dissect them and use their roots and prefixes as devices for remembering their meanings. To that end, wherever possible I point out these roots and prefixes and their meanings.

Macromolecules

Names and Nomenclature: Memorization vs. Understanding

In the discussion of macromolecules, a number of problems crop up that seem to obstruct student understanding. A conflict exists between the tendency of students to memorize facts about macromolecules and the use of previously acquired knowledge to extend their understanding about these molecules. A trivial example of this phenomenon is the difficulty students often seem to have remembering the names of the various bonds involved in holding together macromolecules. For example, many students cannot remember that the bonds that connect monosaccharide monomers in polysaccharides are called glycosidic bonds. Instead of realizing the connection between glycosidic bonds and the root glyco- (*sweet*), many students simply memorize the term. Biology majors, who form the majority of students in our cell biology course, have been repeatedly exposed to the meaning of this root; if they realized this connection, memorization (at least in its classic sense) would seem to be virtually unnecessary. This does not mean that memorization is a tool that cannot be used. The ester linkage (named by German chemist L. Gmelin) that arises from a reaction between an acid and an alcohol with the elimination of water, does not have an association with a root that would facilitate connecting the compounds involved with the name. Here memorization makes sense. Once the meaning of "ester" has been memorized, however, the names of other bonds derived from it can be remembered by using roots. The phosphate ester (P—O—C) linkage by which a phosphate group is attached to a phospholipid resembles an ester (C—O—C) with one of the carbon atoms replaced by a phosphorus atom. In DNA and RNA, the bond that holds together successive nucleotides is the 3'- 5'-phosphodiester bond. The atoms involved in this linkage can be lined up as follows: 3'C—O—P—O—5'C. Clearly, these are two sequential phosphate ester bonds connected by the middle phosphorus atom (phospho**di**ester). The numbers simply identify the carbons of the ribose or deoxyribose sugar participating in the interaction.

When memorization in place of understanding occurs, students often forget the material by the time the next exam rolls around; we call this a "core dump". Some students are seemingly unaware that most courses build on information throughout the semester. If they assume that they can forget previously tested material, they often do not realize that they are severely decreasing their chances of understanding subsequent topics. Realizing this, I emphasize as much as possible the importance of making associations rather than pure memorization. I also frequently stress, in as humorous a manner as possible, that principles covered earlier in the semester will reappear time and time again during the semester and that "dumping" them after the test on the material has been administered is inadvisable.

Carbohydrates

Describe different types of polysaccharides and emphasize the differences in their structure and function (see Chapter outline and table above). For example, mention the presence of α(1→ 4) glycosidic linkages in glycogen, amylose and amylopectin and stress that the branch points in amylopectin and glycogen involve α(1→ 6) glycosidic linkages. Point out the ability of mammals to digest these molecules, while another polymer of glucose, cellulose, cannot be digested by mammals; tell them that the reason for this seeming oddity stems from the glucose monomers of cellulose being connected by β (1→ 4) glycosidic linkages instead of α (1→ 4) glycosidic linkages. The shapes of these two polysaccharides are significantly different, especially in the region of these bonds. In glycogen and starch, each glucose monomer is oriented in the same direction. In contrast, the β– linkages in cellulose cause each glucose monomer to be upside-down relative to its two immediate neighbors. This significantly alters the shape of the polysaccharide in the area of the linkages. The enzyme(s) that hydrolyze(s) the α (1→ 4) glycosidic linkages cannot recognize the analogous regions in cellulose and, hence, it cannot be digested in mammals like ourselves. You can then ask students to explain how mammals that don't

possess the enzyme to digest cellulose can use cellulose as a food source (enteric bacteria digest the cellulose for them). Usually, they come up with the correct answer. I also have them consider the relevance of cellulose indigestibility in humans as it relates to the necessity for roughage or fiber in the human diet.

Lipids

It is important to emphasize that the lipids, while considered macromolecules, differ from the others (proteins, nucleic acids and carbohydrates) because they are not formed by polymerization. Emphasize the structures of the triglyceride and phospholipid building blocks: glycerol and fatty acids. Glycerol is viscous and seems greasy. Yet it is really hydrophilic. I ask my students the reason for that. Usually, they note the three hydroxyl groups (one attached to each of the three carbon atoms of glycerol) and come up with the right answer. They are also often able to recognize the hydrophobic (the chain of carbons) and hydrophilic (the terminal carboxyl group) regions of fatty acids. Once the assembly of these building blocks into triglycerides has been described, ask why the product of this assembly is completely hydrophobic. This occurs, of course, because the hydrophilic portions of both molecules are obliterated during triglyceride formation.

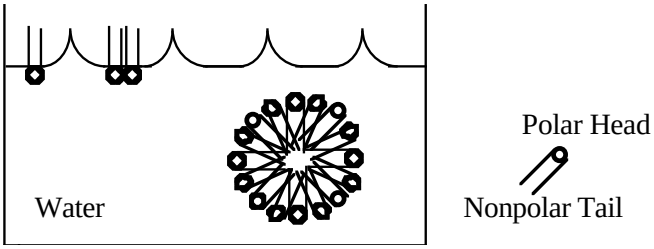
This is also a good time to address the issue of saturation and unsaturation in the fatty acids of triglycerides. Once again, examples from the students' common experiences are valuable; they have all encountered animal fat and salad oil but have probably never been aware of their structures and the effect of those structures on their properties. Fats that are solid at room temperature contain predominantly fatty acids that are saturated (no double bonds). As a result, these substances can pack more closely together when present in higher quantities since their carbon chains are straight. On the other hand, lipids that contain largely unsaturated (with one or more double bonds) fatty acids are fluid and classed as oils. because their carbon chains have kinks at the position of each double bond. This makes closer packing unlikely and explains the liquid nature of oils. Remind students that they are inundated by advertisements lauding the virtues of polyunsaturated fats like those from plants, the dangers of animal fats in their diets and the user-friendliness of soft margarine. It is instructive to ask students to compare butter that contains animal fats with soft margarine, as both are removed from the refrigerator. Butter is, of course, harder and soft margarine much softer. This clearly illustrates the difference between saturated and polyunsaturated fats.

Contrast the hydrophobicity of triglycerides to the amphipathic nature of phospholipids. Since these lipids are built of a negatively charged phosphate group (to which other hydrophilic groups normally attach) along with the glycerol and fatty acids found in triglycerides, the resultant molecule possesses a polar head and nonpolar tail consisting of two fatty acids. Such a molecule has a split personality that leads to interesting structural and functional implications. Indeed, they form a significant portion of the cell membrane lipid bilayer. The behavior of these molecules in water serves as a good example of these implications (see Analogy box below).

ANALOGY

The Behavior of Phospholipids in Water

If phospholipids are spread on the surface of a water solution as they would be in a bowl of Grandma's chicken soup, they will arrange themselves so that their polar heads are in contact with the water, while their hydrophobic tails project into the air above. If you push separate puddles of floating phospholipid together with your spoon, these puddles will fuse into one in order to expose as little phospholipid to water as possible. This behavior is similar to what occurs when secretory vesicles approach the cell membrane during exocytosis. The membrane of a secretory vesicle fuses with the cell



membrane, opening the vesicle to the extracellular space and releasing its contents. If the soup is stirred vigorously, some of these lipids are driven beneath the surface of the water, exposing the hydrophobic tails to water. As phospholipids encounter each other in this underwater environment, they are able to enhance their collective stability by forming spheres with the hydrophilic heads facing and interacting with the surrounding water, while the hydrophobic tails interact with each other in the center of the structure away from the water. Such a structure is called a micelle. If there are many phospholipids around, these structures can get larger, in which case they form a double-layered structure with two layers of phospholipid tails in the center and each surface exposed to water and covered with hydrophilic head groups, a structure called a liposome. Micelles are quite useful in the shower. Imagine that you have just spent the day mowing a few acres of grass without the luxury of a riding mower. You are as dirty as it gets and you decide to take a shower to get squeaky clean. After stepping into the stream and wetting yourself down, you soap up and create an appreciable lather. This lather contains micelles. Since dirt is often oily, the micelles will pick up the hydrophobic dirt molecules and since hydrophobic materials are soluble in other hydrophobic materials, the dirt will be attracted into the region of the micelle occupied by the phospholipid tails. When you step back into the shower spray, the water carries the micelles and the greasy dirt they contain off your body and down the drain, because the outer surface of the micelles interacts well with the water.

Proteins

Structure and Chemistry of Amino Acids

Proteins are the most varied of macromolecules in structure and versatility of function. Emphasize the varied roles played by proteins in living cells and then proceed to a discussion of their structure. Macromolecular structure is often complicated for students to grasp, especially when emphasis is placed on the levels of protein structure. This is another area where students are tempted to memorize the material. Naturally, I start with the

amino acids. Some instructors feel it is important to have their students memorize the structures of R groups. I don't feel that this is particularly advantageous, especially since the students will usually be asked to do this when they take a Biochemistry course. I tell my students that they will not be required to memorize the structure of the amino acid R groups. The resultant classwide sigh of relief is startlingly audible. Since R groups are responsible for the different chemical properties of individual amino acids, I do require the students to recognize the chemical nature of the R groups. If students are shown an amino acid, I expect them to know whether it is polar or nonpolar, polar, charged or polar, uncharged. Most students will be able to do this, given what they have already been taught about hydrophobicity and hydrophilicity.

Polymerization of Amino Acids

Polypeptides, polysaccharides and polynucleotides form through polymerization. This is a concept that can, at times, be difficult for students to grasp. They have trouble understanding how such seemingly different molecules can be hooked together in an orderly fashion. By stressing that amino acids have both a constant portion (the α-carbon, the amino group and carboxyl group) and a variable portion (the R group), it is easy to make the point that the peptide bond is formed by interactions between the constant portions of each amino acid. This naturally allows the bonds to form by the same process no matter which two of the twenty amino acids are being connected. It is not unlike a freight train that has a wide variety of cars, all of which are fitted with identical couplers.

Emphasize that during polymerization, the monomers are connected by condensation reactions with the release of water as a byproduct. Also stress that the introduction of a water molecule across the bond leads to the breaking of the bond, a reaction called **hydrolysis** (hydro - *water*; lysis - *loosening*).

Levels of Structure in Proteins

The levels of structure in proteins are often a source of confusion for students. Perhaps the biggest problem is the tendency of some students to confuse the structural levels in proteins with those in nucleic acids. This is partially due to the structural level names of the two groups of macromolecules and the similarities in their definitions. Perhaps the most common error on exams that I have administered is confusion about the monomer building blocks for proteins and nucleic acids. Many students have said that proteins are polymers of nucleic acids or that amino acids constitute the building blocks of polynucleotides. This occurs even if a special effort is made to warn them that such errors are common.

It is valuable for students to learn the levels of structure of proteins; spend a significant amount of time on the topic. Emphasize the definition of each level, the bonds involved in maintaining the structure and the role that each level plays in the overall structure and function of polypeptides. *Primary structure* is the sequence of amino acids in the polypeptide from the N-terminal end (with its free amino group) to the C-terminal end (with its free carboxyl). Stress that the primary structure determines all of the higher levels of structure. There are a number of examples of this, but the classic example of sickle cell anemia hemoglobin is hard to beat. A change in one amino acid in a chain of 146, the β chain of hemoglobin, alters the higher levels of structure enough to cause sickle cell anemia. Sickle cell hemoglobin crystallizes under conditions of low oxygen tension leading to the disease symptoms. The sixth amino acid in the normal chain, which is a glutamic acid (polar, charged amino acid) is changed to a valine (a hydrophobic amino acid) in the mutant hemoglobin β chain. Due to the chemical differences of the component amino acid R groups, the normal formation of the higher levels of structure is disrupted.

Secondary structure is structure caused by interactions between adjacent or nearly adjacent areas of the backbone, primarily in the region of the peptide bonds. It is stabilized by H bonds and comes in two varieties: the α-helix and the β-pleated sheet. Both of these structures make the maximal number of H bonds possible within the area of the backbone participating in the interaction and are, thus, equally stable. The major difference between them is that the H bonds in an α–helix are oriented parallel to the molecular axis, while the H bonds in a β-pleated sheet are oriented perpendicular to the axis. The regular arrangement of amino acids in the α–helix is a result of the rigidity of the peptide bond. This rigidity is, in turn, the product of resonance between the carbonyl group and its adjacent peptide bond. Electrons from the doubly bonded oxygen of the carbonyl occasionally move to the peptide bond. This prevents rotation around the bond as is typical of double bonds. Illustrate this concept to students using molecular models, pencils and Styrofoam balls or simply your fingers. The resultant rigidity leads to a structure called the amide plane within which lay the atoms of the polypeptide backbone from the central carbon of one amino acid to that of the next in line. The amino acid chain thus resembles successive, rigid planar structures joined by a single bond about which they can rotate.

Analogy

The Slinky Analogy

In space, the α-helix resembles a Slinky with the rungs of the Slinky corresponding to the polypeptide backbone, while the H bonds would connect successive rungs and thus be parallel to the axis of the molecule running through the center of the Slinky. The H bonds in an α–helix connect every fourth amino acid (i.e., amino acid 1 is H bonded to amino acid 5, etc.) as the chain makes one full turn around the helix every 3.6 amino acids. This brings, for example, the imino group (N-H) of amino acid 5 beneath the carbonyl group (C=O) of amino acid 1 so that an H bond can form.

Analogy

The Jacob's Ladder Analogy

The nearest example of the amide plane from common experience that comes to mind is a Jacob's Ladder, a child's toy. It consists of planar slats connected to each other by flexible straps. If you hold the top slat with the rest of the toy suspended below it, you can tilt it. Suddenly, the slat below it will be released and it will swivel downward causing the slat below it to be released. This chain reaction continues to the bottom of the ladder. If one lays this toy on its side and tries to pull it into as tight a circle as possible, there is a limit as to how tight a circle can be formed. This is due to the planar slats that constrain the structure from forming too tight a circle. A similar constraint prevents the α–helix from having fewer than 3.6 amino acids per turn.

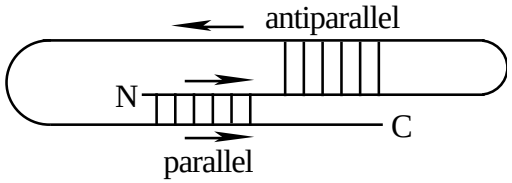
Keratin from hair is an example of a protein that contains lots of α–helix. It is extremely strong and extensible. If enough force is applied to hair to exceed the strength of the collective H bonds in the α–helices of the hair, they will separate and the polypeptide backbones will break, tearing the hair. Otherwise, once the force is removed, the helix will return to its original shape like a Slinky.

Image

The Analogy of Superman's Hair

When I talk about keratin and α-helices, I am reminded of reading Superman comic books as a child. Occasionally in those stories, visits to the Superman museum were illustrated. One of the displays at the museum consisted of a hair generously contributed by the Man of Steel to which was tied a one ton weight. Obviously, Superman's hair contains more H bonds than normal human hair.

The β-pleated sheet, as its name suggests, resembles a pleated sheet. It is strong, as well as flexible, but not extensible. Fibroin, the major protein constituent of silk, has large amounts of β-pleated sheet. The H bonds that join adjacent parts of the backbone are oriented perpendicular to the molecular axes. If the two chains connected run in the same direction, the β-pleated sheet is called parallel. In contrast, if the H-bonded chains run in opposite directions, the structure is named antiparallel. This is illustrated in the drawing below:



Image

The Use of Silk During World War II

When I was a kid, I would often frequent a local Army - Navy surplus store. The store sold maps of China printed on silk priced at 50 cents. I bought a couple. The owner of the store, obviously trying to make the sale, told me why they were used. He said that pilots would take these maps with them during bombing missions over Japan. Since they could not land on the aircraft carriers from which they had taken off, they would continue over Japan and land in China. The maps were handy because they could be removed from pockets, quickly perused and then jammed quickly back into a pocket. Folding the maps carefully was unnecessary because of the strength and flexibility of the silk (fibroin). If the maps were paper, like those we can get from an auto club, even if there was time to fold them carefully, they would eventually tear. If they were crumpled up like the silk, they would be irretrievably wrinkled. Furthermore, there was always a chance that the pilots could land in an area of China occupied by the Japanese. The military would not want these maps to fall into enemy hands. To prevent this, the maps could be rolled into a little ball and swallowed. Since they are protein, they would be digested and unavailable to the enemy. This could not be easily done with a paper road map, which consists of polysaccharide fibers that could not be digested due to their β-glycosidic linkages.

Tertiary structure results from interactions between R groups within the same polypeptide chain. It involves hydrophobic interactions, H bonds, van der Waals forces, ionic bonds and a type of covalent bond, the disulfide linkage. The disulfide links form when two cysteine residues are moved into close proximity by protein folding. Cysteine residues have at the end of their R group a sulfhydryl (—SH) group. When two of these groups approach one another, the H atoms leave and a link forms between the two sulfur atoms, hence the name disulfide link. The disulfide links stabilize the folded structure.

Folding of the protein chains begins before their synthesis is completed. Since it involves adjacent or nearly adjacent regions of the polypeptide backbone, secondary structure forms first after the primary structure is laid down. This introduces some twists and turns into the chain, changes the shape of the molecules to some degree, and is followed by interactions between R groups that serve to fold the molecule even further. Eventually, the protein is folded into its proper three-dimensional shape. Once this has occurred, the disulfide linkages form and the molecule becomes fully functional. All the information needed by proteins to attain their proper shape is encoded within the protein sequence and many can fold completely without any help whatsoever. However, some proteins require the help of molecular chaperones or chaperonins to fold up. They speed up the folding of some molecules that would fold up much more slowly on their own. Furthermore, as the proteins fold into their final tertiary conformation, hydrophobic amino acid residues tend to end up in the center of the protein, while the polar

amino acids tend to wind up on the protein's outer surface exposed to water. When proteins fold, they do so in steps. After the first step, there are clues in the new structure to the next step. This process continues step-by-step, until the protein is completely folded. Each of these steps is called a nucleation state.

Analogy

The Garden Shears - Hedge Clippers Analogy
Garden shears and hedge clippers have a safety feature that keeps them closed when they are stored in your garage. It is a latch on one handle of the shears that fits over a post on the other handle. When the latch is engaged, the handles are secured and the shears cannot be opened until the latch is released. If the shears are open, it is impossible to close the latch over the post since the two of them are just not close enough together. On the other hand, if the shears are closed, the latch can be engaged. This is analogous to what happens in proteins. The disulfide links do not form until the protein has essentially attained its final folded structure. The cysteine residues that will participate in the disulfide linkages are then close enough for the bonds to form.

Analogy

The Jig-Saw Puzzle Analogy
The folding of a protein into its final three-dimensional shape (its tertiary structure) is a lot like a jigsaw puzzle. When most people assemble a jigsaw puzzle, they look first for the corner pieces with two straight sides and then for other pieces with one straight side. These pieces are then assembled into the frame of the puzzle (the first nucleation state). Once the frame is completed, it provides clues for the next group of pieces to be laid in. These provide another set of assembly cues and so on until the puzzle is completed.

Quaternary structure involves interactions between R groups. All of the bonds that participate in tertiary structure can participate in quaternary structure. The only difference is that the interactions are between R groups on different polypeptide chains in quaternary structure, not between R groups on the same chain as in tertiary structure.

Nucleic Acids

The good news about nucleic acids is that many of the general features are understood by the bulk of Cell Biology students. This may be due to increased exposure to these basic principles in high school and introductory level courses in college or to the inherent, elegant simplicity of A pairing with T and G with C. They are also familiar with the idea of the double helix, although confusion does arise once they have been introduced to the α-helix. They do, however, tend to get a little dim when the finer points of function are introduced.

General Structure of Nucleotide Monomers

Introduce the students to the components of nucleotides (phosphate + 5-carbon sugar + nitrogenous base) and nucleosides (5-carbon sugar + nitrogenous base). Once again, I do not require my students to memorize the structure so that they can draw it. I am content if they can look at a drawing and recognize the sugar, base and phosphate group of a nucleotide. Emphasize the negative charge on the phosphate group making it hydrophilic. When discussing the sugars, point out the important elements of their structure. Remind the students that sugars are generally hydrophilic. Stress the roles of each carbon in the sugar. Nitrogenous bases are attached to the 1'-carbon of the 5-carbon sugar. The 2'-carbon can be used to identify the sugar as ribose or deoxyribose, with the presence of an oxygen atom below the plane of the ring at that position indicating ribose and its absence

indicating deoxyribose. The 3'- and 5'- carbons participate in the bonds that attach adjacent nucleotides. The 4'-carbon connects the 3'- and 5'-carbons. The nitrogenous bases are a bit harder for some students to grasp. Point out that these structures are composed of carbon-containing rings that contain an occasional nitrogen. Ask the students what this reveals about nucleotide chemistry and, if necessary, guide them to the answer (hydrophobic due to the ring structure). Tell them that the bases will occasionally have a group protruding from them that is capable of engaging in H bonds (—NH_2, =O). This, of course, means that bases are partially hydrophilic as well and, therefore, amphipathic. Once again, I do not expect the students to memorize the structures of bases, but I do expect them to be able to recognize the purines (adenine, guanine) with their two rings and distinguish them from the pyrimidines (thymine, cytosine, uracil) with their single ring structures. To aid them in remembering the differences between pyrimidine and purine structure, suggest that they associate the smaller name (purine) with the larger structure and the larger name (pyrimidine) with the smaller structure. Remind them that they may distinguish DNA and RNA by the presence of uracil in RNA replacing thymine (only found in DNA).

Biochemical analyses by Erwin Chargaff revealed that the number of adenine bases in a given DNA sample was equivalent to the amount of thymine bases and that the amount of cytosine was equal to the amount of guanine. Furthermore, he found that the G+C/A+T ratio differs from organism to organism, while the pairing rules hold for all organisms. As an aside, I have found two books about the search for DNA structure and the origins of molecular biology enlightening: The Double Helix by James Watson and The Eighth Day of Creation by Horace Freeland Judson, a somewhat broader study.

Polymerization of Nucleotides and the Double Helix

Individual nucleotide monomers, whether RNA or DNA, are joined together by condensation reactions between the 5'-phosphate group of one nucleotide and the 3'-OH group on the sugar of another nucleotide. The resultant bond is called a 3'-5' phosphodiester bond. Break down the name of this bond to illustrate how seemingly intimidating lengthy words can be dissected to extract their meaning. The bond connects the 3' and 5' carbons of successive nucleotides in the chain. It is a combination of two phosphate esters which accounts for the rest of the name. The resultant chain has a 5' → 3' polarity. In DNA, these polynucleotide chains usually come in pairs and they are antiparallel. A base in one chain pairs with the complementary base in the opposite chain. For example, a thymine nucleotide in one chain will pair with an adenine nucleotide in the opposite chain and a cytosine nucleotide will pair with a guanine nucleotide. This complementary relationship also explains the ability of DNA to replicate, an essential feature of the genetic material and its ability to pass the genetic code on to complementary RNA molecules that carry the code to the site of protein synthesis. The double stranded DNA that forms as a result of the pairing resembles a spiral staircase in shape. The phosphate-sugar backbone, which is hydrophilic, will form the railings of the staircase. The paired nitrogenous bases represent the actual stairs. The hydrophilic backbone essentially covers the outside surface of the molecules, while the hydrophobic paired bases are sheltered from water in the interior of the helix. They stack one on top of the other by way of hydrophobic interactions. As in proteins, the arrangement of hydrophilic groups on the outside and hydrophobic groups on the inside contributes much stability to the double helix.

The regularity of the width of the polynucleotide is also striking. Ask your students how they might have guessed about the lack of variation in the width even if they had not been told about it. Sometimes, a student will realize that the number of rings that fit across the helix at the site of each base pair is always three, since purines (2 rings) always pair with pyrimidines (1 ring).

Levels of Structure in Nucleic Acids

The levels of structure of polynucleotides, as mentioned previously, are often confused with those of polypeptides and it is important to stress the definitions of each level. *Primary structure* of polynucleotides simply refers to the nucleotide sequence from one end of the molecule to the other (usually from the 5' to 3' end). This applies to both DNA and RNA polynucleotides. Mention the effort to sequence the entire human genome and raise issues about the cost, potential benefits and potential disadvantages of this mammoth and now completed effort.

Students will frequently get secondary structure in proteins and nucleotides confused. This is invariably due to the similarity in the names of the α-helix and the double helix, which students easily mix up. I have not found a good way around this problem other than warning students each semester of the confusion that has plagued their predecessors. *Secondary structure* in DNA is the double helix. In RNA, secondary structure refers to local areas within the single RNA strand where intrastrand pairing can and does occur. It is useful to expose students to the likelihood of this happening. *Tertiary structure* in polynucleotides is easy to define, but sometimes difficult for students to visualize. It is defined as any further coiling or twisting of the polynucleotide chain above the level of the double helix. The most common example of this is called supercoiling.

Analogy

The Telephone Cord Analogy

The easiest way to describe supercoiling to students is by using an example they are sure to have experienced, the telephone cord. If you have a phone cord, especially a long one on your kitchen phone, you will often walk aimlessly around the kitchen while talking on the phone. When you hang the phone up, you will usually discover that your helical phone cord (analogous to the double helix although it has only one "strand") will be hopelessly twisted. This extra twisting above and beyond that of the regularly coiled phone cord is analogous to supercoiling or tertiary structure in nucleic acids.

Analogy

The Balsa Wood Airplane with Rubber Band Analogy

Another toy from childhood also demonstrates supercoiling. When we were kids, many of us at one time or another had balsa wood airplanes with propellers powered by rubber bands. We would turn the propeller that turns the rubber band and begin to twist it. At first, the rubber band would adopt a conformation reminiscent of the double helix. However, as we turned the band tighter and tighter, the rubber band would become supercoiled. This would be reflected by irregular bumps along the length of the rubber band. Interestingly, this supercoiling can be relieved by releasing the tension being applied to the propeller. The same sort of thing happens to relieve supercoiling in DNA. An enzyme (topoisomerase) will cut one of the strands of DNA and allow the supercoiling to be relieved before reconnecting the severed ends of the chain.

RNA Structure and Function

Briefly summarize the structure and functions of the different major types of RNA: ribosomal RNA (rRNA), transfer RNA (tRNA) and messenger RNA (mRNA). Remind the students of the differences between RNA and DNA summarized above, including the fact that RNA is usually single-stranded. Also mention here and elaborate in later lectures on the idea that RNA can sometimes serve as an enzyme, a ribozyme.

CRITICAL THINKING QUESTIONS

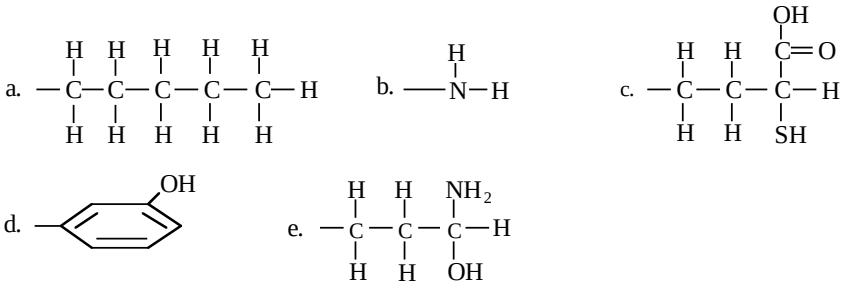
1. When are the valence electrons participating in a covalent bond shared equally between the two bonded atoms? *The valence electrons are shared equally between the two atoms when the two atoms are atoms of the same element or if the two atoms are not electronegative, as in a bond between a carbon and a hydrogen atom.* What is created when a bond has an asymmetric distribution of charge? *A dipole.* What kind of molecule is made up of atoms connected by covalent bonds that exhibit no unequal charge distribution? *A hydrophobic molecule.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

2. What kind of bond is created when one atom transfers an electron completely to another atom which accepts it? *This would be an ionic bond.* Why are stable ionic bonds relatively weak (and rare) in a living cell? *Ionic bonds in a cell are generally weak due to the presence of water in cells.* Where in a living cell can ionic bonds be stronger and more stable? *Ionic bonds can be more stable and stronger deep within the core of a protein or anyplace where water can be excluded.*

Learning Objective: 2.3 Describe the role of noncovalent bonds in the structure of molecules such as water.
Section Reference: 2.3

3. Which of the groups below is capable of only hydrophobic interactions? Explain your answer. *Answer a. is capable of only hydrophobic interactions. It contains no ionizable or hydrophilic groups.* Which group is capable of only hydrophilic interactions? Explain your answer. *Answer b. is capable of only hydrophilic interactions, since it has no component with a long carbon chain or a carbon-containing ring and no nonpolar covalent linkages. It is also capable of ionization.*



Learning Objective: 2.3 Describe the role of noncovalent bonds in the structure of molecules such as water; and 2.5 Describe the general structure and functions of biological molecules.
Section Reference: 2.3 and 2.5

4. You are studying two acids. One of them, acid X, donates its proton to water fairly easily and the other one, Acid Z, remains undissociated when dissolved in water. Which one is the stronger acid? *Acid X is the stronger acid because it dissociates fully, while Acid Zremains undissociated when dissolved in water.*

Learning Objective: 2.4 Explain the characteristics of acids, bases, and buffers.
Section Reference: 2.4

5. You treat a partially purified preparation of protein with a reagent that breaks bonds between sulfur atoms. Which level(s) of protein structure are likely to be affected the most? *Both the tertiary and quaternary levels of structure would be affected since those levels are the only ones in which disulfide bonds are prominent.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

6. Not all proteins are able to renature. Some proteins when exposed to heat or some other denaturing treatment are irreversibly denatured. What is an example of such a protein? (Hint: Think of a daily meal.) *Egg white protein and yolk are examples of proteins that are irreversibly denatured by heat.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

7. You are working with an enzyme altase that you denature in the presence of urea. If altase were denatured no further by the addition of mercaptoethanol, what would that suggest to you about the enzyme? *The enzyme probably contained no disulfide linkages since mercaptoethanol breaks such linkages.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

8. Would all proteins be likely to require exposure to mercaptoethanol in order to accomplish full denaturation? If not, what trait would a protein that did not require mercaptoethanol possess? *Not all proteins would require mercaptoethanol to accomplish full denaturation. If a protein has no disulfide linkages, it probably would not require mercaptoethanol for full denaturation.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

9. An enzyme is placed in a solution containing urea. Assuming that this protein contains no disulfide linkages, is it reasonable to suspect that it will be totally denatured by the treatment? *Placement in a urea solution should totally denature the enzyme, especially since there are no disulfide linkages.* How could you know that the enzyme has, in fact, been denatured? *If there are extensive hydrophobic interactions between enzyme R groups, total denaturation may be difficult to accomplish. If the enzyme activity disappears, there is a good chance the enzyme has been denatured.* Why does the urea denature the tertiary structure of the enzyme? *Urea breaks up the tertiary structure by interfering with hydrophilic interactions like H bonds.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

10. Which of the following tripeptides would be most likely to be soluble in an organic (hydrophobic) solvent like benzene: N - phenylalanine - alanine - glycine - C, N - leucine - alanine - lysine - C, N - proline - phenylalanine - leucine - C, N - arginine - lysine - proline - C, N - glutamate - aspartate - glycine - C? Explain your answer. *N - proline - phenylalanine - leucine - C would be most soluble in a hydrophobic solvent. All three amino acids are classed as nonpolar amino acids and would be likely to be soluble in benzene. In the other tripeptides, at least one of the amino acids does not belong to the nonpolar class.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

11. What structural feature of RNA would be disrupted by a reagent that breaks apart hydrogen bonds? *RNAs generally consist of a continuous single strand, but they often fold back on themselves to produce molecules having extensive double-stranded segments. The double-stranded regions are held together by hydrogen bonds between the bases.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

12. What would happen to a double-stranded DNA molecule if you treated it with a reagent that breaks apart hydrogen bonds? *The DNA would separate into two individual and separate strands. Hydrogen bonds are responsible for holding together the two strands of the DNA double helix, just like they hold together the double-stranded regions of RNA molecules.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

13. Mammals lack the enzyme that hydrolyzes cellulose. Yet many mammals are herbivores, and they eat grass and other plant material for nutrition. How can this be, given that they cannot digest the food they are eating? *While these animals lack the enzyme that digests cellulose, bacteria that reside within their digestive tracts possess that enzyme. There is a symbiotic relationship between the two organisms. The herbivores seek out and eat the grass; the bacteria in their digestive tract digest it. What they don't use, the herbivore does.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

14. What are some possible explanations for the branched structure of glycogen? *First, branching allows more efficient storage of energy. More glucose monomers can be stored in a smaller space. Second, branching creates more free ends on the structure. This would allow glycogen to be disassembled more rapidly when free glucose is needed and would also allow quicker assembly when glycogen is being constructed.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

15. A molecule consists of a long, unbranched chain of carbons connected by single bonds and bound to the maximum number of hydrogen atoms possible. There is a carboxyl group at one end of the molecule. What kind of molecules is it likely to be? *It is likely to be a fatty acid.* Is this molecule hydrophobic or hydrophilic? *It has both hydrophilic (the carboxyl group) and hydrophobic (the long chain of carbons) character; thus, it is amphipathic.* Is it saturated or unsaturated? *It is saturated because the carbons are connected only by single bonds and the chain is bound to the maximum number of hydrogen atoms that is possible.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

16. Scientists have sequenced proteins by using specific proteases to "clip" a purified protein preparation between two specific amino acids, thus forming a number of moderately sized fragments; they have used acid hydrolysis to produce smaller fragments. Each fragment can then be sequenced by breaking the moderate fragments into dipeptides that are easily sequenced. The fragments below are obtained after the initial enzymatic cleavages. Can you deduce the sequence of the original polypeptide? (**HINT:** the original cleavages at specific locations differ depending on which proteolytic enzyme was used to create each fragment; this causes an overlap in the fragments' sequences.) The final polypeptide should have 18 amino acid residues.

    N - ala - ala - gluN - aspN - met - C
    N - iso - pro - aspA - try - thr - C
    N - met - cys - leu - lys - phe - arg - aspA - C
    N - aspN - met - cys - leu - lys - C
    N - aspA - try - thr - phe - tyr - ala - ala - C

N- iso - pro - aspA - try - thr - phe - tyr - ala - ala - gluN - aspN - met - cys - leu - lys - phe - arg - aspA - C

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

17. Many so-called temperature-sensitive mutations have been discovered in a wide variety of organisms. These are proteins that are non-functional at higher temperatures, while, at lower temperatures (often just a few degrees lower), they function normally. For example, the coloration patterns in Siamese Cats arise from a temperature-sensitive mutation. An enzyme required for the synthesis of dark pigment is unable to function in areas close to the body where normal physiological temperatures prevail. However, at the tips of the ears, paws, the tip of the tail and other extremities where the temperature is slightly lower, the enzyme works correctly and dark pigment is produced. What is happening at the molecular level that explains this? *In warmer areas of the organism, the temperature is just high enough to denature the enzyme in question. Since it is denatured, it will not work properly and dark pigment will not be produced in those areas.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

18. Which of the following tripeptides would be most likely to be soluble in an organic (hydrophobic) solvent like benzene? Explain your answer.
    a. N - phenylalanine - alanine - glycine – C
    b. N - leucine - alanine - lysine - C
    c. N - methionine - valine -isoleucine - C
    d. N - arginine - lysine - proline - C
    e. N - glutamate - aspartate - glycine – C

    Answer c. N - methionine - valine - isoleucine - C would be most soluble in a hydrophobic solvent. All three amino acids are classed as nonpolar amino acids and would be soluble in benzene. In the other tripeptides, at least one of the amino acids does not belong to the nonpolar class.

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

19. You have purified and characterized a protein and discover that it is composed of 2 subunits, an α-subunit and a β-subunit. What word describing its subunit structure would apply to this molecule? *It is a heterodimer.* What is the structure of hemoglobin? *Hemoglobin is composed of 4 subunits (2 α-globin and 2 β-globin subunits); each subunitcontains one heme group.* How many molecules of oxygen does a complete hemoglobin molecule bind? *A complete hemoglobin molecule binds to 4 oxygen molecules, one binding to each subunit.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

20. You have purified all of the proteins in a cell and constructed a list of the collection of those proteins. What is this list of proteins called? *It is a proteome.* How could proteomics be used to determine the effect of glucocorticoid (a steroid hormone) exposure on embryonic chick retina cells? *You could expose embryonic chick retinas in culture to an appropriate concentration of glucocorticoid hormone and, at the same time, culture control retinas that were not exposed to the hormone. You could then isolate proteins from the hormone-treated retinas and construct a proteome for the hormone-treated cells. You could then repeat the exercise for the control, untreated embryonic chick retina cells. Once you had the proteomes for the hormone-treated and control cells, you could compare them to determine any changes in the protein complement from the two proteomes. A protein could increase in amount, decrease in amount or remain unchanged in amount after hormone treatment. If there is an increase or a decrease, it is likely that the hormone had an effect.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

21. Distinguish between a proteome and an interactome. *A proteome is a list of all the proteins made by a particular cell in a particular organism at a particular time of development. An interactome for that same cell would be a map that shows all of the proteins from that same cell that co-purify with each other and, therefore, presumably interact inside the cell. It would be a subset of the proteome for that particular cell.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

22. An investigator has isolated an individual gene from human chromosomes and altered its information content in a precisely determined way. The modified gene was then inserted into a bacterium, and the modified protein was synthesized with its altered amino acid sequence. What is this technique, in which a specific site within the protein was altered by changing the corresponding nucleotide within the gene encoding that protein? *The technique is called site-directed mutagenesis.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

23. Different versions of a protein are adapted to function in different tissues or at different stages of development. As an example, humans have six different genes that encode different versions of the contractile protein actin. Two are found in smooth muscle, one in skeletal muscle, one in heart muscle and two more in virtually all other types of cells. What is the term that is used to describe these various, closely-related versions of the same protein? *Isoforms.* What word is used to describe evolutionarily related proteins in organisms living in very different environments? *They are homologous proteins.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

24. What is the general name of a molecule that consists of a phosphate group, ribose and uracil? *It is a nucleotide.* What is the general name of a molecule that consists of deoxyribose and guanine? *It is a nucleoside.* What is the name of the nucleotide that binds to a variety of proteins and acts as a switch to turn on their activities and what is the general name of the proteins to which this nucleotide binds? *The nucleotide is GTP and the proteins are called G proteins.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

25. What macromolecules self-assemble to form ribosomes? *RNA and proteins.* What macromolecules self-assemble to form tobacco mosaic viruses? *RNA and proteins.*

Learning Objective: 2.8 Analyze the evidence supporting the idea that bacterial ribosomal subunits are capable of self-assembly.

Section Reference: 2.8

26. What type of radiation released by atoms is similar to X-rays? *Gamma radiation.* What type of radiation released by atoms is equivalent to an electron? *A beta particle.* What type of radioaction released by atoms is similar to a helium atom nucleus? *An alpha particle.*

Learning Objective: 2.2 Explain the chemical basis of the use of radionuclides in imaging and treatment.
Section Reference: 2.2

27. Although technetium-99 (^{99}Tc) is routinely used as an injectable tracer in patients, why would hospitals purchase molybdenum-99 (^{99}Mo) instead? *The half-life of ^{99}Tc is so short (i.e. 6 hours) that most of it would decay during shipping before injecting into the patient. Molybdenum-99 has a longer half-life of 66 hours, decays into ^{99}Tc, and the two can be separated so that ^{99}Tc can be injected into the patient.*

Learning Objective: 2.2 Explain the chemical basis of the use of radionuclides in imaging and treatment.
Section Reference: 2.2

28. How does the gamma knife make use of cobalt-60 (^{60}Co) to more effectively treat brain tumors than a single gamma beam? *Chunks of ^{60}Co are placed in a hemisphere design around the patient and selectively focused on the tumor. The individual beams are not damaging to tissues, but the region of their overlap at the tumor site experiences an additive and large effect.*

Learning Objective: 2.2 Explain the chemical basis of the use of radionuclides in imaging and treatment.
Section Reference: 2.2

29. Photosynthetic plants are able to fix carbon from cabon dioxide in the atmosphere to build sugar molecules for energy. Are they able to similarly fix nitrogen as well from the air to make amino acids? *No, but a symbiotic relationship occurs between legumes and the bacteria rhizobia. Rhizobia colonizes the root nodules and converts nitrogen to ammonia, which the plant then uses to make nitrogen-containing biochemicals.*

Learning Objective: 2.6 Identify the chemicals in fertilizer that are crucial to plant growth.
Section Reference: 2.6

HUMAN PERPECTIVES QUESTIONS: FREE RADICALS AS A CAUSE OF AGING

1. What kinds of conditions can cause free radicals? *Free radicals may form when a covalent bond is broken such that each atom that had participated in the bond retains one of the two shared electrons that comprised the bond. They may also form when an atom or molecule accepts a single electron transferred during an oxidation - reduction reaction. Water, for example, can be converted into free radicals when exposed to solar radiation.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

2. Why are free radicals capable of altering molecules, such as proteins, nucleic acids, and lipids? *They are extremely reactive which makes them well suited for chemically altering these molecules.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

3. What is the name of substances that are able to destroy free radicals in a test tube? These substances are sold in health food and grocery stores because of their reputed ability to destroy free radicals and thus retard aging, although there is no convincing evidence that they actually retard the aging process. *Such substances are called antioxidants.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

4. What is some specific evidence that demonstrates the importance of superoxide dismutase (SOD) in getting rid of superoxide free radicals? *Mutant bacteria and yeast cells that lack SOD activity are unable to grow in the presence of oxygen. Furthermore, mice that lack the mitochondrial version of the enzyme (SOD2) are not able to survive more than a week or so after birth.*

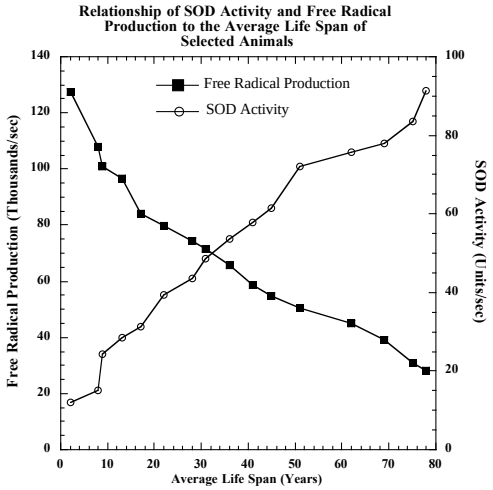
Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

5. Why might an organism that had functional SOD but mutant catalase and/or glutathione peroxidase be at a disadvantage? *SOD converts two superoxide free radicals and two hydrogen ions into hydrogen peroxide and oxygen. Hydrogen peroxide is also a highly destructive substance; it is a potentially reactive oxidizing agent that is often used as a disinfectant and bleaching agent. Hydrogen peroxide is normally destroyed in the cell by catalase and glutathione peroxidase. Without one or both of these enzymes, the organism would be less able to get rid of hydrogen peroxide. It would be likely to build up and could, if its concentration rose high enough, damage important molecules within the cell.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

6. a. It has been hypothesized that aging results from tissue damage caused by free radicals. What are free radicals? *Free radicals occur when atoms or molecules have orbitals containing a single unpaired electron. They are highly unstable and extremely reactive chemical groups and are produced during normal metabolic processes. They can chemically alter many types of molecules, including proteins, nucleic acids and lipids; they may also damage tissues.*

b. Some time later, the enzyme superoxide dismutase (SOD) was discovered. The sole function of this enzyme is the destruction of the superoxide free radical. A connection between free radicals and aging has not been firmly established, but a few predictions assuming their involvement in aging have been made. Below are graphs depicting hypothetical data collected testing some of these hypotheses. Interpret the results of each graph. What do they tell you about the effect of free radicals on aging and the role of enzymes that neutralize free radicals? Consider each graph separately. Do not try to combine the results to form a coherent model.



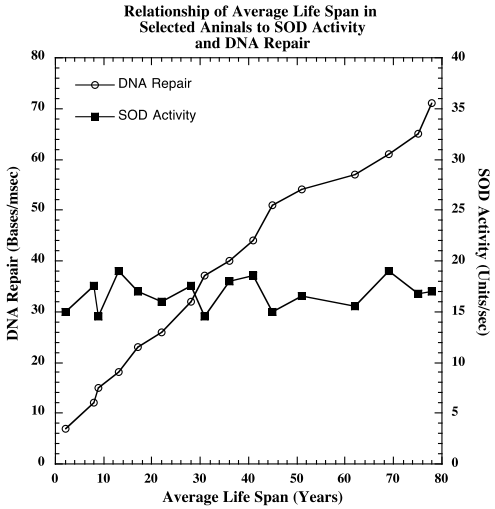


Figure 1. *Animals that live longer have higher SOD activity and correspondingly lower free radical production. This suggests that higher levels of free radical production correlate with shorter life spans. In addition, the ability to destroy the superoxide free radical reflected in higher levels of SOD activity correlates closely with longer life span. Therefore, the graphs suggest that organisms with higher SOD activity and/or lower free radical production will live longer.*

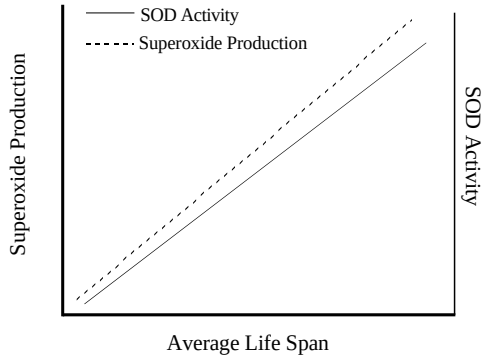
Figure 2. *This graph suggests that increased SOD activity has no influence on the life span of the organisms being monitored. However, an ability to repair DNA more efficiently appears to give organisms a better chance at having longer lives. Some of the DNA damage being repaired by the elevated repair enzymes may be caused by the chemical activity of free radicals.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

7. You isolate superoxide dismutase from two cell culture lines. One line (SOD1) has a level of SOD activity similar to that in liver, the tissue from which the cell line was originally obtained. The other line (SOD10) has elevated SOD activity. The enzyme in SOD10 is very efficient at converting the superoxide free radical to hydrogen peroxide. In a check of other critical enzyme activities, catalase was found to have activity levels that were severely depressed in SOD10, while they appeared normal in SOD1. Observations of SOD10 reveal that this cell line cannot be maintained as easily as SOD1. SOD10 cells appear to die at an accelerated rate. What if anything can you conclude from these data? *While SOD10 is very efficient at neutralizing superoxide free radicals by producing hydrogen peroxide, the peroxide is toxic. SOD10 also has a relatively ineffective catalase, which detoxifies hydrogen peroxide. SOD10 causes hydrogen peroxide to build up rapidly, but lacks the ability to neutralize it just as rapidly. The result is that these cells die at an accelerated rate.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

8. What would a graph similar to those above in question 6 look like if one could conclude from it that organisms that exhibit longer life spans also exhibit proportionately higher production of the superoxide free radical and correspondingly higher levels of SOD activity?



Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

9. What is meant by programmed aging? *"Programmed aging" posits that organisms have evolved mechanisms to induce their own decline after they have passed reproductive age.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

10. What are some common antioxidants found in the body? *Glutathione, vitamins E and C, beta-carotene (the orange pigment in carrots and other vegetables), and the parent compound for vitamin A.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

HUMAN PERSPECTIVES QUESTIONS: PROTEIN MISFOLDING CAN HAVE DEADLY CONSEQUENCES

1. What human disease was found to be similar to kuru in the brain abnormalities it caused? *Creutzfeld-Jakob disease (CJD) is similar to kuru.* What disease in sheep contributes its name to the abnormal prion molecule, PrP^{Sc}? *The disease in sheep that contributes its name to the prion molecule is scrapie.* What have been the causes of outbreaks of acquired CJD? *A number of patients acquired CJD because they ate contaminated beef; this form of the disease was known commonly as "mad cow disease." Acquired CJD has also been seen in recipients of organs and organ products that were donated by a person with undiagnosed CJD. Contaminated beef that the infected individuals had eaten years before has also been implicated as a cause of acquired CJD.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

2. What is spongiform encephalopathy? *This is a pathology in which certain brain regions are riddled with microscopic holes called vacuolations. It causes the tissue to resemble a sponge.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

3. When it was discovered that CJD could be acquired in addition to being inherited, it was at first assumed that the infectious agent was a virus. What led Prusiner to propose that the infectious agent was not a virus? *The infectious agent was found to lack nucleic acid and instead was composed solely of protein.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

4. How was it proved that CJD could be passed to another organism? *Extracts from tissues of diseased individuals can be proved to be infectious if they transmit the disease to another individual. For CJD, this was shown across species with extracts from a brain biopsy of a human CJD victim causing disease in lab animals.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

5. An infectious agent is discovered that causes a particular disease. It has a relatively low molecular weight. Treatment with phenol or proteolytic enzymes, treatments that destroy proteins, render the infectious agent harmless, while treatment with nucleases and ultraviolet radiation, treatments that damage polynucleotides, have no effect. What is your interpretation of the above data and why? *Sensitivity to protein-destroying treatments means the agent contains protein and that the protein is important to the infectious process. The lack of effect of nucleic acid-destroying treatments suggests that nucleic acids are not important for infection and that the infectious agent is not a virus, because nucleic acids are essential when viruses are responsible for an infection. The active part of the infectious agent above is clearly protein and not nucleic acid.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

.6. How was it proved that the brains of patients suffering from CJD, an inherited disease, contain an infectious agent? *Carlton Gajdusek prepared extracts from a biopsy of the brain of a CJD victim. The extract was injected into a suitable laboratory animal. The animal developed a spongiform encephalopathy similar to that of kuru or CJD.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

7. Since replication is a property characteristic of nucleic acids, how might a prion, which lacks nucleic acids, "replicate" itself? *The mutant form of the protein in patients suffering from inherited CJD may act as a template that causes the conformation of the normal protein to convert to the abnormal form. The resultant two abnormal proteins could then convert two others, etc. The conversion of PrP^{C} to PrP^{Sc} has been accomplished in a test tube. Presumably, the appearance of the abnormal protein in the body, by whatever means, starts a chain reaction in which normal protein molecules in the cells are gradually converted to the abnormal prion form.* How can the inherited form of CJD be transmitted to another person? *A person who has the inherited form of CJD could transmit the disease to another person if they donate tissue or blood to a person who does not have the disease. The proteins in the donated tissue could then cause normal proteins in the recipient to shift conformation to the abnormal form. This could eventually lead to clinical CJD.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

8. How was kuru passed from one native of Papua-New Guinea to another? *During a funeral ritual, the mourners would eat the brain tissue of recently deceased relatives. If the deceased had suffered from kuru, the disease could, and often would, be passed from the deceased relative to the mourners.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

9. What is the derivation of the name prion for the agent that can transmit diseases like CJD and kuru*? Disease transmission is by "**prot**ein **on**ly".*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

10. Knockout mice are mice that have had one specific gene removed from their genome. This allows the role of the missing gene and its protein product to be assessed. Given this information, how would you explain the inability of mouse scrapie prions, which cause a malady similar to CJD, to cause the CJD-like disease scrapie in PrP knockout mice? *Since PrP knockout mice lack the PrP^{C} protein, there are no normal proteins in these mice to be converted to the mutant form; thus they do not develop the disease. With no PrP protein at all, normal or abnormal, the mice can still survive, since there appears to be no adverse effect, if the protein is missing.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

11. What physical properties of the abnormal form of the PrP protein probably account for its ability to cause CJD? *Normal PrP (PrP^C) is monomeric, soluble in salt solutions and readily destroyed by proteolytic enzymes. Abnormal PrP^{Sc} molecules are able to interact with each other to form insoluble fibrils that are resistant to enzymatic digestion.* What is odd about these differences given what is known about the structures of the two proteins? *The two proteins can have the same amino acid sequence, but fold up differently to form significantly different three-dimensional structures. PrP^C consists almost entirely of α-helical segments and interconnecting coils. In contrast, PrP^{Sc} consists largely of β-pleated sheet. The shift from a soluble, protease-sensitive conformation to an insoluble, protease-insensitive aggregate can be accomplished* in vitro.

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

12. How does inherited CJD lead to the production of abnormal forms of PrP? *Under normal circumstances, the newly synthesized PrP polypeptide almost invariably folds into the PrP^C conformation. People with inherited CJD have a gene that encodes a mutant protein whose amino acid sequence is different from that of the normal protein. The mutant protein is presumed to be less stable in the PrP^C conformation than the normal version of the protein and more likely to fold into the abnormal β-pleated sheet-rich conformation. Once formed, the β-rich proteins produce aggregates, which lead to disease.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

13. In what ways are CJD and Alzheimer's disease similar? *Both are fatal neurodegenerative diseases that can occur in an inherited or sporadic form. Brains of CJD and Alzheimer's disease patients contain fibrillar deposits of an insoluble material. In both diseases, toxic fibrillar deposits result from self-association of a protein primarily made of β-pleated sheet.* What are the differences between the two diseases? *The proteins that form the disease-causing aggregates are completely unrelated. The parts of the brain that are affected are distinct and the protein responsible for Alzheimer's disease does not act like an infectious agent; it is nontransmissable from one person to another, although it may spread from cell to cell in the brain.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

14. What surprising potential treatment for Alzheimer's disease has been demonstrated in a mouse animal model for the disease? I*nvestigators created transgenic mice that developed amyloid brain plaques by introducing a mutant gene for human amyloid precursor protein (APP). They were able to block amyloid plaque formation by repeatedly injecting the animals with the same substance that causes the problem, the Aβ42 peptide. This caused the animals to produce antibodies against the peptides made in the brains of the mice by cleavage of the mutant APP protein. They were immunizing the animals against the disease. If the mice were injected when they were younger, they did not develop the amyloid deposits. If older mice whose brains already contained deposits were injected, a significant fraction of the deposits was cleared out of the nervous system.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

15. What approaches other than immunization are being developed as treatments for Alzheimer's disease? *One approach now being tried, instead of immunization, is administration of antibodies directed against Aβ that have been produced outside the body; this approach is known as passive immunization because the patient does not produce the therapeutic antibodies. Passive immunization with an anti-Aβ42 antibody (bapineuzumab) has already proven capable of restoring memory function in transgenic mice. It has been shown to safe and effective in Phase I and II clinical trials. Phase III trials have, however, been disappointing. Another antibody that recognizes a different part of the Aβ protein (Solanezumab) has shown more favorable results in Phase III trials. Investigators are also looking into initiating antibodies much earlier to patients who are thought to be at risk for developing early-onset AD. These are preventive trials. Drugs are being developed that bind Aβ peptides and block molecular aggregation and fibril formation (Alzhemed and scyllo-inositol). The effort here is to stop the production of Aβ peptides. They are also looking into inhibition of the β- or γ-secretases. These enzymes cleave the* APP *precursor to release the Aβ peptides. They are also investigating the tau protein as apossible cause of AD and a target of treatment. Other drugs being tried (methylthioninium chloride and a modification of that drug leuco mwthylioninium) appear to dissolve neurofibrillary tangles (NFTs).*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

Experimental Pathways Questions: Chaperones — Helping Proteins Reach Their Proper Folded State

1. Why would heat shock proteins be present at low levels in cells or organisms raised at normal temperatures and then increase in number after a brief exposure to elevated temperatures? *At lower, normal temperatures, proteins would occasionally denature during the course of cell metabolism. Thus, a small number of these heat shock proteins would be enough to help them renature. At higher temperatures, proteins would probably denature at an accelerated rate. Therefore an increase in the number of the heat shock proteins will help the cell or organism survive the elevated temperature.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

2. When a protein denatures in a cell, what causes it to aggregate with other denatured cell proteins? *When proteins in the cell are denatured their interior hydrophobic amino acids are exposed to the aqueous cytoplasm of the cell. If the protein encounters other proteins that have become denatured in the same fashion, their hydrophobic amino acids are likely to aggregate rather than interact with the aqueous environment.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

3. Two seemingly unrelated lines of investigation eventually converged to lead to a fuller understanding of molecular chaperones. What were those lines of investigation and what discovery unified them into one field, the study of molecular chaperones? *The two lines of investigation were the studies of the heat-shock response and the proteins that promote protein assembly. The two fields came together in 1986, when it was shown that one of the proteins that figured most prominently in the heat-shock response, a protein that had been named heat-shock protein 70 (hsp70) because of its molecular mass, was identical to BiP, the protein implicated in the assembly of antibody molecules.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

4. You are working with the enzyme maltase that you obtain in a cell-free extract by homogenizing the cells that normally contain it. You heat the extract, a treatment known to denature the protein. When you cool the extract, the enzyme apparently renatures since most of its activity, which had disappeared when the extract was heated, returns rapidly. It appears that maltase is capable of self-assembly. Does this prove that maltase can renature without any assistance from other molecules like molecular chaperones? If not, describe an experiment that could prove that maltase is capable of self-assembly in the absence of molecular chaperones. *No, the experiment does not prove that the enzyme can renature without help from molecules like molecular chaperones. It is possible that some molecular chaperones were extracted along with maltase and that they aid the renaturation. If you purify maltase to homogeneity, there will be no other proteins present. If the completely purified protein renatures by itself after it has been denatured, one would conclude that no molecular chaperone is needed.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

5. The bacterial GroEL heat-shock protein and the plant Rubisco assembly proteins are said to be homologous proteins. What is it about these proteins that has caused them to be described as members of the same protein family and suggests that they have the same function? *The two proteins share the same amino acids at nearly half of the more than 500 amino acid residues in their respective molecules. The fact that they have retained so many of the same amino acids reflects their similar and essential function in the two types of cell.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

6. A molecular chaperone named GroEL is found in the cytoplasm of the bacterium *E. coli*. An homologous protein, the function of which is thought to aid in Rubisco assembly, is found in plant chloroplasts. Why would it not be surprising to find a protein so similar to GroEL, a prokaryotic protein, in these eukaryotic cell organelles? *According to the Endosymbiotic Theory, chloroplasts are derived from ingested photosynthetic bacteria that were not digested. It would not, therefore, be surprising to find a protein in chloroplasts that is similar to a protein found in prokaryotes like* E. coli.

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

7. Describe the evidence that convinced investigators that chaperones do not assist the assembly of already-folded subunits into larger complexes but that they instead assist polypeptide chain folding. *It has been known for some time that newly-made mitochondrial proteins produced in the cytosol have to cross the outer mitochondrial membrane in an unfolded, extended, monomeric form. During a study of molecular chaperones in mitochondria, a mutant was discovered that altered the activity of another member of the Hsp60 chaperone family that resided inside mitochondria. In cells containing this mutant chaperone, proteins that were transported into mitochondria failed to fold into their native conformation. Even proteins consisting of a single polypeptide chain failed to fold into its native conformation. This finding changed the perception of chaperone function from the notion that they assist assembly of already-folded subunits into larger complexes, to the current understanding that they assist polypeptide chain folding within the crowded confines of the cell.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

8. What are the two major known families of molecular chaperones? *The two major families of molecular chaperones are the Hsp70 chaperones, like BiP and the Hsp60 chaperones, also called chaperonins, like Hsp60, GroEL and the Rubisco assembly protein. The Hsp60 chaperones like GroEL are the best understood.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

9. The GroEL complex is a chaperone protein that is shaped like a cylinder with a central cavity large enough to enclose a polypeptide undergoing folding. The central cavity is lined by a ring of hydrophobic residues. Why are non-native polypeptides able to bind to such a structure and what function(s) might that binding serve? *Since non-native polypeptides are denatured, their hydrophobic groups are exposed and will bind to the hydrophobic wall in the center of the GroEL cylinder. The binding of these proteins in the central cavity may remove a non-native polypeptide from an environment where it is likely to become aggregated. It may also unfold a polypeptide that has become misfolded, thus giving it a chance to refold properly.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

10. What is the apparent purpose of the GroES subunit binding to the GroEL cylinder when it is occupied by a nonnative polypeptide? *GroES binds to the end of the GroEL cylinder and causes a conformational shift in the molecule. The shift increases the volume of the enclosed chamber. In addition, the conformational shift buries the hydrophobic residues on the GroEL wall and exposes a number of polar residues. The non-native polypeptide is then released from the GroEL wall and displaced into the newly enlarged space where it can continue its folding in a protected environment. After about 15 sec, the GroES cap dissociates from the GroEL ring, and the polypeptide is ejected from the chamber. If the polypeptide has not attained its native conformation by the time it is ejected, it can rebind to the same or another GroEL, and the process is repeated.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

11. Why can it be said that a protein that folds inside the GroEL protein self-assembles? *It can be said to self-assemble because the GroEL chaperone only provides a safe location where folding can occur. The protein contains within its sequence all the information it needs to adopt its proper conformation.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

12. As many as 10% of the nonnative soluble proteins of a bacterial cell can interact with GroEL. Given the fact that interactions between proteins are often highly specific, how is it possible that a single protein, like GroEL can bind to so many different polypeptides? *The GroEL binding site consists of a hydrophobic surface formed largely by two α-helices of the apical domain. This portion of the molecule is capable of binding virtually any sequence of hydrophobic residues that might be accessible in a partially folded or misfolded polypeptide. Comparison of the crystal structures of unbound GroEL and GroEL bound to several different peptides revealed that the binding site on the apical domain of a GroEL subunit can locally adjust its positioning when bound to different partners. This suggests that the binding site has structural flexibility that allows it to adjust its own shape to fit the shape of the particular polypeptide with which it has to interact.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

13. What experimental procedure has suggested that binding of a nonnative protein to GroEL is followed by a forced unfolding of the substrate protein? *FRET (fluorescence resonance energy transfer) is a technique that allos researchers to determine the distance between different parts of a protein molecule at different times during a given process. FRET tags on the refolding protein revealed that the protein experienced some unfolding before refolding. The distance between the tags increased after the GroEL conformational shift as the protein denatured. Initially, when the protein first entered the GroEL ring, the FRET tags were in close proximity, indicating that the bound protein was compact in nature.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

ART QUESTIONS

1. Which atom(s) in Figure 2.1 are the least reactive and why? *Neon and argon are the least reactive atoms because they have full outer electron shells. Consequently, they are called inert (noble) gases.*

Learning Objective: 2.1 Describe the role of electrons in the formation of covalent bonds.
Section Reference: 2.1

2. In Figure 2.3, you see a drawing of a salt crystal held together by ionic bonds. Are ionic bonds plentiful in living organisms? Why or why not? *Ionic bonds are not very common in living organisms because their cells contain so much water, which interferes with such bonds. They can exist in areas of a living cell that restrict exposure to water, e.g., in the interior of proteins where hydrophobic R groups congregate.*

Learning Objective: 2.3 Describe the role of noncovalent bonds in the structure of molecules such as water.
Section Reference: 2.3

3. If H bonds are about 180 picometers long (Figure 2.5) and the strongest attraction between molecules participating in a single van der Waals interaction occurs when the molecules involved are separated by about 3.6 Å (Figure 2.7), which interaction is the strongest? *Since longer bonds are weaker as a general rule, a single H bond would be somewhat stronger than a single van der Waals interaction.*

Learning Objective: 2.3 Describe the role of noncovalent bonds in the structure of molecules such as water.
Section Reference: 2.3

4. a. Figure 2.7 exhibits the effect of distance on the attraction between two atoms. How would you describe the attraction or repulsion of two such atoms at a separation distance of 8 Å? *At a separation distance of 8 Å, there is virtually no attraction between two atoms.* What about at 4.5 Å? *At 4.5 Å, there is a slight attraction between two atoms.* How can you tell from this graph the distance at which the optimal attraction between the two atoms occurs? *The optimal attraction between the two atoms occurs at a distance of about 3.6 Å. This would be the separation distance on the x-axis corresponding to the point on the graph that projects farthest below 0 on the y-axis.*

   b. It is not unusual for a mutation to disrupt interactions such as these significantly. In these cases, one amino acid is often substituted for another. For example, a change in one amino acid in the hemoglobin β chains leads to the molecular shape changes that cause sickle cell anemia. How could such a change eliminate van der Waals interactions, such as those illustrated in Figure 2.7? *A mutation making a significant change in the polypeptide chain, for example a hydrophobic amino acid residue exchanged for a polar, charged residue, would be likely to change the secondary and/or the tertiary structure of the protein significantly. This might move normally adjacent parts of the molecule farther apart. If the distance between these two parts of the molecule were increased by 2-3 Å, the effect would be large enough to abolish the van der Waals attractions completely or nearly so.*

Learning Objective: 2.3 Describe the role of noncovalent bonds in the structure of molecules such as water.
Section Reference: 2.3

5. Figure 2.9 shows the interaction between two subunits of a clam hemoglobin molecule. What molecule plays an important role in the interaction between the two subunits? *Water plays an important role in the interaction between the two depicted subunits of clam hemoglobin; it occupies highly ordered locations between the two subunits.*

Learning Objective: 2.3 Describe the role of noncovalent bonds in the structure of molecules, such as water.
Section Reference: 2.3

6. Figure 2.10 shows a ball and stick drawing of the structure of cholesterol. Are there any functional groups on this molecule that could allow even the slightest interaction with water? *Yes.* If so, what are they? *There is a hydroxyl group on the left side of the first ring in the figure. It could form H bonds with water.* If cholesterol is present in a membrane, which end is most likely to be directed to the outer surfaces of the membrane that are closer to the hydrophilic environments of the cell cytoplasm or the extracellular space? *The end with the hydroxyl group would be most likely to be exposed to the polar heads of the membrane phospholipids and the hydrophilic environments surrounding the membrane.*

Learning Objective: 2.5 Describe the general structure and functions of biological molecules.
Section Reference: 2.5

7. a. Figure 2.11 demonstrates that water is released as a byproduct during condensation reactions and reintroduced across the same bond during hydrolysis reactions, resulting in the breakage of the bond. If 1000 glucose molecules ($C_6H_{12}O_6$) were hooked together by condensation reactions, how many glycosidic bonds would be found in the resulting polymer? *999 glycosidic bonds.* After the formation of this polysaccharide, how many carbon, hydrogen and oxygen atoms are found in the polymer? *There would be 6,000 carbons, 12,000 hydrogens and 6,000 oxygens in 1000 glucose molecules. Since 999 glycosidic linkages would connect the 1,000 glucose molecules, 999 water molecules would be removed from the entire structure (1,998 hydrogens and 999 oxygens). Therefore, in the polysaccharide, there should be 6,000 carbons, 10,002 hydrogens, and 5,001 oxygens.*

   b. If a protein consists of 454 amino acids, how many hydrolysis reactions would be required to fully degrade the protein? *453 hydrolysis reactions.*

Learning Objective: 2.5 Describe the general structure and functions of biological molecules.
Section Reference: 2.5

8. Some people are born with or can develop a condition known as lactose intolerance that causes them to suffer intestinal discomfort when they eat lactose-containing dairy products. This occurs because the lactose that can normally be metabolized and passed through the intestinal lining cannot do so in these individuals. Can you suggest an explanation for this? The structure of lactose is shown in Figure 2.18. *People affected by lactose intolerance lack the enzyme that breaks the bond between glucose and galactose, the two sugars that combine to form the disaccharide. Therefore, lactose remains in the intestine leading to the symptoms of lactose intolerance.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.
Section Reference: 2.7

9. In Figure 2.19, there are schematic drawings of glycogen, starch, and cellulose.
   a. Which of these polysaccharides would be most likely to allow the quickest release of glucose monomers during hydrolysis? *Glycogen.* Why? *It's branched and possesses more free ends from which glucose can be released.*
   b. Which polysaccharide(s) could be used as a fuel source by an organism that lacks enzymes that break α-glycosidic linkages, but possesses enzymes that can break β-glycosidic bonds? *Cellulose*.
   c. What kind of bond is marked by the number 3 in the figure? *β (1→ 4) glycosidic bond.*
   d. What kind of bond is denoted by the number 1 in the figure? *α (1→6) glycosidic bond.* What feature of the molecule in question requires this bond? *This bond is required for branching.* What is (are) advantages of this feature? *The advantages are more efficient packing of more glucose residues in a smaller space and more free ends on the molecule to facilitate more efficient and rapid release of glucose monomers when they are needed.* If this bond could not be formed, what would the molecule look like? *Without this bond, the molecule would be linear with no branching. All monomer glucose units would be connected with α (1→ 4) glycosidic bonds. It would like the starch in Figure 2.19b.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

10. Which molecule in Figure 2.21 contains double bonds in at least some of its fatty acid chains? *Linseed oil.* Which molecule contains no double bonds in its fatty acids? *Tristearate.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

11. Is the phospholipid in Figure 2.24, saturated or polyunsaturated? How do you know? *It is saturated because the fatty acid tails contain no double bonds. Also, the chains are straight; they would be kinked if they were unsaturated. A kink would occur at each double bond.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

12. Which amino acid in Figure 2.28 would be most likely to form covalent bonds between two different polypeptide chains? *Cysteine.* Which amino acid would be most likely to be found at a kink in an amino acid chain? *Proline.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.

Section Reference: 2.7

13. In Figure 2.31, there is a scanning electron micrograph of a sickled red blood cell. If hemoglobin were isolated from this cell and others like it and subjected to chromatographic separation after enzymatic digestion, one spot, representing a single peptide, would differ on the chromatograms of normal and sickle cell hemoglobin. How many amino acids are changed in the mutant form of hemoglobin to cause the difference in the two chromatograms? *The single affected peptide has a one amino acid difference between the normal and mutant forms of hemoglobin. A glutamic acid at position 6 in β-globin is replaced with valine.*

Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.
Section Reference: 2.7

14. Which of the structures shown in Figure 2.32 contains H bonds oriented perpendicular to the molecule's axis? *None of them. All of the structures are α-helix, the H bonds of which are oriented parallel to the molecule's axis.*
Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.
Section Reference: 2.7

15. Which of the structures shown in Figure 2.33 contains H bonds oriented parallel to the molecule's axis? *None of them. All of the structures are β-pleated sheet or simply β-sheet, the H bonds of which are oriented perpendicular to the molecule's axis.*
Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.
Section Reference: 2.7

16. The denaturation of ribonuclease is depicted in Figure 2.45. What role in denaturation is played by β-mercaptoethanol? *β-mercaptoethanol breaks disulfide bonds making denaturation occur more readily.*
Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.
Section Reference: 2.7

17. In Figure 2.56a, what kind of nitrogenous base appears in the drawing? To which carbon of the nucleotide sugar is it attached? *It is a purine base, specifically adenine; it is attached to the 1'- carbon of the nucleotide sugar.* In Figure 2.56b, which end of the polynucleotide shown is the 5' end and which the 3' end? *The end of the polynucleotide nearest to the top of the page is the 5' end. The end nearest to the bottom of the page is the 3' end.*
Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.
Section Reference: 2.7

18. Consult Figure 2.57. What is the difference between the pyrimidine nitrogenous bases uracil and thymine? *Thymine differs from uracil by having a methyl group attached to the ring, instead of a hydrogen at one position.*
Learning Objective: 2.7 Identify the monomers, the synthesis, and the functions in the cell of the four types of biological molecules.
Section Reference: 2.7

2 The Chemical Basis of Life

CASE STUDY: Defects in Hemoglobin Structure and Function

Hemoglobin is the major oxygen carrier that is used to deliver oxygen to our tissues. It is a heterotetrameric protein that is composed of two alpha subunits and two beta subunits. Each subunit has the ability to bind and release oxygen, and its ability to do so is influenced by the structure of the other subunits. Defects in hemoglobin structure or synthesis are collectively termed hemoglobinopathies. This group of diseases results from defects in the synthesis of one of the hemoglobin chains or in defects in the structure of the hemoglobin molecule itself. Patients with defective hemoglobin have characteristic anemia, which leads to pallor, fatigue, and shortness of breath. Other clinical manifestations include reticulocytosis (elevation of the number of young red blood cells), splenomegaly (enlarged spleen), and urobilinuria (excess urobilins, which are breakdown products of hemoglobin, in the urine).

Sickle Cell Anemia is a specific type of hemoglobinopathy caused by mutation of a single glutamic acid residue on the surface of hemoglobin to a valine, which results in a change in the surface properties of hemoglobin. This mutant hemoglobin is referred to as HbS. Presence of HbS causes protein aggregation under conditions of deoxygenation. The protein aggregates lead to malformed red blood cells that inhibit capillary flow.

Questions:

1. The mutation in hemoglobin is a change from a glutamic acid to a valine. What are the chemical features of these two amino acids that may result in the defects caused by HbS?

Answer: Glutamic acid has a carboxylic group as a side chain. This side chain is normally negatively charged at physiological pH. In contrast, valine has a non-polar side chain that is hydrophobic.

2. According to the principles of the hydrophobic effect, where should glutamic acid and valine normally be found in proteins?

Answer: Glutamic acid is normally charged and therefore can exist on the outside surface of proteins where it can interact with water. This makes sense because the glutamic acid is normally on the surface of the hemoglobin molecule. In contrast, valine should be found on the interior of proteins because it wants to be shielded from polar water molecules.

3. How then do you think that the mutated valine residue can contribute to the aggregation of hemoglobin molecules in HbS?

Answer: The valine residue is found on the surface of each of the 4 subunits of hemoglobin. These valine residues are then attracted to the valine residues on other hemoglobin molecules so that the hemoglobin

molecules tend to stick together and become non-functional. In reality the situation is even more complicated because based on conformational changes in the hemoglobin molecule the valine residues are only exposed on the surface of deoxygenated hemoglobin molecules. This leads to a cycle in which deoxygenated hemoglobins aggregate and begin to clog the blood vessels, which in turn leads to a decrease in oxygen transport by any oxygen-bound hemoglobin molecules.

Where can I learn more?

1. Schechter, A.N., Hemoglobin research and the origins of molecular medicine. Blood, 2008. 112: p. 3927-3938.
2. Berg, J.M, Tymoczko, J. and L. Stryer, Biochemistry, Sixth Edition, 2007. Chapter 7. Hemoglobin: A portrait of a protein in action. W.H. Freeman Publisher.

2 The Chemical Basis of Life

CASE STUDY: Protein Conformational Diseases

Newly synthesized polypeptides must be properly folded in order to achieve their proper three-dimensional protein structure. Defects in the folding process can lead to improper intramolecular and intermolecular interactions including the formation of protein clusters, or aggregates. These aggregates can become large enough to disrupt cell and tissue function.

Huntington's Disease (HD) is a neurological disease that is caused by a mutation in the huntingtin gene. It is an autosomal dominant disorder, meaning that if a person has just one mutant copy of the huntingtin allele, they will be at much higher risk of having the disease and of passing it along to their children. Unlike many genetic diseases, HD does not arise from point mutations in a gene that cause a loss of function. Instead, it is a member of a family of diseases known as the trinucleotide repeat disorders. The huntingtin gene *(HTT)* contains a repeated section of the sequence CAG, which codes for a long stretch of the amino acid glutamine. This is known as a polyglutamine, or "polyQ" tract. In HD patients, the polyQ tract of the protein huntingtin (HTT) is much longer than in people who are not at risk for the disease. The glutamine side chains do not fold properly and become sticky to each other and to other proteins, which leads to clustering of the protein, precipitous aggregates, and ultimately cellular dysfunction. Major symptoms of the disease, including involuntary movements and dementia, are consistent with neurodegeneration from defective deposits in brain cells.

One challenge to the study of studying human neurological diseases is the requirement for good model systems to ask questions about the disease. A useful model can often be found in animal systems, such as the mouse, where genes can either be overexpressed or deleted. The questions below are based on several mouse models that have been developed to look at HD.

Questions:

1. Mice that are heterozygous for the *HTT* null mutation do not display Huntington's disease. (This means that they have only one copy of the gene encoding huntingtin, and it is the normal, wild type *HTT* allele.) What does this tell you about the mechanism by which the normal HTT protein functions versus the mHTT mutant protein (with expanded polyQ tract at its N-terminus)?

Answer: Because heterozygous HTT null mice display no evidence of HD suggests that the mice do not undergo neurodegeneration and disease if they express less of the HTT protein. This must mean that the expression of the abnormal HTT protein with the expanded glutamine stretch is what causes the cellular dysfunction and neuronal degeneration associated with the disease.

2. Scientists have generated a "conditional" model for HD in mice. In this study the researchers generated a mouse model where they could selectively induce the expression of an abnormal version of the HTT protein that contains the N-terminus of HTT with a very long polyQ tract

(94 CAG repeats). These mice developed neuropathological and behavioral defects associated with HD. Does this result support or refute what you answered in Question 1 about the mechanism by which the HTT protein causes cellular dysfunction?

Answer: This result supports the answer in Question 1 because overexpression of the truncated version of the protein containing only a polyQ tract also causes cellular dysfunction. This result means that it is the presence of excess glutamines in the htt protein that is responsible for the defects in HD.

3. One very exciting finding from this conditional mouse model was that if they turned off the expression of the mutant form of the truncated HTT protein, they could halt progression of the disease and even reverse aggregate formation and progressive motor decline. What do these results imply about the mechanism of the function of the abnormal HTT protein, and what are the implications for therapy for this disease?

Answer: These results imply that the abnormal HTT protein must be continuously expressed for the defects associated with HD to occur. This suggests that drug targets or therapy aimed at degrading or sequestering abnormal HTT protein should be effective in reversing the symptoms of HD.

Where can I learn more?

1. Dickey AS, La Spada AR. Therapy development in Huntington disease: From current strategies to emerging opportunities. *Am J Med Genet A*. 2018;176(4):842–861. doi:10.1002/ajmg.a.38494