chapter

# 1

# The Foundations of Biochemistry

**1. The Size of Cells and Their Components**

**(a)** If you were to magnify a cell 10,000-fold (typical of the magnification achieved using an electron microscope), how big would it appear? Assume you are viewing a "typical" eukaryotic cell with a cellular diameter of 50 $\mu$m.

**(b)** If this cell were a muscle cell (myocyte), how many molecules of actin could it hold? (Assume the cell is spherical and no other cellular components are present; actin molecules are spherical, with a diameter of 3.6 nm. The volume of a sphere is $4/3\ \pi r^3$.)

**(c)** If this were a liver cell (hepatocyte) of the same dimensions, how many mitochondria could it hold? (Assume the cell is spherical; no other cellular components are present; and the mitochondria are spherical, with a diameter of 1.5 $\mu$m.)

**(d)** Glucose is the major energy-yielding nutrient for most cells. Assuming a cellular concentration of 1 mM, calculate how many molecules of glucose would be present in our hypothetical (and spherical) eukaryotic cell. (Avogadro's number, the number of molecules in 1 mol of a nonionized substance, is $6.02 \times 10^{23}$.)

**(e)** Hexokinase is an important enzyme in the metabolism of glucose. If the concentration of hexokinase in our eukaryotic cell is 20 $\mu$M, how many glucose molecules are present per hexokinase molecule?

**Answer**

**(a)** The magnified cell would have a diameter of $50 \times 10^4\ \mu\text{m} = 500 \times 10^3\ \mu\text{m} = 500$ mm, or 20 inches—about the diameter of a large pizza.

**(b)** The radius of a globular actin molecule is 3.6 nm/2 = 1.8 nm; the volume of the molecule, in cubic meters, is $(4/3)(3.14)(1.8 \times 10^{-9}\ \text{m})^3 = 2.4 \times 10^{-26}\ \text{m}^3$.*

The number of actin molecules that could fit inside the cell is found by dividing the cell volume (radius = 25 $\mu$m) by the actin molecule volume. Cell volume = $(4/3)(3.14)(25 \times 10^{-6}\ \text{m})^3 = 6.5 \times 10^{-14}\ \text{m}^3$. Thus, the number of actin molecules in the hypothetical muscle cell is

$$(6.5 \times 10^{-14}\ \text{m}^3)/(2.4 \times 10^{-26}\ \text{m}^3) = 2.7 \times 10^{12}\ \text{molecules}$$

or 2.7 trillion actin molecules.

**Significant figures:* In multiplication and division, the answer can be expressed with no more significant figures than the least precise value in the calculation. Because some of the data in these problems are derived from measured values, we must round off the calculated answer to reflect this. In this first example, the radius of the actin (1.8 nm) has two significant figures, so the answer (volume of actin = $2.4 \times 10^{-26}\ \text{m}^3$) can be expressed with no more than two significant figures. It will be standard practice in these expanded answers to round off answers to the proper number of significant figures.

**(c)** The radius of the spherical mitochondrion is 1.5 $\mu$m/2 = 0.75 $\mu$m, therefore the volume is $(4/3)(3.14)(0.75 \times 10^{-6}\ \text{m})^3 = 1.8 \times 10^{-18}\ \text{m}^3$. The number of mitochondria in the hypothetical liver cell is

$$(6.5 \times 10^{-14}\ \text{m}^3)/(1.8 \times 10^{-18}\ \text{m}^3) = 36 \times 10^3\ \text{mitochondria}$$

**(d)** The volume of the eukaryotic cell is $6.5 \times 10^{-14}\ \text{m}^3$, which is $6.5 \times 10^{-8}\ \text{cm}^3$ or $6.5 \times 10^{-8}$ mL. One liter of a 1 mM solution of glucose has (0.001 mol/1000 mL)$(6.02 \times 10^{23}$ molecules/mol) $= 6.02 \times 10^{17}$ molecules/mL. The number of glucose molecules in the cell is the product of the cell volume and glucose concentration:

$$(6.5 \times 10^{-8}\ \text{mL})(6.02 \times 10^{17}\ \text{molecules/mL}) = 3.9 \times 10^{10}\ \text{molecules}$$

or 39 billion glucose molecules.

**(e)** The concentration ratio of glucose/hexokinase is 0.001 M/0.00002 M, or 50/1, meaning that each enzyme molecule would have about 50 molecules of glucose available as substrate.

**2. Components of *E. coli*** *E. coli* cells are rod-shaped, about 2 $\mu$m long and 0.8 $\mu$m in diameter. The volume of a cylinder is $\pi r^2 h$, where $h$ is the height of the cylinder.

**(a)** If the average density of *E. coli* (mostly water) is $1.1 \times 10^3$ g/L, what is the mass of a single cell?

**(b)** *E. coli* has a protective cell envelope 10 nm thick. What percentage of the total volume of the bacterium does the cell envelope occupy?

**(c)** *E. coli* is capable of growing and multiplying rapidly because it contains some 15,000 spherical ribosomes (diameter 18 nm), which carry out protein synthesis. What percentage of the cell volume do the ribosomes occupy?

**Answer**

**(a)** The volume of a single *E. coli* cell can be calculated from $\pi r^2 h$ (radius = 0.4 $\mu$m):

$3.14(4 \times 10^{-5}\ \text{cm})^2(2 \times 10^{-4}\ \text{cm}) = 1.0 \times 10^{-12}\ \text{cm}^3 = 1 \times 10^{-15}\ \text{m}^3 = 1 \times 10^{-15}\ \text{L}$

Density (g/L) multiplied by volume (L) gives the mass of a single cell:

$$(1.1 \times 10^3\ \text{g/L})(1 \times 10^{-15}\ \text{L}) = 1 \times 10^{-12}\ \text{g}$$

or a mass of 1 pg.

**(b)** First, calculate the proportion of cell volume that does *not* include the cell envelope, that is, the cell volume *without* the envelope—with $r = 0.4\ \mu\text{m} - 0.01\ \mu\text{m}$; and $h = 2\ \mu\text{m} - 2(0.01\ \mu\text{m})$—divided by the total volume.

$$\text{Volume without envelope} = \pi(0.39\ \mu\text{m})^2(1.98\ \mu\text{m})$$

$$\text{Volume with envelope} = \pi(0.4\ \mu\text{m})^2(2\ \mu\text{m})$$

So the percentage of cell that does *not* include the envelope is

$$\frac{\pi(0.39\ \mu\text{m})^2(1.98\ \mu\text{m}) \times 100}{\pi(0.4\ \mu\text{m})^2(2\ \mu\text{m})} = 90\%$$

(Note that we had to calculate to one significant figure, rounding down the 94% to 90%, which here makes a large difference to the answer.) The cell envelope must account for 10% of the total volume of this bacterium.

**(c)** The volume of all the ribosomes (each ribosome of radius 9 nm) $= 15{,}000 \times (4/3)\pi(9 \times 10^{-3}\ \mu\text{m})^3$

The volume of the cell $= \pi(0.4\ \mu\text{m})^2(2\ \mu\text{m})$

So the percentage of cell volume occupied by the ribosomes is

$$\frac{15{,}000 \times (4/3)\pi(9 \times 10^{-3}\ \mu\text{m})^3 \times 100}{\pi(0.4\ \mu\text{m})^2(2\ \mu\text{m})} = 5\%$$

**3. Genetic Information in *E. Coli* DNA** The genetic information contained in DNA consists of a linear sequence of coding units, known as codons. Each codon is a specific sequence of three deoxyribonucleotides (three deoxyribonucleotide pairs in double-stranded DNA), and each codon codes for a single amino acid unit in a protein. The molecular weight of an *E. coli* DNA molecule is about $3.1 \times 10^9$ g/mol. The average molecular weight of a nucleotide pair is 660 g/mol, and each nucleotide pair contributes 0.34 nm to the length of DNA.

**(a)** Calculate the length of an *E. coli* DNA molecule. Compare the length of the DNA molecule with the cell dimensions (see Problem 2). How does the DNA molecule fit into the cell?

**(b)** Assume that the average protein in *E. coli* consists of a chain of 400 amino acids. What is the maximum number of proteins that can be coded by an *E. coli* DNA molecule?

**Answer**

**(a)** The number of nucleotide pairs in the DNA molecule is calculated by dividing the molecular weight of DNA by that of a single pair:

$$(3.1 \times 10^9 \text{ g/mol})/(0.66 \times 10^3 \text{ g/mol}) = 4.7 \times 10^6 \text{ pairs}$$

Multiplying the number of pairs by the length per pair gives

$$(4.7 \times 10^6 \text{ pairs})(0.34 \text{ nm/pair}) = 1.6 \times 10^6 \text{ nm} = 1.6 \text{ mm}$$

The length of the cell is 2 $\mu$m (from Problem 2), or 0.002 mm, which means the DNA is (1.6 mm)/(0.002 mm) = 800 times longer than the cell. The DNA must be tightly coiled to fit into the cell.

**(b)** Because the DNA molecule has $4.7 \times 10^6$ nucleotide pairs, as calculated in (a), it must have one-third this number of triplet codons:

$$(4.7 \times 10^6)/3 = 1.6 \times 10^6 \text{ codons}$$

If each protein has an average of 400 amino acids, each requiring one codon, the number of proteins that can be coded by *E. coli* DNA is

$$(1.6 \times 10^6 \text{ codons})(1 \text{ amino acid/codon})/(400 \text{ amino acids/protein}) = 4{,}000 \text{ proteins}$$

**4. The High Rate of Bacterial Metabolism** Bacterial cells have a much higher rate of metabolism than animal cells. Under ideal conditions some bacteria double in size and divide every 20 min, whereas most animal cells under rapid growth conditions require 24 hours. The high rate of bacterial metabolism requires a high ratio of surface area to cell volume.

**(a)** Why does surface-to-volume ratio affect the maximum rate of metabolism?

**(b)** Calculate the surface-to-volume ratio for the spherical bacterium *Neisseria gonorrhoeae* (diameter 0.5 $\mu$m), responsible for the disease gonorrhea. Compare it with the surface-to-volume ratio for a globular amoeba, a large eukaryotic cell (diameter 150 $\mu$m). The surface area of a sphere is $4\pi r^2$.

**Answer**

**(a)** Metabolic rate is limited by diffusion of fuels into the cell and waste products out of the cell. This diffusion in turn is limited by the surface area of the cell. As the ratio of surface area to volume decreases, the rate of diffusion cannot keep up with the rate of metabolism within the cell.

**(b)** For a sphere, surface area = $4\pi r^2$ and volume = 4/3 $\pi r^3$. The ratio of the two is the surface-to-volume ratio, $S/V$, which is $3/r$ or $6/D$, where $D$ = diameter. Thus, rather than calculating $S$ and $V$ separately for each cell, we can rapidly calculate and compare $S/V$ ratios for cells of different diameters.

$$S/V \text{ for } N.\ gonorrhoeae = 6/(0.5\ \mu\text{m}) = 12\ \mu\text{m}^{-1}$$

$$S/V \text{ for amoeba} = 6/(150\ \mu\text{m}) = 0.04\ \mu\text{m}^{-1}$$

$$\frac{S/V \text{ for bacterium}}{S/V \text{ for amoeba}} = \frac{12\mu\text{m}^{-1}}{0.04\ \mu\text{m}^{-1}} = 300$$

Thus, the surface-to-volume ratio is 300 times greater for the bacterium.

5. **Fast Axonal Transport** Neurons have long thin processes called axons, structures specialized for conducting signals throughout the organism's nervous system. Some axonal processes can be as long as 2 m—for example, the axons that originate in your spinal cord and terminate in the muscles of your toes. Small membrane-enclosed vesicles carrying materials essential to axonal function move along microtubules of the cytoskeleton, from the cell body to the tips of the axons. If the average velocity of a vesicle is 1 $\mu$m/s, how long does it take a vesicle to move from a cell body in the spinal cord to the axonal tip in the toes?

   **Answer** Transport time equals distance traveled/velocity, or

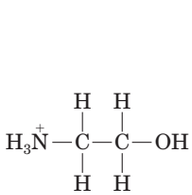
   $$(2 \times 10^6\ \mu\text{m})/(1\ \mu\text{m/s}) = 2 \times 10^6\ \text{s}$$

   or about 23 days!

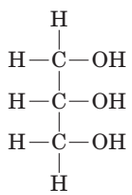
6. **Is Synthetic Vitamin C as Good as the Natural Vitamin?** A claim put forth by some purveyors of health foods is that vitamins obtained from natural sources are more healthful than those obtained by chemical synthesis. For example, pure L-ascorbic acid (vitamin C) extracted from rose hips is better than pure L-ascorbic acid manufactured in a chemical plant. Are the vitamins from the two sources different? Can the body distinguish a vitamin's source?

   **Answer** The properties of the vitamin—like any other compound—are determined by its chemical structure. Because vitamin molecules from the two sources are structurally identical, their properties are identical, and no organism can distinguish between them. If different vitamin preparations contain different impurities, the biological effects of the *mixtures* may vary with the source. The ascorbic acid in such preparations, however, is identical.

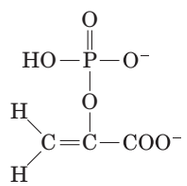
7. **Identification of Functional Groups** Figures 1–15 and 1–16 show some common functional groups of biomolecules. Because the properties and biological activities of biomolecules are largely determined by their functional groups, it is important to be able to identify them. In each of the compounds below, circle and identify by name each functional group.



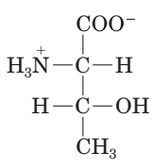
Ethanolamine
**(a)**



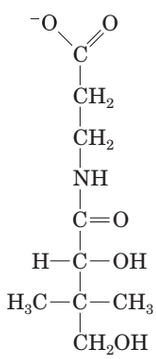
Glycerol
**(b)**



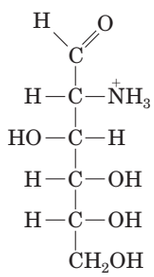
Phosphoenolpyruvate, an intermediate in glucose metabolism
**(c)**



Threonine, an amino acid
**(d)**



Pantothenate, a vitamin
**(e)**



D-Glucosamine
**(f)**

**Answer**

**(a)** $—NH_3^+$ = amino; —OH = hydroxyl

**(b)** —OH = hydroxyl (three)

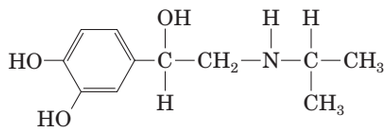
**(c)** $—P(OH)O_2^-$ = phosphoryl (in its ionized form); $—COO^-$ = carboxyl

**(d)** $—COO^-$ = carboxyl; $—NH_3^+$ = amino; —OH = hydroxyl; $—CH_3$ = methyl (two)

**(e)** $—COO^-$ = carboxyl; —CO—NH— = amide; —OH = hydroxyl (two); $—CH_3$ = methyl (two)

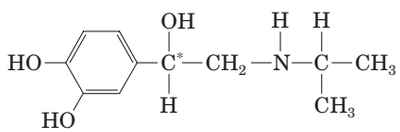
**(f)** —CHO = aldehyde; $—NH_3^+$ = amino; —OH = hydroxyl (four)

**8. Drug Activity and Stereochemistry** The quantitative differences in biological activity between the two enantiomers of a compound are sometimes quite large. For example, the D isomer of the drug isoproterenol, used to treat mild asthma, is 50 to 80 times more effective as a bronchodilator than the L isomer. Identify the chiral center in isoproterenol. Why do the two enantiomers have such radically different bioactivity?



Isoproterenol

**Answer** A chiral center, or chiral carbon, is a carbon atom that is bonded to four different groups. A molecule with a single chiral center has two enantiomers, designated D and L (or in the RS system, *S* and *R*). In isoproterenol, only one carbon (asterisk) has four different groups around it; this is the chiral center:



The bioactivity of a drug is the result of interaction with a biological "receptor," a protein molecule with a binding site that is also chiral and stereospecific. The interaction of the D isomer of a drug with a chiral receptor site will differ from the interaction of the L isomer with that site.

**9. Separating Biomolecules** In studying a particular biomolecule (a protein, nucleic acid, carbohydrate, or lipid) in the laboratory, the biochemist first needs to separate it from other biomolecules in the sample—that is, to *purify* it. Specific purification techniques are described later in the text. However, by looking at the monomeric subunits of a biomolecule, you should have some ideas about the characteristics of the molecule that would allow you to separate it from other molecules. For example, how would you separate **(a)** amino acids from fatty acids and **(b)** nucleotides from glucose?

**Answer**

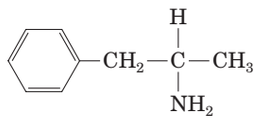
**(a)** Amino acids and fatty acids have carboxyl groups, whereas only the amino acids have amino groups. Thus, you could use a technique that separates molecules on the basis of the properties (charge or binding affinity) of amino groups. Fatty acids have long hydrocarbon chains and therefore are less soluble in water than amino acids. And finally, the sizes and shapes of these two types of molecules are quite different. Any one or more of these properties may provide ways to separate the two types of compounds.

**(b)** A nucleotide molecule has three components: a nitrogenous organic base, a five-carbon sugar, and phosphate. Glucose is a six-carbon sugar; it is smaller than a nucleotide. The size difference could be used to separate the molecules. Alternatively, you could use the nitrogenous bases and/or the phosphate groups characteristic of the nucleotides to separate them (based on differences in solubility, charge) from glucose.

**10. Silicon-Based Life?** Silicon is in the same group of the periodic table as carbon and, like carbon, can form up to four single bonds. Many science fiction stories have been based on the premise of silicon-based life. Is this realistic? What characteristics of silicon make it *less* well adapted than carbon as the central organizing element for life? To answer this question, consider what you have learned about carbon's bonding versatility, and refer to a beginning inorganic chemistry textbook for silicon's bonding properties.

**Answer** It is improbable that silicon could serve as the central organizing element for life under such conditions as those found on Earth for several reasons. Long chains of silicon atoms are not readily synthesized, and thus the polymeric macromolecules necessary for more complex functions would not readily form. Also, oxygen disrupts bonds between two silicon atoms, so silicon-based life-forms would be unstable in an oxygen-containing atmosphere. Once formed, the bonds between silicon and oxygen are extremely stable and difficult to break, which would prevent the breaking and making (degradation and synthesis) of biomolecules that is essential to the processes of living organisms.

**11. Drug Action and Shape of Molecules** Several years ago two drug companies marketed a drug under the trade names Dexedrine and Benzedrine. The structure of the drug is shown below.

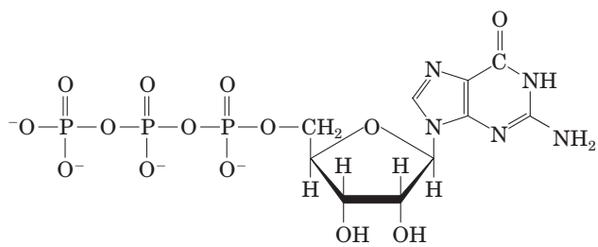


The physical properties (C, H, and N analysis, melting point, solubility, etc.) of Dexedrine and Benzedrine were identical. The recommended oral dosage of Dexedrine (which is still available) was 5 mg/day, but the recommended dosage of Benzedrine (no longer available) was twice that. Apparently, it required considerably more Benzedrine than Dexedrine to yield the same physiological response. Explain this apparent contradiction.

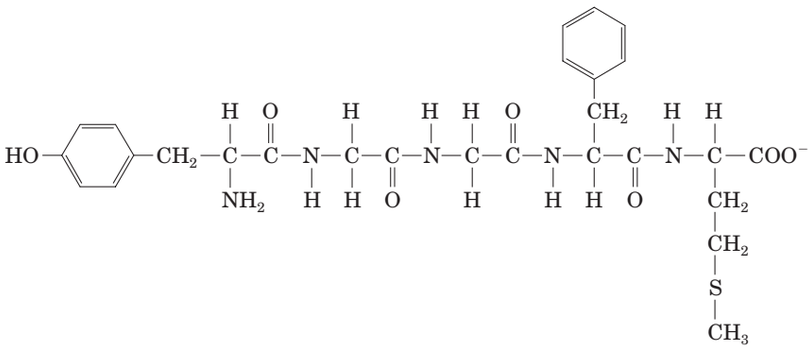
**Answer** Only one of the two enantiomers of the drug molecule (which has a chiral center) is physiologically active, for reasons described in the answer to Problem 3 (interaction with a stereospecific receptor site). Dexedrine, as manufactured, consists of the single enantiomer (D-amphetamine) recognized by the receptor site. Benzedrine was a racemic mixture (equal amounts of D and L isomers), so a much larger dose was required to obtain the same effect.

**12. Components of Complex Biomolecules** Figure 1–10 shows the major components of complex biomolecules. For each of the three important biomolecules below (shown in their ionized forms at physiological pH), identify the constituents.

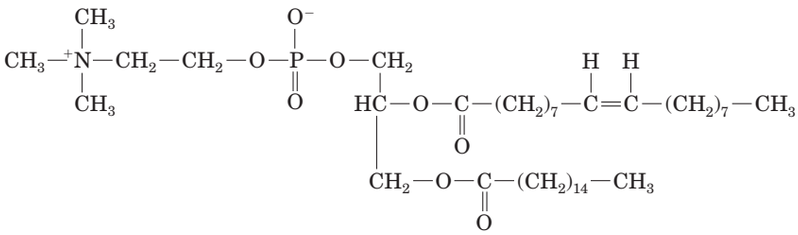
**(a)** Guanosine triphosphate (GTP), an energy-rich nucleotide that serves as a precursor to RNA:



**(b)** Methionine enkephalin, the brain's own opiate:



**(c)** Phosphatidylcholine, a component of many membranes:



**Answer**

**(a)** Three phosphoric acid groups (linked by two anhydride bonds), esterified to an $\alpha$-D-ribose (at the 5′ position), which is attached at C-1 to guanine.

**(b)** Tyrosine, two glycine, phenylalanine, and methionine residues, all linked by peptide bonds.

**(c)** Choline esterified to a phosphoric acid group, which is esterified to glycerol, which is esterified to two fatty acids, oleic acid and palmitic acid.

**13. Determination of the Structure of a Biomolecule** An unknown substance, X, was isolated from rabbit muscle. Its structure was determined from the following observations and experiments. Qualitative analysis showed that X was composed entirely of C, H, and O. A weighed sample of X was completely oxidized, and the $H_2O$ and $CO_2$ produced were measured; this quantitative analysis revealed that X contained 40.00% C, 6.71% H, and 53.29% O by weight. The molecular mass of X, determined by mass spectrometry, was 90.00 u (atomic mass units; see Box 1–1). Infrared spectroscopy showed that X contained one double bond. X dissolved readily in water to give an acidic solution; the solution demonstrated optical activity when tested in a polarimeter.

**(a)** Determine the empirical and molecular formula of X.

**(b)** Draw the possible structures of X that fit the molecular formula and contain one double bond. Consider *only* linear or branched structures and disregard cyclic structures. Note that oxygen makes very poor bonds to itself.

**(c)** What is the structural significance of the observed optical activity? Which structures in **(b)** are consistent with the observation?

**(d)** What is the structural significance of the observation that a solution of X was acidic? Which structures in **(b)** are consistent with the observation?

**(e)** What is the structure of X? Is more than one structure consistent with all the data?

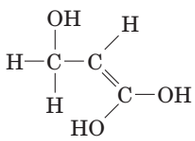
**Answer**

**(a)** From the C, H, and O analysis, and knowing the mass of X is 90.00 u, we can calculate the relative atomic proportions by dividing the weight percents by the atomic weights:

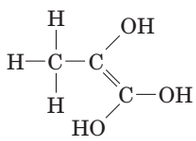
| Atom | Relative atomic proportion | No. of atoms relative to O |
|---|---|---|
| C | (90.00 u)(40.00/100)/(12 u) = 3 | 3/3 = 1 |
| H | (90.00 u)(6.71/100)/(1.008 u) = 6 | 6/3 = 2 |
| O | (90.00 u)(53.29/100)/(16.0 u) = 3 | 3/3 = 1 |

Thus, the empirical formula is $CH_2O$, with a formula weight of 12 + 2 + 16 = 30. The molecular formula, based on X having a mass of 90.00 u, must be $C_3H_6O_3$.

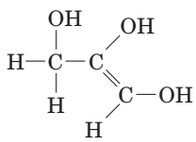
**(b)** Twelve possible structures are shown below. Structures **1** through **5** can be eliminated because they are unstable enol isomers of the corresponding carbonyl derivatives. Structures **9**, **10**, and **12** can also be eliminated on the basis of their instability: they are hydrated carbonyl derivatives (vicinal diols).



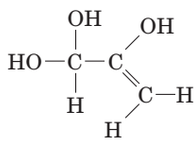
**1**



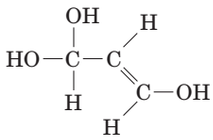
**2**



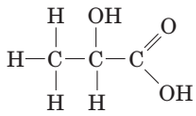
**3**



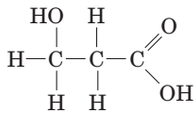
**4**



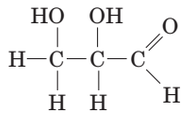
**5**



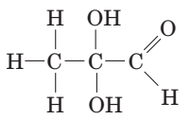
**6**



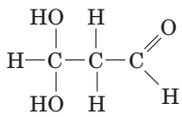
**7**



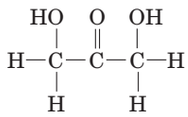
**8**



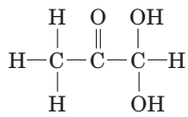
**9**



**10**



**11**



**12**

**(c)** Optical activity indicates the presence of a chiral center (a carbon atom surrounded by four different groups). Only structures **6** and **8** have chiral centers.

**(d)** Of structures **6** and **8**, only **6** contains an acidic group: a carboxyl group.

**(e)** Structure **6** is substance X. This compound exists in two enantiomeric forms that cannot be distinguished, even by measuring specific rotation. One could determine absolute stereochemistry by x-ray crystallography.

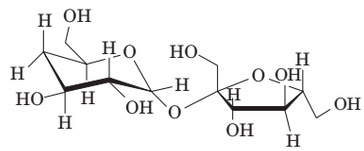
## Data Analysis Problem

**14. Sweet-Tasting Molecules** Many compounds taste sweet to humans. Sweet taste results when a molecule binds to the sweet receptor, one type of taste receptor, on the surface of certain tongue cells. The stronger the binding, the lower the concentration required to saturate the receptor and the sweeter a given concentration of that substance tastes. The standard free-energy change, $\Delta G°$, of the binding reaction between a sweet molecule and a sweet receptor can be measured in kilojoules or kilocalories per mole.

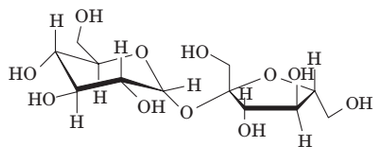
Sweet taste can be quantified in units of "molar relative sweetness" (MRS), a measure that compares the sweetness of a substance to the sweetness of sucrose. For example, saccharin has an MRS of 161; this means that saccharin is 161 times sweeter than sucrose. In practical terms, this is measured by asking human subjects to compare the sweetness of solutions containing different concentrations of each compound. Sucrose and saccharin taste equally sweet when sucrose is at a concentration 161 times higher than that of saccharin.

**(a)** What is the relationship between MRS and the $\Delta G°$ of the binding reaction? Specifically, would a more negative $\Delta G°$ correspond to a higher or lower MRS? Explain your reasoning.

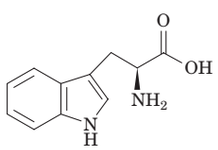
Shown below are the structures of 10 compounds, all of which taste sweet to humans. The MRS and $\Delta G°$ for binding to the sweet receptor are given for each substance.



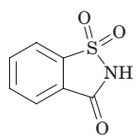
Deoxysucrose
MRS = 0.95
$\Delta G° = -6.67$ kcal/mol



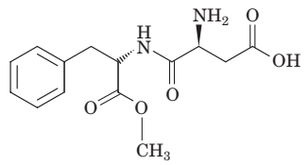
Sucrose
MRS = 1
$\Delta G° = -6.71$ kcal/mol



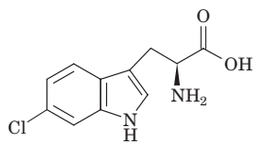
D-Tryptophan
MRS = 21
$\Delta G° = -8.5$ kcal/mol



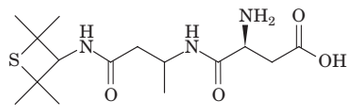
Saccharin
MRS = 161
$\Delta G° = -9.7$ kcal/mol



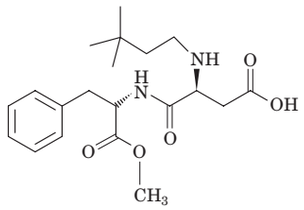
Aspartame
MRS = 172
$\Delta G° = -9.7$ kcal/mol



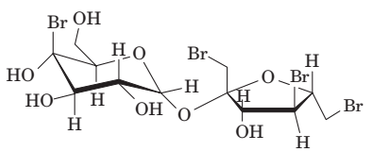
6-Chloro-D-tryptophan
MRS = 906
$\Delta G° = -10.7$ kcal/mol



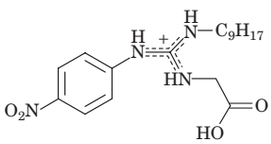
Alitame
MRS = 1,937
$\Delta G° = -11.1$ kcal/mol



Neotame
MRS = 11,057
$\Delta G^\circ = -12.1$ kcal/mol



Tetrabromosucrose
MRS = 13,012
$\Delta G^\circ = -12.2$ kcal/mol



Sucronic acid
MRS = 200,000
$\Delta G^\circ = -13.8$ kcal/mol

Morini, Bassoli, and Temussi (2005) used computer-based methods (often referred to as "in silico" methods) to model the binding of sweet molecules to the sweet receptor.

**(b)** Why is it useful to have a computer model to predict the sweetness of molecules, instead of a human- or animal-based taste assay?

In earlier work, Schallenberger and Acree (1967) had suggested that all sweet molecules include an "AH-B" structural group, in which "A and B are electronegative atoms separated by a distance of greater than 2.5 Å [0.25 nm] but less than 4 Å [0.4 nm]. H is a hydrogen atom attached to one of the electronegative atoms by a covalent bond" (p. 481).

**(c)** Given that the length of a "typical" single bond is about 0.15 nm, identify the AH-B group(s) in each of the molecules shown above.

**(d)** Based on your findings from **(c),** give two objections to the statement that "molecules containing an AH-B structure will taste sweet."

**(e)** For two of the molecules shown above, the AH-B model *can* be used to explain the difference in MRS and $\Delta G^\circ$. Which two molecules are these, and how would you use them to support the AH-B model?

**(f)** Several of the molecules have closely related structures but very different MRS and $\Delta G^\circ$ values. Give two such examples, and use these to argue that the AH-B model is unable to explain the observed differences in sweetness.

In their computer-modeling study, Morini and coauthors used the three-dimensional structure of the sweet receptor and a molecular dynamics modeling program called GRAMM to predict the $\Delta G^\circ$ of binding of sweet molecules to the sweet receptor. First, they "trained" their model—that is, they refined the parameters so that the $\Delta G^\circ$ values predicted by the model matched the known $\Delta G^\circ$ values for one set of sweet molecules (the "training set"). They then "tested" the model by asking it to predict the $\Delta G^\circ$ values for a new set of molecules (the "test set").

**(g)** Why did Morini and colleagues need to test their model against a different set of molecules from the set it was trained on?

**(h)** The researchers found that the predicted $\Delta G^\circ$ values for the test set differed from the actual values by, on average, 1.3 kcal/mol. Using the values given with the structures above, estimate the resulting error in MRS values.

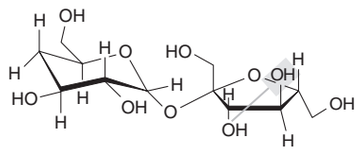
**Answer**

**(a)** A more negative $\Delta G°$ corresponds to a larger $K_{eq}$ for the binding reaction, so the equilibrium is shifted more toward products and tighter binding—and thus greater sweetness and higher MRS.

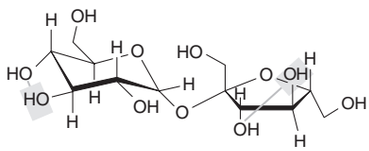
**(b)** Animal-based sweetness assays are time-consuming. A computer program to predict sweetness, even if not always completely accurate, would allow chemists to design effective sweeteners much faster. Candidate molecules could then be tested in the conventional assay.

**(c)** The range 0.25 to 0.4 nm corresponds to about 1.5 to 2.5 single-bond lengths. The figure below can be used to construct an approximate ruler; any atoms in the gray rectangle are between 0.25 and 0.4 nm from the origin of the ruler.

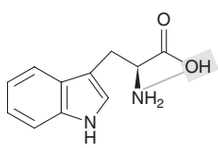
There are many possible AH-B groups in the molecules; a few are shown here.



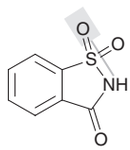
Deoxysucrose



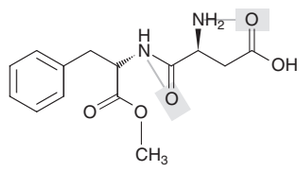
Sucrose



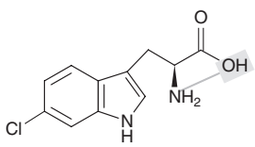
D-Tryptophan



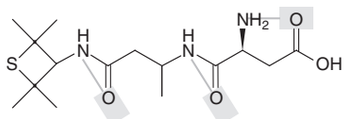
Saccharin



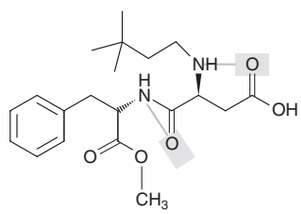
Aspartame



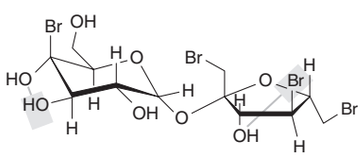
6-Chloro-D-tryptophan



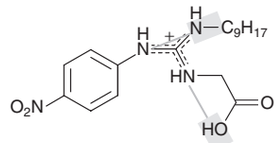
Alitame



Neotame



Tetrabromosucrose



Sucronic acid

**(d)** First, each molecule has multiple AH-B groups, so it is difficult to know which is the important one. Second, because the AH-B motif is very simple, many nonsweet molecules will have this group.

**(e)** Sucrose and deoxysucrose. Deoxysucrose lacks one of the AH-B groups present in sucrose and has a slightly lower MRS than sucrose—as is expected if the AH-B groups are important for sweetness.

**(f)** There are many such examples; here are a few: (1) D-Tryptophan and 6-chloro-D-tryptophan have the same AH-B group but very different MRS values. (2) Aspartame and neotame have the same AH-B groups but very different MRS values. (3) Neotame has two AH-B groups and alitame has three, yet neotame is more than five times sweeter than alitame. (4) Bromine is less electronegative than oxygen and thus is expected to weaken an AH-B group, yet tetrabromosucrose is much sweeter than sucrose.

**(g)** Given enough "tweaking" of parameters, any model can be made to fit a defined dataset. Because the objective was to create a model to predict $\Delta G°$ for molecules not tested in vivo, the researchers needed to show that the model worked well for molecules it had not been trained on. The degree of inaccuracy with the test set could give researchers an idea of how the model would behave for novel molecules.

**(h)** MRS is related to $K_{eq}$, which is related exponentially to $\Delta G°$, so adding a constant amount to $\Delta G°$ multiplies the MRS by a constant amount. Based on the values given with the structures, a change in $\Delta G°$ of 1.3 kcal/mol corresponds to a 10-fold change in MRS.

### References

**Morini, G., Bassoli, A., & Temussi, P.A.** (2005) From small sweeteners to sweet proteins: anatomy of the binding sites of the human T1R2_T1R3 receptor. *J. Med. Chem.* **48,** 5520–5529.

**Schallenberger, R.S. & Acree, T.E.** (1967) Molecular theory of sweet taste. *Nature* **216,** 480–482.