# Physical Interactions That Determine the Properties of Proteins

4

*T*he first three chapters of this volume have discussed proteins only in terms of their covalent structures. Knowledge of this structure is usually adequate for characterizing the chemistry of small molecules, but not for proteins. The large size of polypeptide chains enables them to fold back on themselves so that many simultaneous interactions take place among different parts of the molecule. A complex, three-dimensional structure results, which provides the unique environments and orientations of the functional groups that give proteins their many special properties. The biological activities of proteins are also mediated by their interactions with their environment: with water, salts, membranes, other proteins, nucleic acids, and the numerous other large and small molecules in living systems. All of these interactions arise from a limited set of fundamental noncovalent forces, but with many variations on the theme, so it is important to understand the physical basis of the interactions.

The physical natures of the noncovalent interactions between atoms are understood fairly well for individual molecules in a vacuum and in a regular solid, but not in liquids. This is a consequence of the complexity of the liquid state, with its constantly changing interactions among many molecules in transient ensembles. The complexities of liquids are especially relevant to proteins because their folded conformations usually occur only in a liquid-water environment or in membranes; the latter occur as a result of the relatively poor interactions of lipids with water. In spite of water's biological importance and the many studies of it, water is not one of the best understood liquids, and our limited understanding of water limits our understanding of proteins. The most important characteristic of all forces between molecules dissolved in water is that these forces are often due more to the properties of this extraordinary
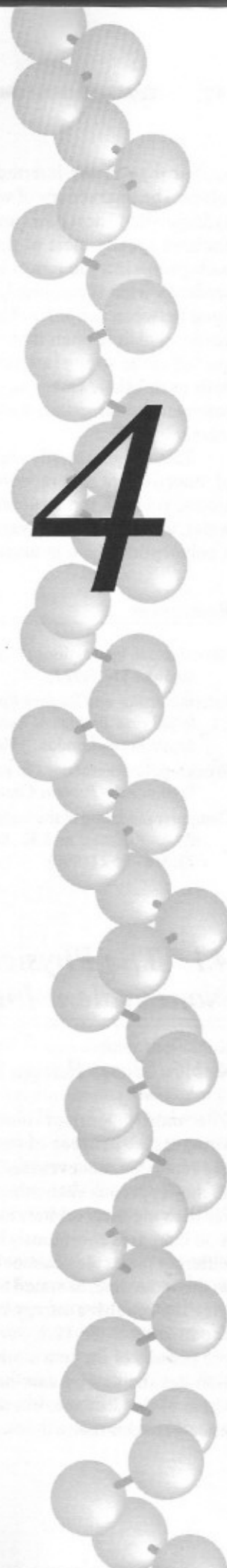
solvent than to the intermolecular interactions themselves. The interactions of water with ions, dipoles, and hydrogen bond acceptors and donors are so strong as to diminish greatly most of the forces that occur among such groups in a vacuum or in a nonpolar solvent. Water produces what is commonly considered to be a unique force between nonpolar atoms, the hydrophobic interaction, which is often thought of as primarily a consequence of the strong interaction of the water molecules with each other, rather than of any direct interaction between the nonpolar, hydrophobic molecules themselves.

This chapter briefly and simply describes the types of noncovalent interactions that take place among atoms, reviews the structure and properties of liquid water, and examines the various physical interactions in a polypeptide chain in an aqueous environment.

References

Noncovalent interactions. P. A. Kollman. *Acc. Chem. Res.* 10:365–371 (1977).

Intermolecular and Surface Forces; With Applications to Colloidal and Biological Systems. J. N. Israelachvili. Academic Press, London, 1985.

Weakly polar interactions in proteins. S. K. Burley and G. A. Petsko. *Adv. Protein Chem.* 39:125–189 (1988).

Computer modeling of the interactions of complex molecules. P. A. Kollman and K. M. Merz, Jr. *Acc. Chem. Res.* 23:246–252 (1990).

4.1 The Physical Nature of Noncovalent Interactions

4.1.1 Short-Range Repulsions

The most important interaction, energetically and structurally, between atoms and between molecules is the repulsion that eventually takes place between them as they approach each other. As they come near enough for their electron orbitals to begin to overlap, the repulsion increases enormously because the electrons on the different molecules cannot be in the same part of space at the same time, as stated by the Pauli exclusion principle. The repulsive energy is often said to increase with the inverse of the 12th power of the distance between the centers of the two atoms. A more realistic description has the energy varying exponentially with the inverse of the distance, but there is little practical difference between the two descriptions.

Table 4.1 *Van der Waals Radii of Atoms Found in Proteins*

Atom	Observed range (Å)	Radius when singly bonded (Å)
Hydrogen	1.0–1.54	1.17
Oxygen	1.4–1.7	1.40
Nitrogen	1.55–1.60	1.55
Carbon	1.70–1.78	1.75
Sulfur	1.75–1.80	1.80

Values from A. Bondi, *J. Phys. Chem.* 68:441–451 (1964) and A. Gavezzotti, *J. Amer. Chem. Soc.* 105:5220–5225 (1983).

Table 4.2 *Van der Waals Surface Areas and Volumes of Chemical Groups When Bonded to Carbon Atoms*

Chemical group	Area (Å^2)	Volume (Å^3)
—C—	1.0	5.5
—CH—	10.9	11.5
—CH$_2$—	20.9	16.8
—CH$_3$	33.4	22.3
Phenyl	94.9	76.1
—OH	19.3	12.6
—C— (C=O)	22.3	18.2
—C—OH (C=O)	43.4	
—SH		24.6
—NH$_2$	26.5	17.5
—NH—	16.4	13.4

Values from A. Bondi, *J. Phys. Chem.* 68:441–451 (1964) and A. Gavezzotti, *J. Amer. Chem. Soc.* 107:962–967 (1985).

Because the repulsive energy rises so steeply, it is possible to consider atoms and molecules as having definite dimensions and occupying volumes that are impenetrable to other atoms and molecules at ordinary temperatures. Individual atoms are usually modeled as

Table 4.3   *Volume Properties of Individual Amino Acid Residues*

Residue	Van der Waals volume[a] ($Å^3$)	Partial volume in solution[b] ($Å^3$)	Partial specific volume[b] (cm^3/g)
Ala (A)	67	86.4	0.732
Arg (R)	148	197.4	0.756
Asn (N)	96	115.6	0.610
Asp (D)	91	108.6	0.573
Cys (C)	86	107.9	0.630
Gln (Q)	114	142.0	0.667
Glu (E)	109	128.7	0.605
Gly (G)	48	57.8	0.610
His (H)	118	150.1	0.659
Ile (I)	124	164.6	0.876
Leu (L)	124	164.6	0.876
Lys (K)	135	166.2	0.775
Met (M)	124	160.9	0.739
Phe (F)	135	187.3	0.766
Pro (P)	90	120.6	0.748
Ser (S)	73	86.2	0.596
Thr (T)	93	113.6	0.676
Trp (W)	163	225.0	0.728
Tyr (Y)	141	190.5	0.703
Val (V)	105	136.8	0.831
Weighted average[c]			0.703

[a] Volume enclosed by van der Waals radius.

[b] Increase in volume of water after adding either one molecule or one gram of residue (A. A. Zamyatnin, *Ann. Rev. Biophys. Bioeng.* 13:145–165, 1984.)

[c] Weighted by frequency of occurrence in proteins, to give the value for an average residue in globular proteins.

spheres, and their impenetrable volumes are usually defined by the **van der Waals radius.** Values of van der Waals radii are usually measured from the smallest distances that can exist between neighboring, but not covalently bonded, atoms in the crystalline state; this distance is the sum of their respective van der Waals radii. Some typical values of van der Waals radii are given in Table 4.1. A range of values is given because the observed radius depends on the way in which the atom is covalently bonded. For example, the van der Waals radius of a hydrogen atom varies from 1.0 Å when bonded to an aromatic carbon atom to 1.54 Å when bonded to a negative ion. Fortunately, these are extreme variations, and two atoms are generally in close van der Waals contact when the distance between them is approximately 0.8 Å greater than when they are covalently bonded. The van der Waals radius is a minimal estimate of the size of an atom or molecule. Optimal van der Waals interactions (Sec. 4.1.3) generally occur at a

distance that is about 1.2 Å greater than the covalent bond length.

The van der Waals radius also defines the van der Waals surface area and volume of an atom or molecule; some pertinent values for various atoms and chemical groups are given in Table 4.2. The volumes and surface areas of entire molecules can be estimated by summing the parameters for the constituent parts, so long as the molecule is not structurally strained. The van der Waals volumes of the amino acid residues are given in Table 4.3.

The van der Waals surface of a molecule is not especially relevant chemically, because it inevitably includes many nooks and crannies between atoms that are not accessible to other atoms or molecules that may be present. A more practical concept of surface is that which is defined by the solvent molecules in contact with the molecule, the **accessible surface area.** It is generally described by the center of a solvent molecule of

Table 4.4 *Accessible Surface Areas of Amino Acid Residues in a Gly-X-Gly Tripeptide in an Extended Conformation*

Residue	Total ($Å^2$)	Main-chain atoms ($Å^2$)	Side-Chain Atoms ($Å^3$)		
			Total	Nonpolar	Polar
Ala	113	46	67	67	
Arg	241	45	196	89	107
Asn	158	45	113	44	69
Asp	151	45	106	48	58
Cys	140	36	104	35	69
Gln	189	45	144	53	91
Glu	183	45	138	61	77
Gly	85	85			
His	194	43	151	102	49
Ile	182	42	140	140	
Leu	180	43	137	137	
Lys	211	44	167	119	48
Met	204	44	160	117	43
Phe	218	43	175	175	
Pro	143	38	105	105	
Ser	122	42	80	44	36
Thr	146	44	102	74	28
Trp	259	42	217	190	27
Tyr	229	42	187	144	43
Val	160	43	117	117	

From S. Miller et al, *J. Mol. Biol.* 196:641–656 (1987).

radius 1.4 Å, representative of a water molecule, in van der Waals contact with the molecule (see Sec. 6.2.5). Estimated values for the accessible surface areas of the amino acid residues are given in Table 4.4.

References

Van der Waals volumes and radii. A. Bondi. *J. Phys. Chem.* 68:441–451 (1964).

Crystal environments and geometries of leucine, isoleucine, valine and phenylalanine provide estimates of minimum nonbonded contact and preferred van der Waals interaction distances. R. O. Gould et al. *J. Amer. Chem. Soc.* 107:5921–5927 (1985).

The size of molecules. A. Y. Meyer. *Chem. Soc. Rev.* 15:449–474 (1986).

Properties of atoms in molecules: atomic volumes. R. F. W. Bader et al. *J. Amer. Chem. Soc.* 109:7968–7979 (1987).

Calibration of effective van der Waals atomic contact radii for proteins and peptides. H. Iijima et al. *Proteins: Struct. Funct. Genet.* 2:330–339 (1987).

4.1.2 Electrostatic Forces

a. Point Charges

All intermolecular forces are thought to be essentially electrostatic in origin, and the most fundamental noncovalent attraction is that between electrostatic charges. Coulomb's law states that the energy of the electrostatic interaction between two atoms A and B in a vacuum is simply the product of their two charges divided by the distance between them r_{AB}:

$$\Delta E = \frac{Z_A Z_B \epsilon^2}{r_{AB}} \tag{4.1}$$

where ϵ is the charge of an electron and Z the number of such charges on each atom. The energy ΔE is expressed relative to the energy when the two charges are very far apart.

If the two charges are of opposite sign, the energy decreases as they approach each other, and the interaction is favorable; if the charges are of the same sign, there is repulsion between them. Because the interac-

tion varies inversely with only the first power of the distance, it is effective over relatively large distances; at twice the distance, the interaction energy is only halved. It is also a very strong interaction. The attractive interaction between an isolated Na^+ ion in optimum contact with a Cl^- ion, 2.76 Å apart, has the value 2.0×10^{-19} calorie, which is equivalent to 120 kcal/mol (502 kJ/mol).

Coulomb's law (Eq. 4.1) is valid only for two point charges in a vacuum. For other environments, the electrostatic interaction is modulated by other interactions. In homogeneous environments, the electrostatic interaction is decreased by the **dielectric constant D**:

$$\Delta E = \frac{Z_A Z_B \epsilon^2}{D r_{AB}} \qquad (4.2)$$

The values of dielectric constants are invariably greater than unity, so the electrostatic interaction in media other than a vacuum is always less than that stated by Coulomb's law. This is especially important to consider when dealing with liquids, which usually have dielectric constants in the range of 2–110 (that of water is about 80). The concept of a dielectric constant is valid only for homogeneous liquids; less homogeneous environments must be treated explicitly. Special problems arise at interfaces between regions with very different bulk dielectric constants, such as between an aqueous solvent and a folded protein. Two charges on opposite sides of a sphere of low dielectric constant immersed in a medium of high dielectric constant interact not through the shortest distance of low dielectric, but around the outside of the sphere. The very long path length through the high dielectric medium results in the energy of this electrostatic interaction being much smaller than might otherwise be expected.

Because Coulomb's law ignores the finite sizes of ions, it is valid only at distances that are significantly greater than atomic dimensions. The charge of an atom is separated between the nucleus and the diffuse electron cloud, so it cannot be treated as a point charge at short distances. This problem of charge separation is even more severe with proteins, in which the charges of the ionized groups of the side chains of Lys, Arg, His, Asp, and Glu residues, plus the α-amino and α-carboxyl groups, are distributed over two or more hydrogen or oxygen atoms. Consequently, interactions between very close, oppositely charged groups in a protein, known as **salt bridges,** usually consist not only of electrostatic interactions but also of some degree of hydrogen bonding (Fig. 4.1).

The charged groups in polypeptide chains titrate in accessible pH regions, so electrostatic effects in proteins invariably involve changes in their ionization
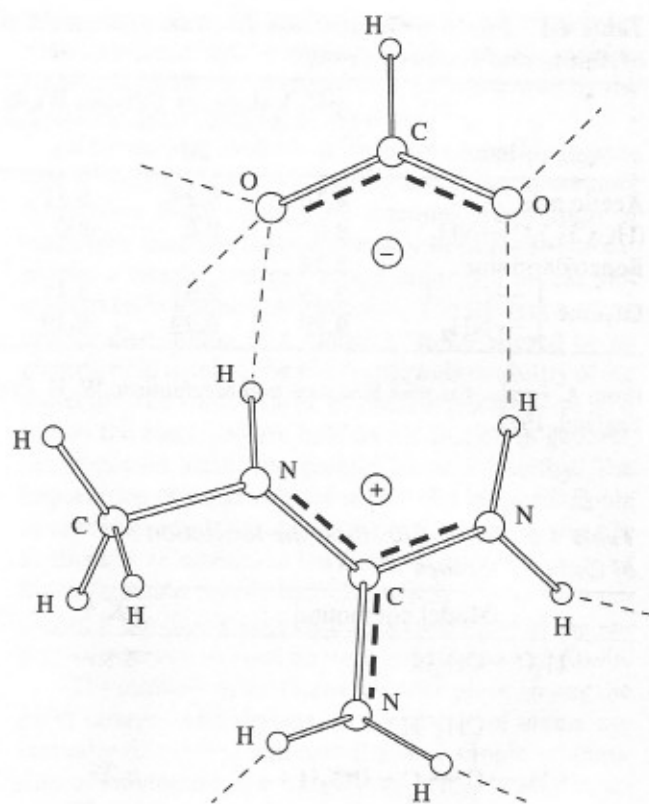

FIGURE 4.1

Hydrogen bonds involved in the interaction between ionized carboxyl and guanidinium groups, illustrated by the crystal structure of methylguanidinium formate. The thin dashed lines indicate hydrogen bonds assumed to be present from the short distances between the adjacent atoms. The thick dashed lines indicate partial double bonds arising from resonance. The distances between the nitrogen and oxygen atoms of the two complete hydrogen bonds shown are 2.81 Å *(left)* and 2.89 Å *(right)*. (Adapted from D. D. Bray et al., *Int. J. Peptide Protein Res.* 24:414–418, 1984.)

tendencies, or pK_a values (see Table 1.2). Favorable electrostatic interactions increase the tendency of any group to ionize. Other factors, however, such as the accessibility to the solvent and the solvent's polarity, also affect the pK_a value. For example, adding dioxane to decrease the polarity of the solvent inhibits ionization of accessible amino and carboxyl groups (Table 4.5). Ionization also becomes less favorable with increasing bulkiness of the surrounding aliphatic groups (Table 4.6).

b. Dipoles

A molecule need not have a net charge to participate in electrostatic interactions because electron density can

Table 4.5 *Effect of Nonaqueous Environment on the pK_a Values of Amino and Carboxyl Groups*

Acid or base	pK_a Values for Various Wt % Dioxane in Water				
	0	20	45	50	70
Acetic acid	4.76	5.29	6.31		8.34
$(HOCH_2)_3C-NH_2$	8.0	8.0	8.0	8.0	8.0
Benzoylarginine	3.34			4.59	4.60
Glycine $\{$ $-CO_2H$	2.35	2.63	3.11		3.96
$-NH_2$	9.78	9.29	8.49		7.42

From A. Fersht, *Enzyme Structure and Mechanism*, W. H. Freeman, Reading, England, 1977.

Table 4.6 *Steric Effects on the Ionization of Carboxyl Groups*

Model compound	pK_a[a]
H_3C-CO_2H	5.55
(branched structure)	6.25
(branched structure)	6.44
(branched structure)	6.71
(branched structure)	6.97

[a] The pK_a values were measured in equal volumes of methanol and water at 40°C by G. S. Hammond and D. H. Hogle, *J. Amer. Chem. Soc.* 77:338–340 (1955).

be localized if the atoms have different electronegativities. Atoms with the greatest electronegativities have an excess of negative charge, $\delta-$, others an excess of positive charge, $\delta+$. The electronegativities of the atoms in proteins are O, 3.45; N, 2.98; C, 2.55; S, 2.53; and H, 2.13. The importance of these values is illustrated by the peptide bond, which has partially double-bonded character due to resonance with a form in which the more electronegative oxygen atom acquires a net partial negative charge, whereas the $-NH-$ group is somewhat positively charged.

$$\underset{C^\alpha}{\overset{O}{\underset{}{}}}C-\underset{H}{\overset{C^\alpha}{N}} \rightleftharpoons \underset{C^\alpha}{\overset{O^-}{\underset{}{}}}C=\underset{H^+}{\overset{C^\alpha}{N}} \qquad (4.3)$$
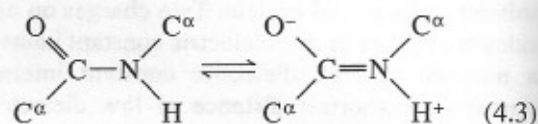
The double-bonded charged species is populated about 40% of the time, so the peptide group can be represented as having partial charges of up to $\pm0.4\epsilon$. Polar oxygen and nitrogen atoms in other molecules have partial charges as great as $\pm0.35\epsilon$, but those in aliphatic amino acid side chains are probably no greater than $\pm0.1\epsilon$.

The π electrons in the aromatic rings of Phe, Tyr, and Trp residues are localized above and below the face of the ring. This excess of electrons gives the face of the ring a small net negative charge of approximately -0.15ϵ, whereas the hydrogen atoms on its edge have a corresponding positive net charge. The electrostatic interactions between these partial charges dominate the interactions between aromatic rings. Such rings preferentially interact with the positively charged edge of one ring pointing at the negatively charged face of another, or with their rings parallel but offset so that the edge of each ring is interacting with the face of the other. The commonly held impression that aromatic rings interact favorably by stacking their rings one above the other is incorrect. Electronegative oxygen and sulfur atoms

tend to interact favorably with the edges of aromatic rings, —NH groups with their faces.

The separation of charge in a molecule determines its **dipole moment** μ_D, which is given by the product of the magnitude of the separated excess charge Z and the distance d by which it is separated:

$$\mu_D = Zd \qquad (4.4)$$

One electron unit of positive and negative charge separated by 1 Å has a dipole moment of 4.8 Debye units (D). The dipole moment of a peptide bond is 3.5 D and that of a water molecule 1.85 D. The dipole moment has direction as well as magnitude and is usually depicted as a vector along a straight line from the negative to the positive charge. The dipole moment of the peptide bond can be represented as

$$(4.5)$$
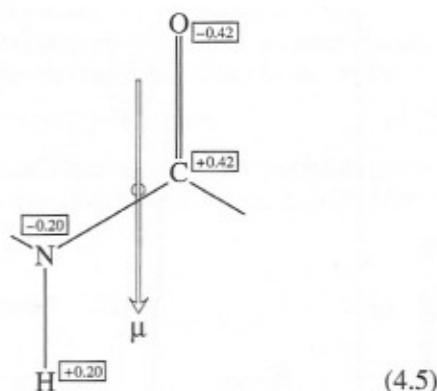
Dipoles interact with point charges, with other dipoles, and with more complex charge separations known as quadrupoles, octupoles, and so on, in a complex manner that is determined by the relative orientations of the two interacting groups. The interactions can be computed by considering the individual charges of the groups, including those of the dipole, quadrupole, and so on, and calculating the coulombic interactions among all of them. The interactions among the four partial charges of two dipoles are analogous to those between two bar magnets. Two side-by-side dipoles repel each other when parallel, whereas there is an equivalent attraction between them when they are antiparallel. Maximum interactions, repulsive or attractive, occur in a head-to-tail orientation. Dipolar interactions are, however, weaker than those between ions because both attractions and repulsions occur between the two separated charges of each of the dipoles. This combination of attraction and repulsion also has the effect of making the energy of the interactions depend inversely on the second to third power of the distance between interacting molecules when they are in fixed orientations, and on the sixth power when they are free to rotate in response to the interaction. Consequently, electrostatic interactions between dipoles fall off much

more abruptly as the distance between them increases than do interactions between point charges; the interactions are, however, modulated in the same way by the dielectric properties of the medium.

Interactions involving dipoles also modify the nature of the dipole charge distribution in the interacting molecules. Being simply an unequal distribution of electrons, that distribution is easily perturbed. For example, a nearby charged group induces a dipole moment even in a spherical molecule. The tendency of the charge distribution of a molecule to be altered by an electric field is called the electronic **polarizability** of the molecule, the value of which depends primarily on how tightly the electrons are held by the nuclei; in general, the larger an atom, the greater its polarizability. The importance of polarizability is that the induced dipole always interacts favorably with the field that induced it, so there is an attraction between them. The energy of this interaction is only half that which would have occurred if the dipole preexisted because some of the energy of interaction must be used in inducing the dipole.

The multiple interactions that take place among the point charges and dipoles on a number of atoms are mutually dependent and turn the very simple relationship of Coulomb's law into a very complex phenomenon. The electrostatic interactions among molecules in a homogeneous liquid can be averaged and expressed as a simple dielectric constant of the liquid. This concept is not valid, however, when the environment is not homogeneous at the molecular level, as is always the case for proteins, whose electrostatic interactions invariably involve interactions among the multiple charges and dipoles of the protein, and between these and the bulk solvent and any ions in it. In this case, interactions between individual charges and dipoles of the protein must be calculated directly. This is impractical with the many molecules of the solvent, and there is still considerable controversy about how to analyze and calculate electrostatic interactions in proteins.

References

Calculations of electrostatic interactions in biological systems and in solutions. A. Warshel and S. T. Russell. *Quart. Rev. Biophys.* 17:283–422 (1984).

Electrostatic effects in proteins. J. B. Matthew. *Ann. Rev. Biophys. Bioeng.* 14:387–417 (1985).

Aromatic–aromatic interactions: a mechanism of protein structure stabilization. S. K. Burley and G. A. Petsko. *Science* 229:23–28 (1985).

The modelling of electrostatic interactions in the function of globular proteins. N. K. Rogers. *Prog. Biophys. Mol. Biol.* 48:37–66 (1986).

Treatment of electrostatic effects in macromolecular model-

ing. S. C. Harvey. *Proteins: Struct. Funct. Genet.* 5:78–92 (1989).

The nature of π–π interactions. C. A. Hunter and J. K. M. Sanders. *J. Amer. Chem. Soc.* 112:5525–5534 (1990).

Electrostatic interactions in macromolecules: theory and applications. K. A. Sharp and B. Honig. *Ann. Rev. Biophys. Biophys. Chem.* 19:301–332 (1990).

4.1.3 Van der Waals Interactions

All atoms and molecules attract each other, even in the absence of charged groups, as a result of mutual interactions related to the induced polarization effects described in the preceding section. These ubiquitous attractions, known as **van der Waals interactions**, are weak and close-range, varying as the sixth power of the distance between them, d^{-6}. They arise from three types of interactions: those between two permanent dipoles, those between a permanent and an induced dipole, and those between two mutually induced dipoles, known as **London** or **dispersion forces**. The first two types have been described, but the third is the most important because it occurs among all atoms and molecules.

The dispersion force is complex because it is quantum mechanical in nature. A greatly simplified description can be derived from the classical picture of an atom with electrons orbiting the nucleus. Such a spherical atom has no net dipole moment, but it can have a transient dipole moment resulting from a temporarily asymmetric orientation of the electrons and nucleus. This transient dipole polarizes any neutral atom nearby, creating an attraction between them. Although the transient dipole of the first atom constantly and rapidly changes, that of the other atom tends to follow it, and the two are correlated. This dispersion force is basically electrostatic in nature and varies as d^{-6}, as do the other two components of van der Waals interactions. This distance dependence breaks down at distances greater than about 50 Å, however, because the correlation between the two electron distributions diminishes as the time it takes the field from one atom to reach the other increases.

Van der Waals interactions are often represented by an energy potential as a function of distance d that includes both the attractive force and the repulsion at close range. The most well known of these is the **Lennard–Jones potential** of the form:

$$E(d) = \frac{C_n}{d^n} - \frac{C_6}{d^6} \qquad (n > 6) \qquad (4.6)$$

where C_n and C_6 are constants. The first term gives the repulsions, the second the van der Waals attractions. The most common potential has $n = 12$, which is

known as the Lennard–Jones 6,12 potential and is computationally efficient because $12 = 2 \times 6$ (Fig. 4.2).

The optimal distance for the interaction of two atoms, given by the minimum in Figure 4.2, is usually 0.3–0.5 Å greater than the sum of their van der Waals radii as measured from the closest contact distance in crystals (Table 4.1). The van der Waals radius is given by the steeply ascending repulsive interaction at closer distances. The van der Waals interaction is generally considered to be independent of the orientation of the interacting molecules, but this is only approximately true, when the interacting molecules are independent and tumbling rapidly in a gas or a liquid. Otherwise, the magnitude of the interaction of even a nonpolar group
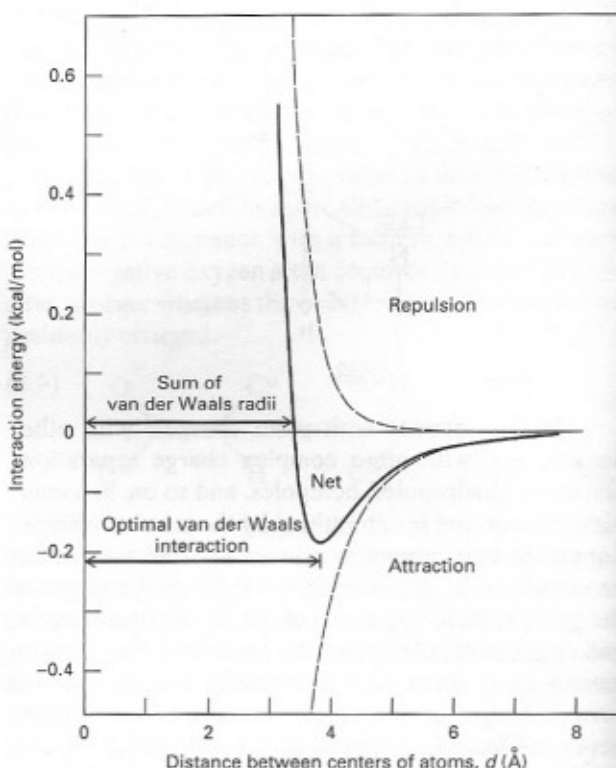
FIGURE 4.2

Representative profile of the energy of the van der Waals interaction as a function of the distance d between the centers of the two atoms. The individual attractive and repulsive components are indicated by the dashed lines, the net interaction by the solid line. The optimal interaction between the two atoms occurs where the energy is at a minimum. The sum of the van der Waals radii of the two atoms is given by the distance at which the energy increases sharply. The interaction energy was calculated using the Lennard–Jones 6,12 potential (Eq. 4.6) with $C_{12} = 2.75 \times 10^6$ Å^{12} kcal/mol and $C_6 = 1425$ Å^6 kcal/mol for the interaction between two carbons atoms (M. Levitt, *J. Mol. Biol.* 82:393–420, 1974).

like —CH_3 can vary with orientation because the polarizability of a C—H bond is nearly twice as great along the bond as perpendicular to it.

References

Van der Waals forces in biological systems. J. N. Israelachvili. *Quart. Rev. Biophys.* 6:341–387 (1973).

Van der Waals picture of liquids, solids and phase transformations. D. Chandler et al. *Science* 220:787–794 (1983).

Theoretical studies of van der Waals molecules and intermolecular forces. A. D. Buckingham et al. *Chem. Rev.* 88:963–988 (1988).

4.1.4 Hydrogen Bonds

A hydrogen bond occurs when two electronegative atoms compete for the same hydrogen atom:

$$—D—H\cdots A— \qquad (4.7)$$

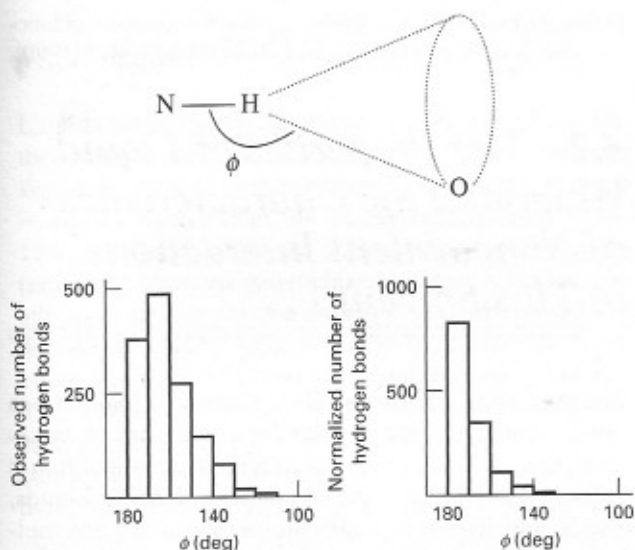The hydrogen atom is formally bonded covalently to one of the atoms, the *donor* D, but it also interacts favorably

FIGURE 4.3

Linearity of N—H $\cdots$ O hydrogen bonds observed in crystal structures of small molecules. The degree of linearity is measured by the angle ϕ *(top)*, which would have a value of 180° for a perfectly linear hydrogen bond. The histogram of observed values of intermolecular hydrogen bonds *(left)* is affected by a geometric factor in which the various ranges of the value of ϕ include different volumes of three-dimensional space. Correcting for this gives the histogram on the *right*, which illustrates the marked tendency of hydrogen bonds to be linear. (Adapted from R. Taylor and O. Kennard, *Acc. Chem. Res.* 17:320–326, 1984.)

with the other, the *acceptor* A. In a few strong, short hydrogen bonds, the hydrogen atom is symmetrically placed between the two electronegative atoms, but usually it is covalently bonded to one of the atoms, with a normal covalent bond length.

The main component of the hydrogen bond is an electrostatic interaction between the dipole of the covalent bond to the hydrogen atom, in which the hydrogen atom has a partial positive charge, and a partial negative charge on the other electronegative atom:

$$—D^{\delta-}—H^{\delta+}\cdots{}^{\delta-}A— \qquad (4.8)$$
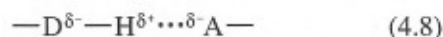
The hydrogen atom is special in being able to interact strongly with one electronegative atom while being covalently attached to another. It can do this because of its small size and its substantial charge, which results from its tendency to be positively polarized. In strong hydrogen bonds, an additional covalent aspect arises from a transfer of electrons. The electrostatic and covalent aspects of the hydrogen bond cause the most common and, presumably, most energetically favorable hydrogen bonds to keep the three bonded atoms collinear (Fig. 4.3). There is considerable uncertainty, however, about how the strength of the hydrogen-bond interaction varies with departures from linearity.

Oxygen atoms are frequently observed to participate simultaneously as acceptors in two hydrogen bonds:

$$\qquad (4.9)$$
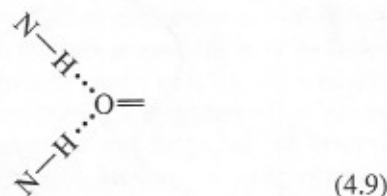
The partial negative charge at the electronegative acceptor atom such as an oxygen atom is localized on the two lone-pair electron orbitals (Sec. 4.2.2), and an intrinsic preference for one hydrogen atom to point toward each of these orbitals might be expected. Hydrogen bonds between two molecules in the gas phase show this geometry, and it is also frequently observed in crystals with carbonyl oxygen acceptors (Fig. 4.4). It is not observed so frequently in other circumstances, however, and the energetic preference for hydrogen atoms to be directed at the lone-pair electrons is probably so small as to be easily overwhelmed by other considerations. Much less frequently, single hydrogen atoms are shared between two acceptors.

The lengths and strengths of hydrogen bonds depend on the electronegativities of the acceptor and donor; the greater their electronegativities, the shorter the distance between them and the stronger the hydrogen bond. Charged groups also give shorter and

stronger hydrogen bonds (Table 4.7). Hydrogen bonds in proteins most frequently involve the $C=O$ and $N-H$ groups of the polypeptide backbone. In this type of hydrogen bond, the $H \cdots O$ distance is most often 1.9–2.0 Å. The hydrogen atom is generally not observed directly in protein crystal structures, however (see Chap. 6), so hydrogen-bond distances in proteins are usually expressed as the distance between the donor and acceptor atoms. The covalent $N-H$ distance is
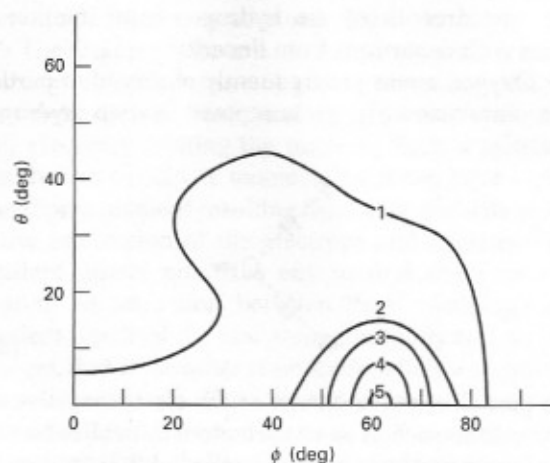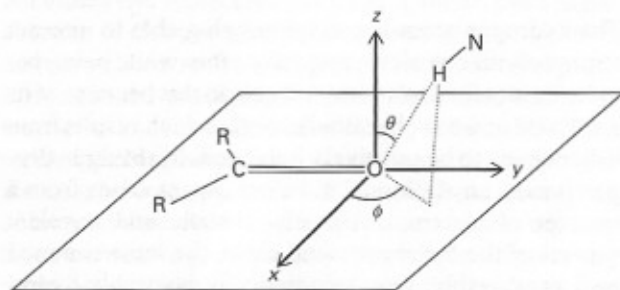
1.03 ± 0.02 Å, so a typical $N-H \cdots O=C$ hydrogen bond has the nitrogen and oxygen atoms 3.0 Å apart.

The strengths of hydrogen bonds are generally said to be within the rather broad range of 2–10 kcal/mol at room temperature. Part of this variation is due to the variety of hydrogen bonds, but much is also a result of uncertainty. The chemical groups in proteins that most commonly serve as hydrogen-bond donors are $N-H$, $O-H$, and, much less frequently, $S-H$ and $C-H$ groups. The most common acceptors are $O=$, $-O-$, $-N=$, and, much less frequently, $-S^-$, $-S-$, and the π electrons of aromatic groups.
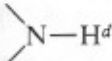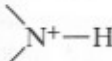
References

Hydrogen bonding. G. C. Pimentel and A. L. McLellan. *Ann. Rev. Phys. Chem.* 22:347–385 (1971).

The origin of hydrogen bonding. An energy decomposition study. H. Umeyama and K. Morokuma. *J. Amer. Chem. Soc.* 99:1316–1332 (1977).

Hydrogen-bond geometry in organic crystals. R. Taylor and O. Kennard. *Acc. Chem. Res.* 17:320–326 (1984).

Directional character, strength, and nature of the hydrogen bond in gas-phase dimers. A. C. Legon and D. J. Millen. *Acc. Chem. Res.* 20:39046 (1987).

Hydrogen bond stereochemistry in protein structure and function. J. A. Ippolito et al. *J. Mol. Biol.* 215:457–471 (1990).

4.2 The Properties of Liquid Water and the Characteristics of Noncovalent Interactions in This Solvent

The preceding discussion of intermolecular interactions concentrated on interactions between pairs of molecules in a vacuum, where the nature of the interaction is relatively straightforward. In condensed media, liquids, solids, and macromolecules, numerous atoms and molecules are interacting simultaneously and usually inducing alterations in each other, so the exact mathematical treatment of interactions becomes much more problematic. For example, involvement of an $O-H$ group as donor in a hydrogen bond increases the negative charge on the donor oxygen atom, so it becomes a better hydrogen-bond acceptor in a second hydrogen bond. Van der Waals interactions among atoms affect their polarizabilities, and the magnitude of the van der Waals interaction between two molecules is about 30% greater when they are part of a liquid than when isolated in a vacuum.

FIGURE 4.4

The geometries of $C=O \cdots H-N$ hydrogen bonds observed in crystal structures of small molecules. The definitions of the angles ϕ and θ are illustrated at the *top*, and the relative frequencies of their observed values in intermolecular hydrogen bonds (R. Taylor et al., *J. Amer. Chem. Soc.* 105:5761–5766, 1983) are given by the contours. The angle ϕ measures departures from linearity of the $C=O$ bond and the H atom; the most frequently observed values are in the region of 50°–60°. The angle θ measures the extent to which the H atom lies out of the plane defined by the R, C, and O atoms; the most commonly observed values are in the region of 0°–7°. The lone-pair electrons of the oxygen atom are believed to project at angles of $\phi = 60°$, $\theta = 0°$. The spherical polar coordinate system used here gives a bias toward small values of θ that could be corrected by plotting $\sin \theta$.

Table 4.7 *Lengths of* H—N $\cdots$ O=C *hydrogen bonds*[a]

Donor	Mean H $\cdots$ O Distance for Different Acceptors (Å)		
	Carboxyl[b]	Carboxylate[c]	Amide
$\diagdown$ N—H[d] $\diagup$	2.002 ± 0.012	1.928 ± 0.012	1.934 ± 0.005
$\diagdown$ N$^+$—H[e] $\diagup$	1.983 ± 0.055	1.869 ± 0.028	1.858 ± 0.043
NH_4^+	1.916 ± 0.041	1.886 ± 0.018	1.988 ± 0.075
R—NH_3^+	1.936 ± 0.014	1.841 ± 0.008	1.891 ± 0.034
R_2—NH_2^+	1.887 ± 0.047	1.796 ± 0.014	1.793 ± 0.070
R_3—NH^+		1.722 ± 0.025	1.845 ± 0.014

[a] The N—H distance is generally 1.03 Å; adding this value to the tabulated distances gives the distance between the N and O atoms.

[b] C=O oxygen atom of unionized carboxylic acids and esters.

[c] Oxygen atom of carboxyl anions (—CO_2^-).

[d] Uncharged donor.

[e] Charged donor with trigonal geometry.

From R. Taylor and O. Kennard, *Acc. Chem. Res.* 17:320–326 (1984).

4.2.1 *Liquids*

Liquids have no clearly defined structure and are further complicated by the many interactions occurring simultaneously among neighboring molecules. A liquid is usually lighter than the corresponding solid by 5–15%, indicating increased distances and decreased interactions between molecules in a liquid. Liquids usually have similar volume densities (the ratio of the van der Waals volume of a molecule to the average total volume occupied by it in the liquid) between 0.48 and 0.61. These densities barely change with pressure, indicating that the molecules are generally in van der Waals contact with their neighbors; for example, doubling atmospheric pressure generally decreases the volume of a liquid by only 0.01%. On the other hand, the liquid volume generally increases by about 0.1% for each 1°C rise in temperature, so the molecules are also fluctuating substantially.

The best experimental description of liquid structure comes from the scattering of X rays or neutrons, which yields a **radial distribution function** $g(r)$ that gives the symmetrically averaged density of atoms as a function of radial distance from the center of one molecule, r, relative to the bulk density of the liquid (Fig. 4.5). The value of $g(r)$ is zero at $r = 0$, and its value becomes substantial when r approaches twice the van der Waals radius of the molecule. At about this distance, $g(r)$ gen-

erally reaches a maximum, with a value indicating that, in a regular liquid consisting of nearly spherical molecules, 9–11 nearest-neighbor molecules are packed in nearly van der Waals contact around the central molecule. At somewhat greater values of r, the value of $g(r)$ drops to a minimum, indicating that there are few spaces for molecules to penetrate the first shell of nearest neighbors. The density again increases to another maximum when r is just under two molecular diameters, corresponding to the second shell of neighboring atoms. The second maximum is markedly lower than the first, indicating that the structural order is diminishing with increasing distance. This trend continues, and a third shell may be apparent, but the atom density rapidly approaches that of the bulk liquid. The lower the temperature, the greater the degree of order in liquids unless the volume is kept constant by decreasing the pressure. More detailed descriptions of liquids come from numerical simulations, but their validity depends on that of the model used for the calculations.

Current models depict liquids as consisting of close-packed, hard-sphere molecules in which the packing is both irregular and constantly changing. The shapes of molecules and the harsh repulsive forces between them largely determine the properties of a liquid. The structure of liquid argon is well represented by a box of marbles, that of liquid benzene by a box of small O-rings. Although the attractive interactions between
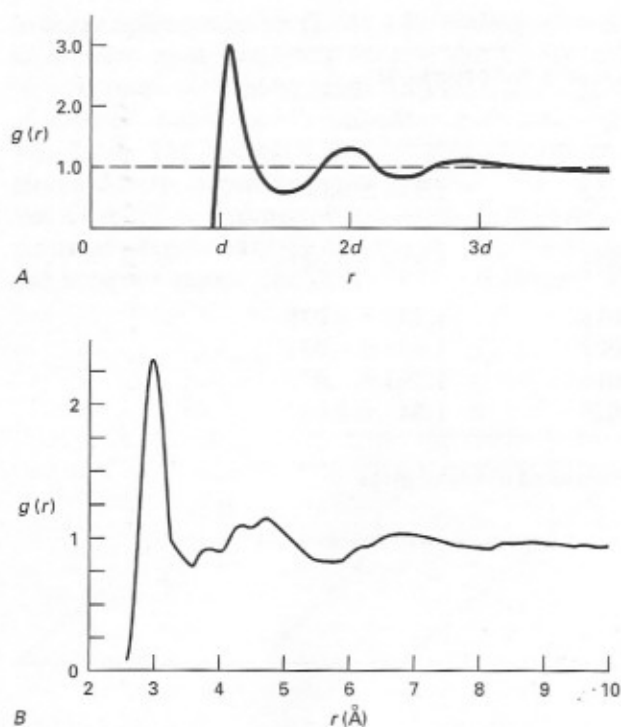
FIGURE 4.5

Radial distribution function of a normal liquid *(A)* and of water at 4°C *(B)*. The probability of finding other atoms at radial distance *r* from a central molecule is given by $g(r)$ times the density of the liquid. The distance *r* is expressed in *A* in terms of the van der Waals diameter of the molecule, *d*. The experimental curve for water measured using X rays is given in *B*; X rays are scattered primarily by the oxygen atom of the water molecule. (From A. H. Narten and H. A. Levy, *J. Chem. Phys.* 55:2263–2269, 1971.)

molecules stabilize the liquid phase, they play a minor role in determining its structure unless they include hydrogen bonds or strong ionic interactions, which is the case with the biologically most important liquid, water.

References

The structure of liquids. J. D. Bernal. *Proc. Roy. Soc. London* A280:299–322 (1964).

Structures of molecular liquids. D. Chandler. *Ann. Rev. Phys. Chem.* 29:441–471 (1971).

4.2.2 Water

The H_2O molecule is unique in having the same number of hydrogen atoms and lone-pair electron acceptors. As a result, hydrogen bonding is preeminent. The H_2O molecule has a bent geometry, with an O—H bond

length of 0.957 Å and a bond angle of 104.5°:

(4.10)
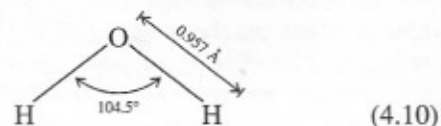
With van der Waals radii of 1.2 Å and 1.45 Å for the hydrogen and oxygen atoms, respectively, the molecule has a volume of 17.7 Å³. The molecule in this stick representation appears very asymmetric, but oxygen has eight electrons and a share of the single electron of both hydrogens, so the electronic structure of the molecule is nearly spherical (Fig. 4.6); it is often represented as a sphere of radius 1.4 Å. The net charge is distributed asymmetrically, however, with excess electrons on the more electronegative oxygen atom (Fig. 4.6). The molecule has a dipole moment of 1.85 D units, and each O—H bond can be considered to have 33% ionic character.

The predominance of hydrogen bonding in determining the properties of water is amply illustrated by the structure of ice (Fig. 4.7). This structure gives the impression of being determined exclusively by the four hydrogen bonds in which each H_2O molecule participates, two as hydrogen donor and two as acceptor. The angle between the hydrogen atoms in the water molecule (104.5°) is very close to the ideal for tetrahedral packing, 109.5°. Due to the hydrogen bonding, the crystal structure of ice is much more open than other crystals because each molecule of ice has only four immediate neighbors, instead of the usual 12 in crystalline closest-packing of spheres. Only 42% of the volume of ice is filled by the van der Waals volume of the molecules, rather than the 74% observed in spherical close-packing. (The hydrogen bonds of ice decrease the van der Waals volume of each water molecule by 4.0 Å³. If this were not taken into account, the density of ice would suggest a packing volume of 54%.) The four hydrogen-bonded water molecules have their oxygen atoms 2.76 Å from the central oxygen atom, with the next nearest neighbors 4.5 Å away. The hydrogen bonds are shorter in ice than between isolated H_2O dimers (2.98 Å) because of cooperativity; participation of the hydrogen atom in a hydrogen bond as a donor causes the oxygen atom to be much more effective as an acceptor in a second hydrogen bond, and vice versa. There may also be a contribution from the order of the crystal, with so many hydrogen bonds being present simultaneously (see Sec. 4.4).

The partial negative charge on the oxygen atom is frequently described as being localized primarily on two lone-pair electrons that effectively project above and below the plane of the molecule, giving water a tetrahedral structure. Hydrogen bonding in ice and in H_2O
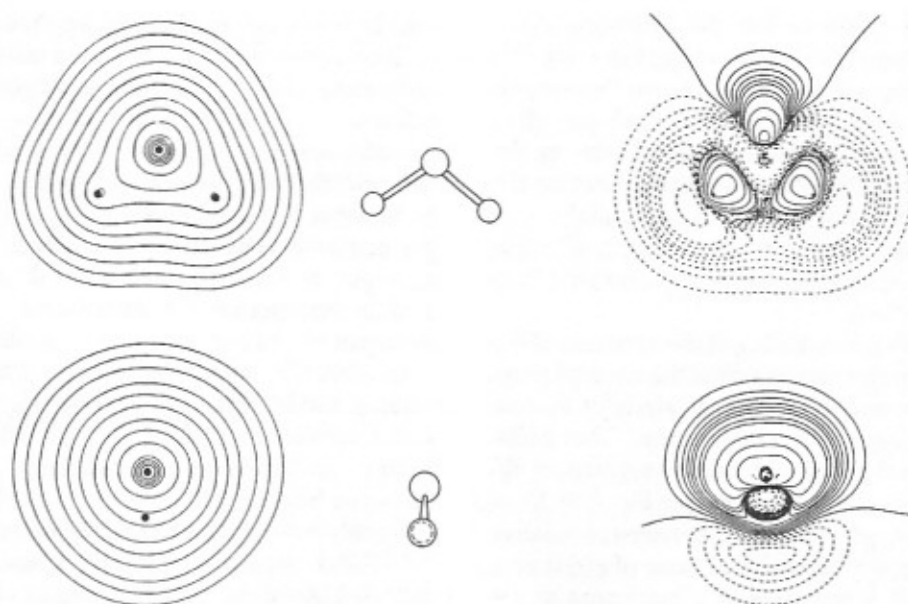
FIGURE 4.6

The electronic structure of the water molecule, shown as contour maps of electron density through the center of the molecule, viewed from two perpendicular angles. At *left* is shown the total electron density, illustrating the nearly spherical shape of the molecule. At *right* is shown the difference between this total electron density of the molecule and the density that would result from the superposition of individual spherical atoms. This gives the effect of covalent bonding on the electron density. The shift of electrons to the oxygen atom is indicated by the positive electron density (solid curves) on the O, the negative electron density (dashed lines) on the hydrogen atoms. (Adapted from I. Olovsson, *Croatica Chem. Acta* 55:171–190, 1982.
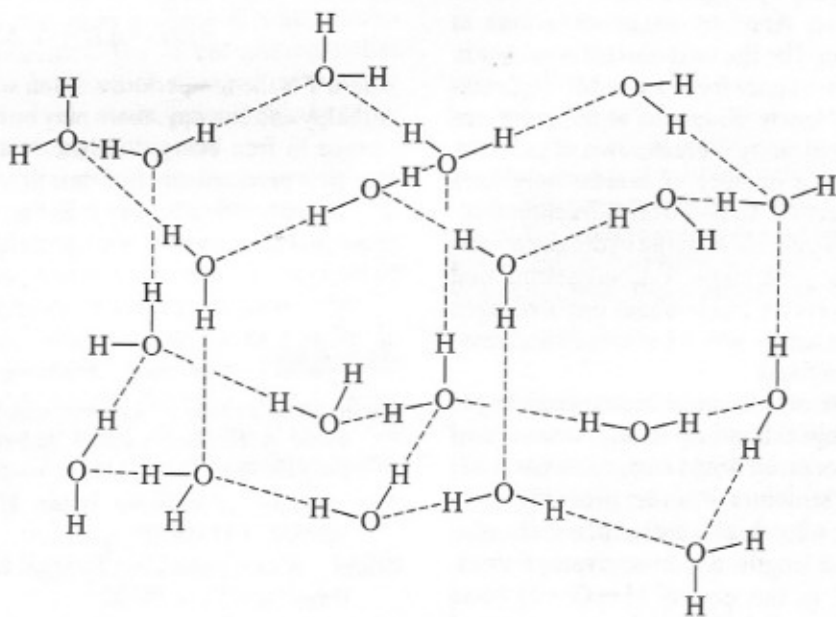

FIGURE 4.7

Structure of normal ice. Each H_2O molecule is involved in four hydrogen bonds (thin, dashed lines), each 2.76 Å between oxygen atoms. The water molecule is donor in two hydrogen bonds, acceptor in the other two. Substantial empty channels run between the molecules.

dimers in the gas phase directs the hydrogen atoms toward the positions of the lone-pair electrons, but this geometry is now thought to result primarily from repulsions between the hydrogen atoms. The lone-pair electrons on the oxygen atom are smeared out between the tetrahedral positions, lying in the plane bisecting the molecule (Fig. 4.6). There is probably only a slight energetic preference for a tetrahedral arrangement of hydrogen bonds, which in ice probably results primarily from packing considerations.

Although our understanding of the structure of liquid water is still somewhat uncertain, the unusual properties of water are well known, particularly its anomalous density change with temperature. The radial distribution function $g(r)$ of water shows substantial differences from those of other liquids (see Fig. 4.5). Measured with X rays, $g(r)$ reveals primarily the relative positions of the oxygen atoms. The value of $g(r)$ is zero for r less than 2.5 Å but rises to a maximum at $r =$ 2.82 Å at low temperatures; r at this maximum is just slightly greater than the distance between hydrogen-bonded neighbors in normal ice, and this distance increases somewhat at higher temperatures, up to 2.94 Å at 200°C. The number of nearest neighbors is approximately 4.4, greater than the 4.0 in ice but less than in most liquids. A second maximum of $g(r)$ occurs at 3.5 Å, but this peak is now thought to be an artifact of the diffraction analysis. The next maximum is at about 4.5 Å, which corresponds to the distance between pairs of oxygen atoms that are hydrogen bonded in ice to the same water molecule. Another maximum occurs at about 7 Å, which would be the next nearest neighbor in ice, but there is little evidence for more order. The maxima at 4.5 and 7 Å largely disappear at temperatures greater than 50°C, indicating a breakdown of even the local structure, and the number of nearest neighbors increases to about 5 at 100°C. Neutron diffraction analysis reveals the relative positions of the hydrogen atoms. They are not fixed beyond about 5 Å, suggesting that the water molecules rotate freely about one hydrogen bond — in contrast to ice, in which the tetrahedral crystal lattice keeps them fixed.

The X-ray and neutron-scattering data confirm the importance of hydrogen bonding in the structure of water, but how a disordered liquid state can exist is not clear. Part of water's structural disorder probably arises from flexibility in the water molecule itself, both in variations of O—H bond lengths and in an average variation of about ±15° in the normal H—O—H bond angle of 104.5°. The local tetrahedral arrangement of four hydrogen-bonded near neighbors of ice appears to persist in the liquid but with a fifth neighbor frequently present. Although in ice the rotation about the hydrogen bond is limited to one of three angles by the crystal

lattice, many orientations of neighboring molecules probably exist in the liquid, so the relative positions of hydrogen atoms on neighboring molecules are not well defined.

Many models of liquid water are built on the assumption that all the molecules are hydrogen bonded all of the time, but with a great variety of hydrogen-bond geometries and energies. In other models, each group is hydrogen bonded only a fraction of the time. Some models incorporate the experimental evidence that water is a mixture of two states in equilibrium: one state with relatively low enthalpy, low entropy, and large volume, similar to hydrogen-bonded ice; the other state with relatively high enthalpy, high entropy, and small volume, analogous to a normal liquid with much less hydrogen bonding. At present, no particular model is obviously more realistic than the others.

The strong hydrogen bonding among water molecules is clearly the basic explanation for many of the peculiar properties of this solvent. Hydrogen bonding also causes thermodynamic studies of phenomena in water to be particularly complex because changes in entropy and enthalpy tend to be mutually compensating, with relatively little change in free energy. For example, formation of hydrogen bonds in water should produce a favorable decrease in enthalpy H, but an unfavorable decrease in entropy S, because the molecules participating in the hydrogen bonds must be relatively fixed in orientation and proximity. These two contributions to the Gibbs free energy G tend to cancel out, because

$$\Delta G = \Delta H - T \Delta S \tag{4.11}$$

where T is the temperature. Even with large changes in enthalpy and entropy, there may be relatively little or no change in free energy. Other, relatively small effects may thus predominate in determining the free energy of any such transition in water. Rationalization of thermodynamic data in water is very tricky indeed.

References

Observed diffraction pattern and proposed models of liquid water. A. H. Narten and H. A. Levy. *Science* 165:447–454 (1969).

The structure of ordinary water. H. S. Frank. *Science* 169:635–641 (1970).

Water—A Comprehensive Treatise. F. Franks, ed. Plenum Press, New York, 1972.

Water revisited. F. H. Stillinger. *Science* 209:451–457 (1980).

Atom pair distribution functions of liquid water at 25°C from neutron diffraction. A. H. Narten et al. *Science* 217:1033–1034 (1982).

Supercooled water. C. A. Angell. *Ann. Rev. Phys. Chem.* 34:593–630 (1983).

The structure of polar molecular liquids. P. J. Rossky. *Ann. Rev. Phys. Chem.* 36:321–346 (1985).

Model for the structure of the liquid water network. E. Grunwald. *J. Amer. Chem. Soc.* 108:5719–5726 (1986).

Equation of state of a random network, continuum model of liquid water. A. R. Henn and W. Kauzmann. *J. Phys. Chem.* 93:3770–3783 (1989).

4.2.3 Aqueous Solutions

a. Solubilities

The solvent properties of liquid water are dominated by the polar character of the water molecule and by its relatively ordered structure. To be soluble in water, a molecule must occupy a certain volume, thereby disrupting the water structure at least within that volume. But the volume that is occupied by such a molecule in solution (the **partial molecular volume**) reflects not only the atomic volume of the molecule but also any changes it produces by rearranging the liquid around it. The measured partial volumes occupied by the various amino acid residues are given in Table 4.3 and are compared with their van der Waals volumes. The partial volumes occupied by the residues are generally 28–38% greater than their van der Waals volumes, which presumably reflects the open packing of water molecules around the molecules. The values are somewhat less when hydrogen bonding or charged groups are present on the residue side chain, presumably because they interact more strongly with water.

The solubility of a molecule in water depends on how much of the unfavorable aspects of creating a cavity in water are compensated by favorable interactions with the surrounding water molecules. A measure of the favorable interactions of a molecule with water, its **hydrophilicity**, is its relative tendency to equilibrate between being dissolved in the aqueous phase and being free in vapor (Fig. 4.8). The relative concentrations $[X]$ of the molecule in the vapor phase and in the aqueous phase at equilibrium gives the **partition coefficient** K_D:

$$K_D = \frac{[X]_{water}}{[X]_{vapor}} \tag{4.12}$$

The free energy of transfer from vapor to water is given by $-RT \ln K_D$, which is a measure of the hydrophilicity of a molecule; the most hydrophilic molecules have the most negative values. Only the relative values of the partition coefficients and hydrophilicities are relevant.
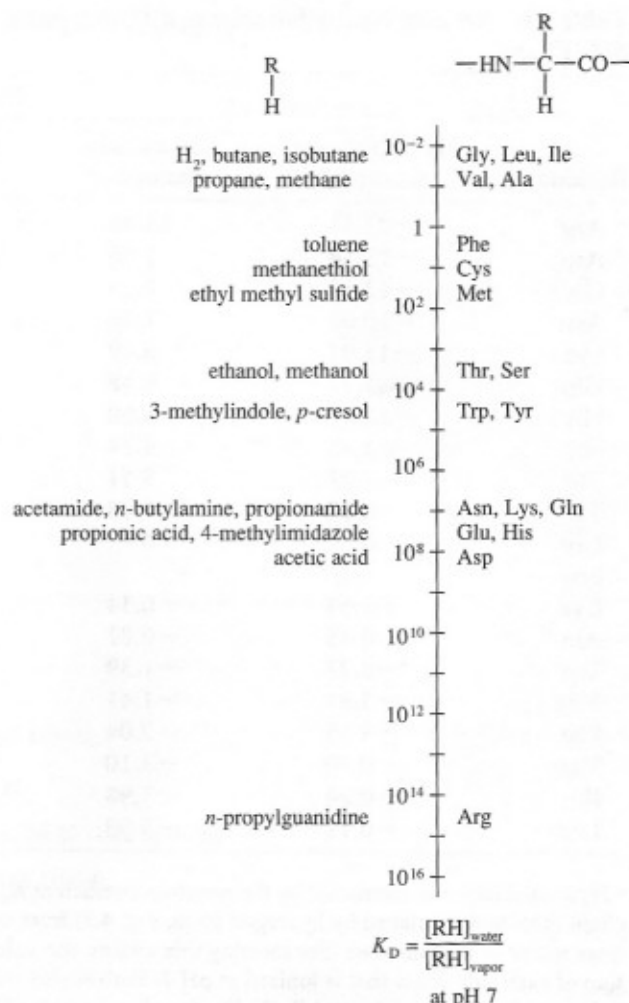

$$K_D = \frac{[RH]_{water}}{[RH]_{vapor}}$$

at pH 7

FIGURE 4.8

Relative hydrophilicities of amino acid side chains, measured by the partition coefficient between vapor of the appropriate small molecule and water at pH 7. The scale gives the equilibrium constant between the vapor and aqueous phase. The model compounds used for the side chains of the amino acid residues *(right)* are given on the *left.* (From R. Wolfenden et al., *Biochemistry* 20:849–855, 1981.)

The hydrophilicities of the amino acid side chains have been measured using model compounds in which the main chain was replaced by a hydrogen atom. For example, CH_4 was the model for the Ala side chain, toluene for Phe. A model for the peptide backbone is *N*-methylacetamide:

$$CH_3-\overset{\overset{\displaystyle O}{\|}}{C}-NH-CH_3 \tag{4.13}$$
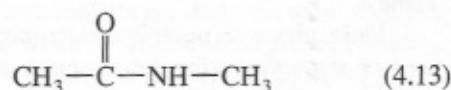
Ionized molecules have a negligible tendency to vaporize, so the observed partition coefficients for the corre-

Table 4.8 *Relative Hydrophilicities and Hydrophobicities of Amino Acid Side Chains*

Residue	Hydrophilicity[a] (kcal/mol)	Hydrophobicity (kcal/mol)			
		Side-chain analogues[a]	Amino acids[b]	N-acetyl amides[c]	Calculated[d]
Arg	−22.31	15.86	3.0	1.01	3.95
Asp	−13.34	9.66	2.5	0.77	3.81
Glu	−12.63	7.75	2.5	0.64	2.91
Asn	−12.07	7.58	0.2	0.60	1.91
Lys	−11.91	6.49	3.0	0.99	2.77
Gln	−11.77	6.48	0.2	0.22	1.30
His	−12.66	5.60	−0.5	−0.13	0.64
Ser	−7.45	4.34	0.3	0.04	1.24
Thr	−7.27	3.51	−0.4	−0.26	1.00
Tyr	−8.50	1.08	−2.3	−0.96	−1.47
Gly	0	0	0	0	0
Pro			−1.4	−0.72	−0.99
Cys	−3.63	−0.34	−1.0	−1.54	−0.25
Ala	−0.45	−0.87	−0.5	−0.31	−0.39
Trp	−8.27	−1.39	−3.4	−2.25	−2.13
Met	−3.87	−1.41	−1.3	−1.23	−0.96
Phe	−3.15	−2.04	−2.5	−1.79	−2.27
Val	−0.40	−3.10	−1.5	−1.22	−1.30
Ile	−0.24	−3.98	−1.8	−1.80	−1.82
Leu	−0.11	−3.98	−1.8	−1.70	−1.82

[a] Hydrophilicity was measured by the partition coefficient K_D of the model for each side chain (backbone replaced by hydrogen atom, Fig. 4.8) from vapor → water; hydrophobicity from water → cyclohexane. For ionizing side chains, the values were corrected for the fraction of each side chain that is ionized at pH 7. Both scales were normalized to zero for the value of Gly (A. Radzicka and R. Wolfenden, *Biochemistry* 27:1664–1670, 1988).

[b] Some values were measured from the relative solubilities of the amino acids in water and ethanol or dioxane (Y. Nozaki and C. Tanford, *J. Biol. Chem.* 246:2211–2217, 1971); others were extrapolated from these data (M. Levitt, *J. Mol. Biol.* 104:59–107, 1976).

[c] Measured from the partition coefficient between water and octanol of the N-acetyl amino acid amides (J. Fauchère and V. Pliska, *Eur. J. Med. Chem.* 18:369–375, 1983).

[d] Calculated from the hydrophobicities of the individual groups that make up each side chain, using data for the partition coefficient between water and octanol of many model compounds (M. A. Roseman, *J. Mol. Biol.* 200:513–522, 1988).

sponding nonionized molecules are corrected to the fraction of nonionized form present at pH 7. The measured hydrophilicity values of the amino acid side chains, normalized so that of Gly is zero, are given in Table 4.8.

Molecules with polar hydrogen-bond donors or acceptors strongly prefer the aqueous environment because they form hydrogen bonds to water approximately as well as other water molecules do. Maximum hydrophilicity is observed with ionized molecules and

with those that can act as both donor and acceptor in hydrogen bonds with water. The peptide bond is as hydrophilic as the side chains of Asn and Gln residues.

Nonpolar molecules do not interact as favorably with water. Some favorable van der Waals interactions probably take place between water and nonpolar solutes, but they are probably relatively weak because of the relatively open structure of water and because they are less favorable than the interactions between water molecules. Nonpolar molecules have low solubilities in

**Table 4.9** *Association in Water of Small Molecules Typical of Noncovalent Interactions in Proteins*

Type of interaction	Example	Association constant (M^{-1})
Salt bridge	$CH_3-CO_2^- \cdot H_2^+N{=}\overset{\overset{\textstyle NH_2}{\textstyle \vert}}{C}-NH_2$	0.5[a] 0.37[b]
	$CH_3-CO_2^- \cdot H_3^+N-(CH_2)_3-CH_3$	0.31[b]
	benzene$-O^- \cdot H_3^+N-CH_2-CH_2-OH$	0.20[c]
Hydrogen bond[i]	Formic acid dimers	0.04[d]
	Urea dimers	0.04[e]
	N-Methylacetamide dimers	0.005[f]
	δ-Valerolactam dimers	0.013[g]
Van der Waals	Benzene dimers	0.4[h]
	Cyclohexane · cyclohexanol	0.9[h]
	Benzene · phenol	0.6[h]

[a] C. Tanford, *J. Amer. Chem. Soc.* 76:945–946 (1954).

[b] B. Spriggs and P. Haake, *Bioorg. Chem.* 6:181–190 (1977).

[c] N. Stahl and W. P. Jencks, *J. Amer. Chem. Soc.* 108:4196–4205 (1986).

[d] A. Katchalsky et al., *J. Amer. Chem. Soc.* 73:5889–5890 (1951).

[e] J. A. Schellman, *Compt. Rendu Trav. Lab. Carlsberg Ser. Chim.* 29:223–229 (1955).

[f] I. M. Klotz and J. S. Franzen, *J. Amer. Chem. Soc.* 84:3461–3466 (1962).

[g] H. Susi et al., *J. Biol. Chem.* 239:3051–3054 (1964).

[h] S. D. Christian and E. E. Tucker, *J. Solution Chem.* 11:749–754 (1982).

[i] Interactions other than hydrogen bonding may contribute to the dimerization of these molecules, so the association constants are maximum values for hydrogen bonding.

water (Fig. 4.8), which can be ascribed to the hydrophobic effect, described later (see Sec. 4.3).

b. Interactions between Molecules in Water

Interactions between model compounds in aqueous solution provide information for analyzing the corresponding interactions in peptides and proteins. They are usually measured by the equilibrium constant K_{AB} for the association of two appropriate molecules A and B:

$$A + B \xrightleftharpoons{K_{AB}} A \cdot B$$

$$K_{AB} = \frac{[A \cdot B]}{[A][B]} \qquad (4.14)$$
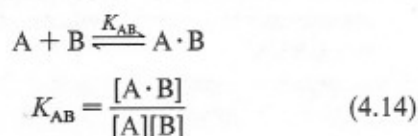
For two molecules to interact favorably in solution, they must overcome a loss of entropy and must interact with each other more strongly than they do individually with

water. For both of these reasons, interactions between individual molecules in solution are usually very weak; some values are given in Table 4.9. For comparison, small molecules that interact with each other as well as they do with water would be expected to have $K_{AB} = \frac{1}{55} M = 0.02\ M^{-1}$, where 55 M is the concentration of water molecules in liquid water.

By altering the water structure between them, molecules in water can interact to varying extents over a distance, without coming into contact and forming a complex. For example, large nonpolar surfaces have been shown to interact in water over distances as great as 25 Å. The interaction energy does not vary gradually with distance between the surfaces but exhibits oscillations with an average periodicity of 2.5 Å, approximately the diameter of a water molecule. The atomic structure of the water between such surfaces largely determines the apparent interaction between them; the
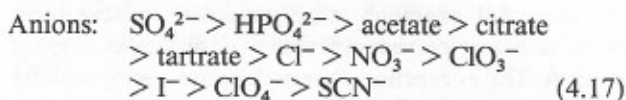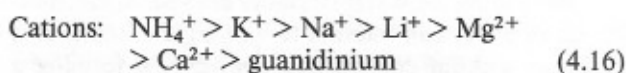
most favorable interactions occur when the distance between the surfaces is compatible with integral numbers of layers of water molecules. In these cases, the apparent interaction between the two surfaces is largely indirect and occurs because of the unfavorable situation of the water molecules between them.

Electrostatic interactions in water are less than those in other solutes because of water's high dielectric constant, which results from the tendency of the large dipoles of water molecules to align with any electric field. The dielectric constant of pure water at 25°C is 78.5, and it decreases at higher temperatures because thermal motion overcomes the orienting effects of the water dipoles. When small diffusible ions such as Na^+ and Cl^- are present in water, the apparent dielectric constant of the solution increases because the ions tend to concentrate in the vicinity of charges of the opposite sign. This **Debye–Hückel screening** is often described by an effective dielectric constant that increases with increasing distance d between the charges:

$$D_{eff} = D_{H_2O} \exp(+\kappa d) \qquad (4.15)$$

where κ is a parameter that is proportional to the square root of the ionic strength. The parameter $1/\kappa$ is known as the *Debye screening distance,* a measure of the distance over which electrostatic effects are damped out by the mobile ions. At physiological ionic strength, about 150 mM, the Debye distance is about 8 Å, but it ranges from about 300 Å in 10^{-4} M salt to 3 Å at 1 M.

The relatively ordered structure of liquid water is susceptible to alteration by high concentrations of other molecules. Changes in the physical properties of water, such as its surface tension, are reflected in the interactions of aqueous solutions with other molecules; both reflect the structure of water at a physical interface, but a detailed explanation is not currently available. Salts in the **Hofmeister series** are known to have such effects when present at concentrations in the range of 0.01 – 1 M. This series was first observed over 100 years ago by Hofmeister in his work on the effectiveness of salts in precipitating serum globulins and has been encountered over and over again in a variety of phenomena. The Hofmeister series for cations and anions are

Cations: $NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+}$
$> Ca^{2+} >$ guanidinium $\qquad (4.16)$

Anions: $SO_4^{2-} > HPO_4^{2-} >$ acetate $>$ citrate
$>$ tartrate $> Cl^- > NO_3^- > ClO_3^-$
$> I^- > ClO_4^- > SCN^- \qquad (4.17)$

The effects of cations and anions are usually independent and additive, with the anion having the larger effect. The first ions of each series disrupt the structure of water, markedly increase its surface tension, and decrease the solubility of nonpolar molecules (i.e., "salt out"). The last ions of each series generally increase the structure of water, have less effect on its surface tension, and increase the solubility of nonpolar molecules (i.e., "salt in"). The dividing points between the two effects are usually taken as Na^+ and Cl^-; NaCl is approximately neutral in this respect.

Nonionic molecules can also affect the physical properties of water. The best example is urea,

$$\begin{array}{c} O \\ \parallel \\ H_2N-C-NH_2 \end{array} \qquad (4.18)$$
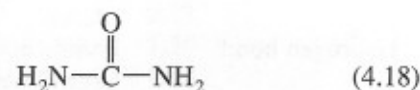
which hydrogen-bonds to water and at high concentrations disrupts the usual aqueous hydrogen-bond network.

Although all these additives seem to affect the properties of bulk water, a major part of their effects on the solubilities of nonpolar compounds arises from the fact that they are excluded from the solvent–nonpolar-solute interface. This interface is crucial for determining aqueous solubility. Exclusion of an additive from the air–water interface is the reason for the increase in surface tension of the bulk solvent.

References

Water and proteins. I. The significance and structure of water; its interaction with electrolytes and non-electrolytes. II. The location and dynamics of water in protein systems and its relation to their stability and properties. J. T. Edsall and H. A. McKenzie. *Adv. Biophys.* 10:137–207 (1978); 16:53–183 (1983).

Waterlogged molecules. R. Wolfenden. *Science* 222:1087–1093 (1983).

Amino acid, peptide, and protein volume in solution. A. A. Zamyatnin. *Ann. Rev. Biophys. Bioeng.* 13:145–165 (1984).

Theory of protein solubility. T. Arakawa and S. N. Timasheff. *Methods Enzymol.* 114:49–77 (1985).

The Hofmeister effect and the behaviour of water at interfaces. K. D. Collins and M. W. Washabaugh. *Quart. Rev. Biophys.* 18:323–422 (1985).

Solvation forces and liquid structure, as probed by direct force measurements. J. Israelachvili. *Acc. Chem. Res.* 20:415–421 (1987).

Forces between surfaces in liquids. J. N. Israelachvili and P. M. McGuiggan. *Science* 241:795–800 (1988).

Characteristic thermodynamic properties of hydrated water for 20 amino acid residues in globular proteins. M. Oobatake and T. Ooi. *J. Biochem.* 104:433–439 (1988).

4.3 The Hydrophobic Interaction

Electrostatic, hydrogen-bond, and van der Waals interactions between two molecules in an aqueous environment are not particularly favorable energetically (Table 4.9) because there are comparable competing interactions between the molecules and the water surrounding them. The interactions with water are not so favorable, however, with nonpolar surfaces. Water is a very poor solvent for nonpolar molecules compared with most organic liquids. Nonpolar molecules cannot participate in the hydrogen bonding that appears to be so important in liquid water, and aqueous solutions of such molecules have many anomalous physical properties. This relative absence of interactions between nonpolar molecules and water causes interactions among the nonpolar groups themselves to be much more favorable than would be the case in other solvents, so nonpolar molecules greatly prefer nonpolar environments. This preference of nonpolar atoms for nonaqueous environments

has come to be known as the **hydrophobic interaction.** It is a major factor in the stabilities of proteins, nucleic acids, and membranes, and it has some unusual characteristics.

4.3.1 The Hydrophobic Interaction in Model Systems

The magnitude of the hydrophobic interaction is usually measured by the free energy of transfer ΔG_{tr} of a nonpolar molecule in the gas, liquid, or solid state to water. The free energy of transfer is positive, indicating that the nonpolar molecule prefers a nonaqueous environment. The thermodynamics of transfer between these phases of a molecule the size of cyclohexane are illustrated in Figure 4.9 at two temperatures; the thermodynamics of transfer to water are anomalous in being markedly temperature dependent. In considering the thermodynamics of this transfer, keep in mind that the enthalpy change ΔH reflects the difference in the magnitude of the noncovalent interactions between molecules that occur in the two phases, whereas the entropy change ΔS
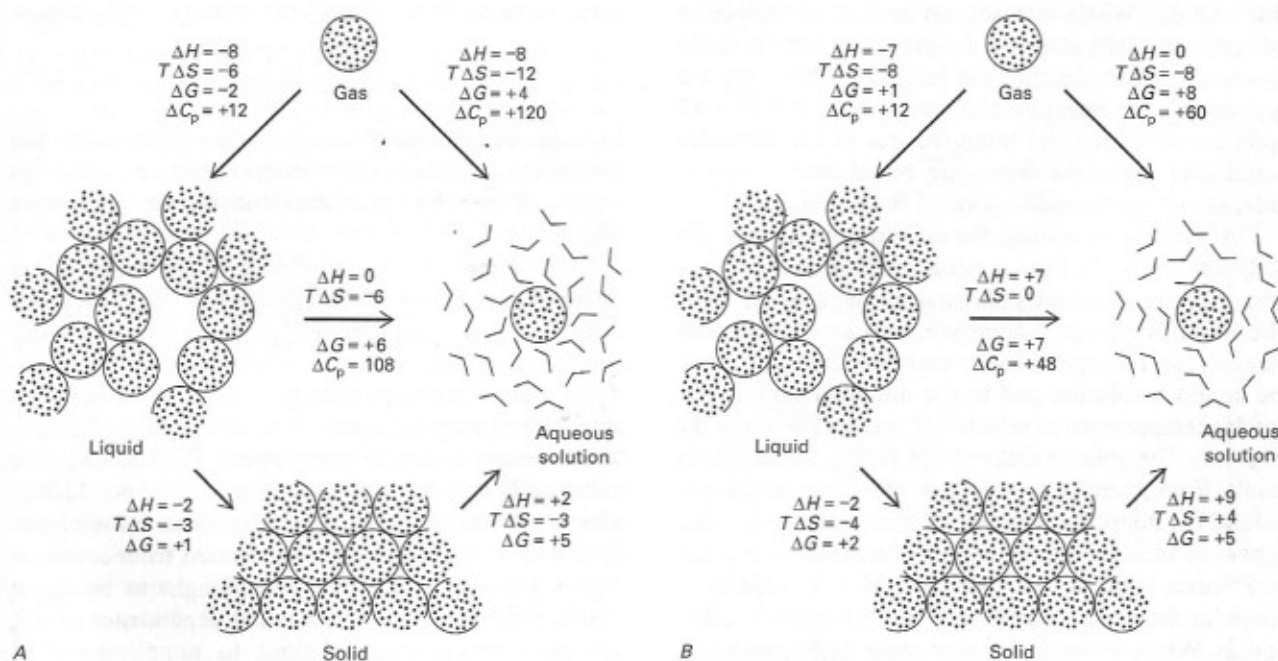

FIGURE 4.9

Typical thermodynamics of transfer of a nonpolar molecule the size of cyclohexane between the gas, liquid, and solid phases and aqueous solution at two temperatures: (A) T_H, approximately 20°C, where $\Delta H_{tr} = 0$ for transfer between liquid and water; and (B) T_S, approximately 140°C, where $\Delta S_{tr} = 0$. The values of ΔH, $T \Delta S$, and ΔG are in units of kcal/mol, that of ΔC_p in units of cal/(K·mol).

reflects the difference in disorder. Transfering a solute molecule to a liquid involves (1) creating a suitable cavity in the liquid, (2) introducing the solute molecule into the cavity, and (3) rearranging the solute and the surrounding liquid molecules to maximize the interactions between them. The observed thermodynamics of transfer between the two phases are the net effect of all three factors, so physical interpretation of thermodynamic parameters for transfer is not always straightforward. This is especially the case with water, in which the high degree of hydrogen bonding results in a more negative enthalpy and also a more positive entropy due to the necessity of fixing the positions of the interacting molecules; the enthalpy and enthropy have compensating effects on the free energy (Eq. 4.11). Nevertheless, analysis of the temperature dependence of the hydrophobic interaction is crucial for understanding its physical basis.

The thermodynamics of transfer data indicate that it is the aqueous solution of nonpolar molecules that shows the anomalous temperature-dependent physical properties. The differences in the thermodynamic parameters for transfer of nonpolar molecules to water from their gaseous, liquid, or solid states simply reflect the normal thermodynamic differences between these three states. For example, a nonpolar liquid has favorable van der Waals interactions among its molecules that are essentially absent in the gas phase, but the liquid also has less disorder than the gas; such differences are apparent in the negative changes of both ΔH and ΔS upon transfer from the nonpolar gas to the nonpolar liquid. Changes of the same type, but of smaller magnitude, occur upon solidification of the liquid.

At room temperature, the unfavorable transfer of a nonpolar molecule from a nonpolar liquid to water is primarily a result of the unfavorable change in entropy. The enthalpy change is approximately zero at this temperature, so there are similar enthalpic interactions in the aqueous solution and in the nonpolar liquid. The precise temperature at which $\Delta H_{tr} = 0$ is known as T_H (Fig. 4.9). The unfavorable entropy change is thought to result from increased ordering of water molecules around the nonpolar molecule. These water molecules appear to be more tightly packed than those of normal bulk water because the measured partial volumes of nonpolar molecules are smaller in water than in other liquids. Water molecules cannot make hydrogen bonds to nonpolar solutes so they are imagined to satisfy their hydrogen-bond potential by forming a hydrogen-bonded "iceberg" network among themselves around the nonpolar surface in a "water-ordering effect." Extreme examples of such ordered water cages are observed in **clathrates**, ordered water structures that incorporate apolar gases at low temperatures and high gas pressures. The water molecules are fully hydrogen-bonded in the clathrates, as they are in ice, although with nonoptimal geometries. A similar ordering of water molecules is thought to take place around a nonpolar solute molecule in aqueous solution in normal conditions, although to a lesser degree; the water molecules become more ordered and lose entropy, but their increased hydrogen bonding compensates by decreasing their enthalpy. Because the entropic factor dominates the unfavorable ΔG_{tr} to water, it was previously thought that the water-ordering effect is responsible for the low solubility of nonpolar molecules in water, but later analysis demonstrated just the opposite, that the water ordering occurs because it makes the interaction between water and a nonpolar molecule more favorable.

As the temperature is increased, the ordered water shell around the nonpolar solute tends to melt out and to become more like bulk water. This melting of the ordered water produces the anomalously large heat capacity C_p of this type of aqueous solution. The large C_p is the thermodynamic hallmark of aqueous solutions of nonpolar molecules. It causes the thermodynamics of such solutions to be markedly temperature dependent (Fig. 4.10) because the heat capacity defines the temperature dependence of both the enthalpy and entropy:

$$C_p = \frac{\partial H}{\partial T} = \frac{T\,\partial S}{\partial T} \qquad (4.19)$$

Its value is generally found to be proportional to the nonpolar surface area of the solute molecule exposed to water, as are the other thermodynamic parameters (Fig. 4.11).

The temperature dependence of the hydrophobic interaction provides important clues to its physical nature. At temperatures above T_H, the entropy of transfer decreases and becomes less unfavorable for transfer to water, but the enthalpy change becomes more unfavorable. The entropy of transfer from the nonpolar liquid to water becomes zero at temperature T_S (Fig. 4.9). The value of T_S was originally thought to be about 110°C, when ΔC_p was thought to be independent of temperature. The value of ΔC_p is now known to decrease at higher temperatures, and T_S is thought to be about 140°C (Fig. 4.10). The temperature dependence of ΔC_p primarily affects extrapolations to nonphysiological temperatures, however, so it is often convenient to approximate ΔC_p as a constant.

The large changes with temperature of ΔH_{tr} and ΔS_{tr} mostly compensate, and the value of ΔG_{tr} changes much less than they do (Fig. 4.10). Nevertheless, the magnitude of the hydrophobic interaction has a maxi-
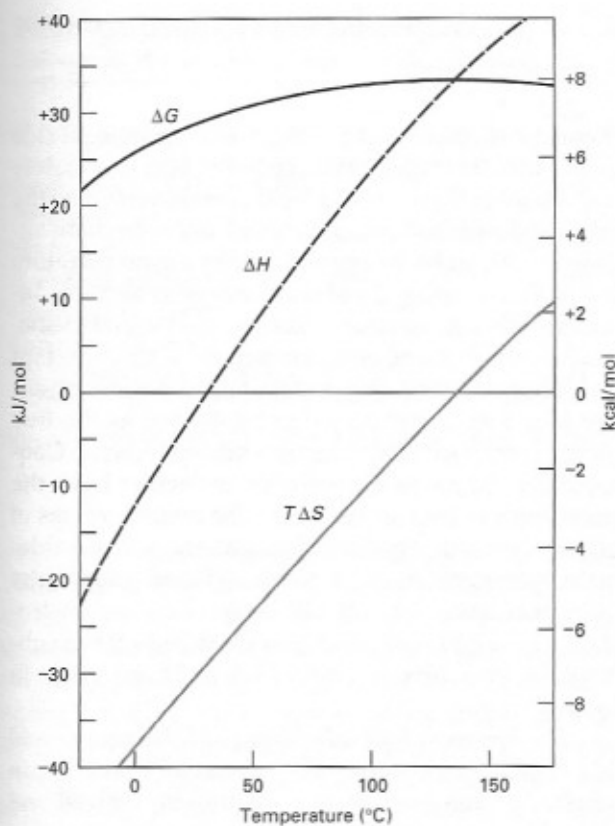
FIGURE 4.10
Typical thermodynamics of the free energy of transfer of a hydrocarbon from the liquid to aqueous solution, using pentane as an example. The strong temperature dependence of both the enthalpy and entropy difference between the two phases is a result of the different heat capacities of the two phases. The free-energy difference is the net difference between the enthalpic (ΔH) and entropic ($T \Delta S$) contributions. It reaches a maximum where $\Delta S = 0$, whereas the equilibrium constant (which is proportional to $-\Delta G/T$) reaches a maximum where $\Delta H = 0$. (Adapted from P. L. Privalov and S. J. Gill, *Adv. Protein Chem.* 39:191–234, 1988.)

mum value at one temperature, above and below which it decreases. Hydrophobicity measured as ΔG_{tr} is at a maximum at temperature T_S, whereas hydrophobicity as measured by the equilibrium constant for transfer is at a maximum at temperature T_H. It is important, therefore, to specify which measure of hydrophobicity is being used.

At the higher temperature T_S, the value of zero for the ΔS_{tr} from nonpolar liquid to water suggests that the net water-ordering effect has disappeared and that the water has now become a normal solvent. The difficulty with this interpretation is that the ΔC_p, which is thought to arise from the water-ordering effect, is still substan-

tial, although diminished (Fig. 4.9). Nevertheless, there appear to be no net interactions between a nonpolar solute molecule and water at this temperature because the ΔH_{tr} from the gas phase to water is approximately zero (Fig. 4.9). A possible explanation for this apparent discrepancy is that totally disordered water would have net repulsions with nonpolar solutes and that sufficient ordering persists at T_S to balance such repulsions. Not all of the details are clear, but the important point is that water becomes a much more normal solvent at high temperatures.

At the high temperature T_S, in the apparent absence of the net water-ordering effect, the low solubility of nonpolar molecules in water is observed to be due to the much weaker enthalpic interactions between the nonpolar molecules and water than between the nonpolar molecules in the nonpolar liquid and between the water molecules in bulk liquid water. The van der Waals interactions in a nonpolar liquid or solid are much greater than those between water and the nonpolar solvent. Likewise, the hydrogen bonds existing in water cannot be transferred to the nonpolar molecule but are disrupted by it when it is dissolved in the water. This is the cause of the hydrophobic interaction.

The hydrophobic interaction does not result from repulsions between water and nonpolar molecules as
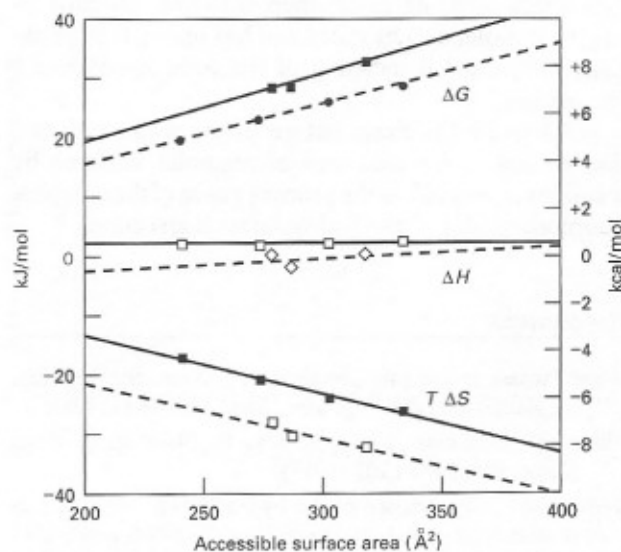

FIGURE 4.11
Thermodynamics of dissolution of hydrocarbon liquids into water at 25°C as a function of the accessible surface area of the hydrocarbon. The enthalpy change is virtually zero at this temperature. The dashed lines are for aliphatic molecules, the solid lines for aromatics. (Kindly provided by S. J. Gill.)

implied by the term. There are, in fact, favorable inter-actions between water and a nonpolar molecule at all accessible temperatures, those less than T_s (Fig. 4.9). The magnitudes of these interactions, however, are less than those of the van der Waals interactions in nonpolar liquids and those of the hydrogen bonding in liquid water.

The hydrophobic interaction results in a tendency of nonpolar atoms to interact with each other rather than with water. This interaction has the unusual prop-erty of decreasing in magnitude at lower temperatures, which results from the increasing tendency of water at lower temperatures to form the hydrogen-bonded clathrate-like structures around the nonpolar molecule. The water ordering effect increases the solubility of nonpolar molecules in water because it seems to be water's attempt to improve its interactions with the nonpolar molecule. The water-ordering effect is not re-sponsible for the low solubility of nonpolar molecules, as has often been assumed from the way the entropy change dominates the thermodynamics of transfer at room temperature. The water-ordering effect is respon-sible for the decrease in ΔG_{tr}, that is, the decrease in magnitude of the hydrophobic interaction, at low tem-peratures. It is important, therefore, not to use terms such as *hydrophobicity, hydrophobic interaction*, or *hy-drophobic effect* to refer to the water-ordering effect or to the resultant anomalous thermodynamic parameters. The water-ordering effect increases the solubility of nonpolar molecules in water and has opposite implica-tions to the usual meaning of the term *hydrophobic interaction.*

Although the exact nature of the water-ordering that occurs in the solvation of nonpolar surfaces by water is uncertain, it is the primary cause of the complex thermodynamics of the hydrophobic interaction.

References

Some factors in the interpretation of protein denaturation. W. Kauzmann. *Adv. Protein Chem.* 14:1–63 (1959).

"Iceberg" formation and solubility. K. Shinoda. *J. Phys. Chem.* 81:1300–1302 (1977).

Temperature dependence of the hydrophobic interaction in protein folding. R. L. Baldwin. *Proc. Natl. Acad. Sci. USA* 83:8069–8072 (1986).

Stability of protein structure and hydrophobic interaction. P. L. Privalov and S. J. Gill. *Adv. Protein Chem.* 39:191–234 (1988).

The hydrophobic effect: a reappraisal. P. L. Privalov and S. J. Gill. *Pure Appl. Chem.* 61:1097–1104 (1989).

Search for a realistic view of hydrophobic effects. N. Muller. *Acc. Chem. Res.* 23:23–28 (1990).

4.3.2 Hydrophobicities of Amino Acid Residues

The hydrophobicities of the individual amino acid side chains have been measured experimentally in a variety of ways, using the free amino acids, amino acids with the amino and carboxyl groups blocked, and side-chain an-alogues with the backbone replaced by a hydrogen atom (Fig. 4.8), and using a variety of nonpolar solvents in-cluding ethanol, octanol, dioxane, and cyclohexane. Because the hydrophobic interaction is an important component in stabilizing protein folded conformations (see Sec. 7.4), it will hereafter be defined as the free energy of transfer from water to a nonpolar liquid. Con-sequently, the more hydrophobic molecules have the more negative hydrophobicities. The relative values of ΔG_{tr} are relevant, not their absolute values, so the side-chain hydrophobicities are obtained by subtracting the value measured for Gly. Unfortunately, the hydro-phobicity values measured in various ways differ sub-stantially, so several representative scales are given in Table 4.8.

The apparent hydrophobicities of the amino acid side chains vary enormously, depending primarily on whether or not polar groups are present. Ionized and polar side chains interact strongly with water and have lower solubilities in nonpolar solvents because of the unfavorable energetics of placing a polar group in a nonpolar environment. The magnitude of this effect varies enormously, depending on the solvent and the molecule, and is probably the main source of the varia-tion in the hydrophobicity scales. In a nonaqueous sol-vent, the polar groups of the side chains can have vary-ing interactions with other polar groups of the peptide backbone and any of the solvent, including any water that is present. In this case, molecules that have polar groups can appear to be more hydrophobic than they really are. Perhaps for that reason, the most extreme values of hydrophobicity have been measured using models for the side chain alone, without the polypeptide backbone, and with the apolar solvent cyclohexane, which has the least polar nature (Table 4.8). Because of the widely varying hydrophobicities of amino acid side chains, the more neutral term **hydropathy** is often used to describe their relative preferences for aqueous and nonpolar environments.

The diversity of polar and nonpolar groups in amino acid side chains and in the polypeptide backbone makes it advisable to consider the individual groups rather than the side chain as a whole (Table 4.10). The free energy of transfer from water to nonpolar solvents of the nonpolar side chains is correlated with their sur-face areas, and there are remarkably large differences

Table 4.10 *Thermodynamic Parameters for the Transfer at 25°C from Nonpolar Solvent to Water of 1 Å² of Accessible Surface Area of Various Chemical Groups*

Chemical group	ΔG_{tr} (cal·mol^{-1}·Å^{-2})	ΔH_{tr} (cal·mol^{-1}·Å^{-2})	ΔC_p (cal·K^{-1}·mol^{-1}·Å^{-2})
Aliphatic: —CH$_3$, —CH$_2$—, CH	+8	−26	0.370
Aromatic	−8	−38	0.296
Hydroxyl	−172	−238	0.008
Amide & amino: —NH—, NH$_2$	−132	−192	−0.012
Carbonyl C: C=	+427	+413	0.613
Carbonyl O: =O	−38	−32	−0.228
Thiol and sulfur: —SH, —S—	−21	−31	−0.001

From T. Ooi and M. Oobataka, *J. Biochem.* 103:114–120 (1988).

among the hydrophobicities as measured in various ways (Fig. 4.12A). The values measured for the transfer of the other amino acids are dominated by their polar groups. That solvation by water is the predominant factor in hydrophobicity is illustrated by the excellent correlation of the surface areas of all the amino acid side chains with their free energy of transfer from vapor to cyclohexane (Fig. 4.12B); the polar groups are presumably equally uncomfortable in both phases. The heat

capacities of aqueous solutions of the side-chain analogues are directly proportional to their nonpolar accessible surface areas (Fig. 4.13), except that the side chains with ring structures, and methionine with its sulfur atom, give slightly lower values. There can be little doubt that the unusually large heat capacities of aqueous solutions of nonpolar molecules and the anomalous thermodynamics of the hydrophobic effect arise from the interactions of water with nonpolar atoms.
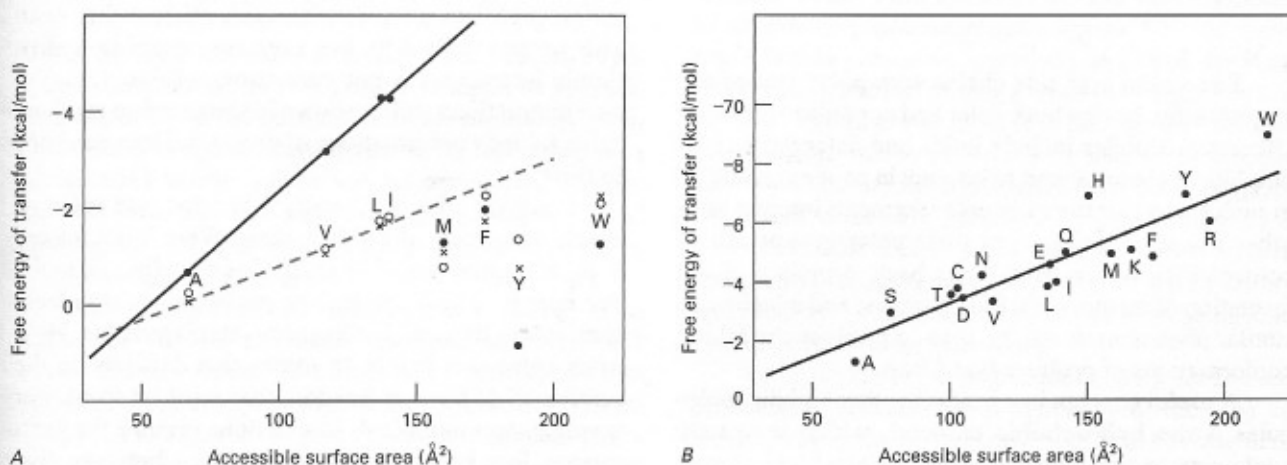

FIGURE 4.12
Relationship between the accessible surface areas of the nonpolar amino acid side chains and their free energies of transfer from water to nonaqueous solvent *(A)* and from vapor to cyclohexane *(B)*. The amino acid residues are designated by the one-letter code (see Table 4.3). The free energies of transfer in *A* are from Table 4.8: the ● points represent measurements with the side-chain analogues, using cyclohexane as the nonpolar solvent; the x points were obtained with the free amino acids and ethanol and dioxane as the nonpolar solvent; the ○ points were calculated from the hydrophobicities of the parts of each side chain. In *A* the slope of the solid line is 43 cal/Å², the dashed line 20 cal/Å². The free energies of transfer in *B* were from A. Radzicka and R. Wolfenden, *Biochemistry* 27:1664–1670 (1988). The slope of the line in *B* is 41 cal/Å².
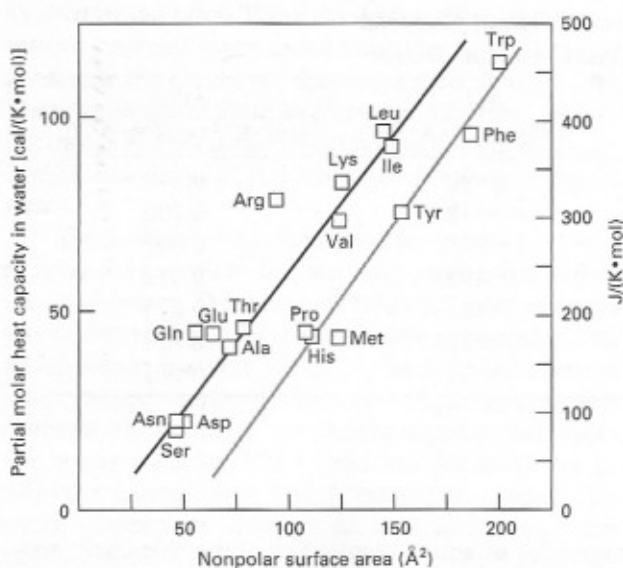
FIGURE 4.13

Correlation between the heat capacities in aqueous solution at 25°C with the accessible surface area of the nonpolar atoms of analogues of the amino acid side chains. The upper straight line fits all the side chains except those with ring structures and the sulfur-containing Met (lower line). The slope of the upper line is 0.72 cal/K·mol Å² (300 J/ K·mol nm²). (Adapted from G. I. Makhatadze and P. L. Privalov, *J. Mol. Biol.* 213:375–384, 1990.)

The amino acid side chains with polar groups are **amphiphilic,** having both polar and nonpolar segments. Other amphiphiles include lipids and detergents. Amphiphilic molecules tend to interact in aqueous solution in such a way that their nonpolar segments interact with other nonpolar groups and their polar groups are in contact with water. This is the basic principle of the formation of membranes, lipid bilayers, and micelles. A similar phenomenon can be seen to produce the folded conformations of proteins (see Chap. 6).

A useful concept in considering amphiphilic molecules is the **hydrophobic moment,** which is exactly analogous to the dipole moment of electrical charge (Sec. 4.1.2.b) but represents a vector from the hydrophilic to the hydrophobic parts of a molecule. The hydrophobic moment of a polypeptide chain in a particular three-dimensional structure is calculated from the vector sum of the contributions of each of the amino acid residues. This contribution is given by a vector that points from the C$^\alpha$ atom to the center of the side chain and whose length is proportional to the hydrophobicity of the side chain. This parameter can often account for the architecture and interactions of large molecules such as proteins.

References

Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients. D. J. Abraham and A. J. Leo. *Proteins: Struct. Funct. Genet.* 2:130–152 (1987).

Characteristic thermodynamic properties of hydrated water for 20 amino acid residues in globular proteins. M. Oobatake and T. Ooi. *J. Biochem.* 104:433–439 (1988).

Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. G. I. Makhatadze and P. L. Privalov. *J. Mol. Biol.* 213:375–384 (1990).

4.4 Intramolecular Interactions

The analysis presented in this chapter of the electrostatic, van der Waals, and hydrogen-bond interactions among the atoms of proteins indicates that all such interactions are weak in the presence of water. Ionized or polar groups interact with water almost as favorably as they interact with other suitable ionized or polar groups (Table 4.9), and it is energetically unfavorable to remove them from aqueous solution (Fig. 4.8). Nonpolar groups prefer to interact with each other rather than with water (Table 4.8), but even the resulting hydrophobic interaction is not very strong (Table 4.9). Yet such interactions will be shown in Chapter 6 to produce stable folded conformations of proteins. How can they do this?

For molecules to interact, they must lose entropy, which is energetically unfavorable. Were it not for entropy, all matter would be solid. This contribution to the free energy, which arises from molecules having freedom, makes the liquid and gaseous states possible. How much entropy is lost in an interaction depends on the number of degrees of freedom that must be fixed. For example, van der Waals interactions require the least entropy loss because only the distance between two atoms is fixed (Fig. 4.2), whereas hydrogen bonding requires that both proximity and orientation be fixed to some extent (Figs. 4.3 and 4.4).

Entropic considerations are especially important when two or more interactions are possible simultaneously in a single molecule because, in favorable cases, much less entropy need be lost in the second and subsequent interactions than in the first. Two interactions that can occur simultaneously can be much more favorable energetically than might be expected from their individual strengths.

4.4.1 Effective Concentrations

The magnitude of entropic cooperativity can be illustrated by intramolecular interactions. Two parts of the same molecule can interact without losing as much entropy as must be lost to bring two independent molecules together. Because the two parts of a molecule are already fixed to some degree in proximity and orientation, only some fraction (which depends on the molecule and the interaction) of the internal flexibility of the molecule has to be lost in the intramolecular interaction.

Intramolecular and bimolecular examples of the same interaction can be compared by means of the ratio of their equilibrium constants, which for the intramolecular interaction is dimensionless and for the intermolecular interaction has dimensions of (concentration)$^{-1}$. Therefore, the ratio of the two has the dimensions of concentration, which can be thought of as the **effective concentration** of the two groups when they are part of the same molecule in the intramolecular interaction:

$$A\!-\!B \underset{K_{intra}}{\rightleftharpoons} A \cdot B \qquad (4.20)$$

$$A + B \underset{K_{inter}}{\rightleftharpoons} A \cdot B \qquad (4.21)$$
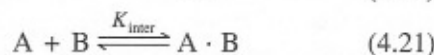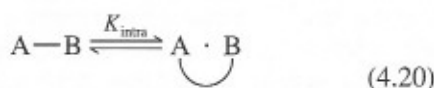
$$\frac{K_{intra}}{K_{inter}} = \text{effective concentration of A}\!-\!\text{B} \qquad (4.22)$$

It was thought for a long time that the maximum effective concentration of two groups in aqueous solution was about 55 M, the concentration of pure water, when one group could be considered to be immersed in a liquid environment of the second component. Consequently, one often finds instances of the magnitudes of intramolecular and intermolecular interactions being interconverted by using the factor of 55 M.

Many experimental measurements have been made for various chemical reactions, however, and much greater values of effective concentrations are generally found; representative examples are given in Table 4.11. These examples represent chemical reactions involving reversible covalent bond formation that can be considered analogous to noncovalent interactions. The covalent nature of these interactions, however, probably increases the magnitude of the entropic effect, due to the more stringent geometrical requirements of covalent bond formation. In any case, the first three examples in the table (A–C) involve flexible molecules with relatively free rotations about three single bonds, which must be restricted to form the product. In spite of this considerable entropic loss, the effective concentrations measured are in the range of 10^2 to 10^5 M. Therefore, merely keeping two groups in reasonable proximity by linking them covalently through several bonds causes their concentration relative to each other to be much higher than would be possible if the groups were on separate molecules, even in the most concentrated liquid state. The last example of Table 4.11 (D) has an enormous effective concentration of 5×10^9 M, which is undoubtedly due primarily to the small entropy difference between the molecule with and without the anhydride interaction. In this case, the planar aromatic structure of the molecule keeps the carboxyl groups in close proximity whether or not the anhydride is present. The very small increase in flexibility and entropy that occurs when the anhydride interaction is broken results in an enormous effective concentration that is close to the maximum considered possible theoretically (approximately 10^{10} M). Of course, other factors can cause large apparent effective concentrations, such as strain in the molecule that is relieved upon forming the interaction, but numerous examples of large values are known that illustrate the entropic contribution.

When there is no entropic difference between the molecules with and without the interaction, the effective concentration is at its maximum value. This value depends on the type of interaction. Those in which the proximity and orientation of the interacting groups are very important, as in a hydrogen bond and especially when a covalent bond is formed, have very high maximum effective concentrations, up to 10^{10} M. When these factors are not so important, as in van der Waals interactions, the interacting groups have significant degrees of freedom and have less entropy to gain upon dissociating, so lower values of maximum effective concentrations apply. Even in this last case, though, the maximum values are substantially greater than 55 M. The reason for this high maximum value is that the molecules of a liquid have a high degree of translational and rotational freedom, so they are far from being in the optimal situation for interacting.

Unfortunately, the magnitudes of the effective concentrations expected for interactions of the type observed in proteins are not known. Only in the case of the disulfide interaction between thiol groups have values been measured in proteins (Sec. 7.5.4). The maximum value observed is somewhat greater than 10^5 M, but the disulfide bond is a covalent interaction, which tends to enhance the effective concentration. Hydrogen bonds are moderately sensitive to orientation and probably have a partial covalent character, so substantial maximum values would be expected, but probably much less than 10^{10} M and less than those involving disulfide bonds. Ionic and hydrophobic interactions are not stereochemically very stringent, so maximum values of 10^2–10^3 M may apply in these instances.

Table 4.11 *Selected Examples of Measured Values of Effective Concentrations of Two Reactive Groups in Small Molecules*

Example	Equilibrium reaction	Effective concentration (M)
A		3.7×10^3 [a]
B		2×10^2 [b]
C		1.9×10^5 [a]
D		5.4×10^9 [a]

[a] From A. J. Kirby, *Adv. Phys. Org. Chem.* 17:183–278, (1980).

[b] From M. H. Chau and J. W. Nelson, *FEBS Letters* 291:296–298 (1991).

In contrast to the very high effective concentrations that are possible when interacting groups are held in the appropriate proximity and orientation, constituent groups that are kept apart by the structure of their molecules have very low, or zero, effective concentrations. Intramolecular interactions are much more sensitive to their environment than are interactions between independent molecules in the liquid state.

Detailed explanations of most values of the effective concentrations measured are complicated by the presence of unfavorable steric or physical interactions in the molecules with or without the interaction. Conse-quently, there is no ideal example with which to illustrate the solely entropic contribution to the effective concentration, but the many experimental examples available indicate that the effect is very substantial.

References

Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. M. I. Page and W. P. Jencks. *Proc. Natl. Acad. Sci. USA* 68:1678–1683 (1971).

Effective molarities for intramolecular reactions. A. J. Kirby. *Adv. Phys. Org. Chem.* 17:183–278 (1980).

effects on water activity. M. T. Record et al. *Quart. Rev. Biophys.* 11:183–278 (1980).

4.4.2 Multiple Interactions

Multiple groups on a molecule can behave very differently from the same groups in isolation. For example, individual ions do not associate very strongly in aqueous solution (Table 4.9) because of their favorable entropies and strong interactions with water; but a **polyelectrolyte** molecule that has a number of such charged groups binds ions of the opposite charge very tightly due to interactions among the charged groups. Being part of the same molecule, the charged groups are constrained to be close to each other by the covalent bonds. The electrostatic repulsion between groups with the same charge is compensated by the counterions that they attract from the solution and bind very tightly; just how tightly depends on the charge density of the polyelectrolyte and the valence of the counterions. The degree of binding is almost independent of the concentration of counterions in the bulk solvent because the diminution of the electrostatic repulsions in the polyelectrolyte is energetically much more important than is the equilibration of the ions with the bulk solvent. The counterions are not necessarily bound at specific sites, and they can retain their water of hydration and move in an unrestricted and random manner along the polyelectrolyte chain. This phenomenon can be very important for binding other ligands that have the same charge as the counterions; displacement of the counterions into a very dilute bulk solution by the binding of the ligand can provide a strong driving force for its binding.

Another means of compensating unfavorable electrostatic repulsions in a polyelectrolyte is suppression of the ionization of a fraction of its groups. Consequently, groups on a polyelectrolyte can have pK_a values that are very different from those found when they are isolated. These electrostatic effects among multiple groups on a polyelectrolyte are important for the function of proteins, but especially for the function of nucleic acids, with their numerous phosphate groups, and for the interactions of nucleic acids with proteins (Sec. 8.3.2).

References

The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. G. S. Manning. *Quart. Rev. Biophys.* 11:199–246 (1978).

Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening and ion

4.4.3 Cooperativity of Multiple Interactions

The simultaneous presence of multiple interactions in a single molecule produces cooperativity between them, and together they can be much stronger than might be expected from the sum of their individual strengths. Cooperativity is essential for proteins, in which the noncovalent interactions are intrinsically very weak (Table 4.9). Only when such interactions cooperate is a stable single conformation produced.

Consider an unfolded polypeptide chain in which two groups A and B are capable of interacting favorably as in a hydrogen bond, a salt bridge, or a nonpolar hydrophobic interaction:


$$unfolded \qquad\qquad (4.23)$$
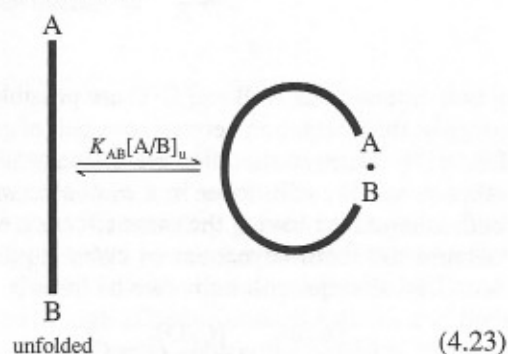
The observed equilibrium constant for interaction of the two groups, $K_{obs,U}$, can be expressed as

$$K_{obs,U} = K_{AB}[A/B]_U \qquad (4.24)$$

where K_{AB} is the association constant measured with groups A and B on individual molecules (Table 4.9) and $[A/B]_U$ is the effective concentration of the two groups relative to each other on the unfolded polypeptide U. Groups attached to moderate sized random polypeptides have effective concentrations in the range 10^{-2}–10^{-5} M, depending on their relative positions in the polypeptide chain (see Sec. 5.2). With typical values of K_{AB} (Table 4.9), values for the observed equilibrium constant $K_{obs,U}$ of between 4×10^{-3} and 10^{-7} are expected for individual hydrogen bonds, salt bridges, and so on. Consequently, a single interaction between two groups on a polypeptide chain is not expected to be stable unless the groups are close in the covalent structure in such a way that they have an especially high effective concentration.

Multiple interactions among two or more pairs of groups on the same molecule are often not independent but assist or interfere with each other. The following

equilibria are possible with two pairs of groups on a polypeptide:

$$(4.25)$$
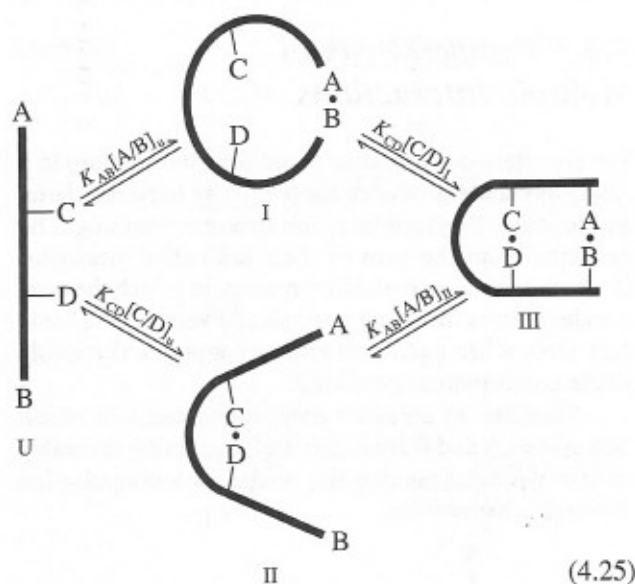
If both interactions $A \cdot B$ and $C \cdot D$ are possible simultaneously, the interaction between one pair of groups will frequently increase the effective concentration of the other pair. This will occur in a mutual manner, with both interactions having the same effect on each other because the thermodynamics of cyclic equilibria (see Sec. 8.4.1.d) require that the two be linked:

$$\frac{[A/B]_{II}}{[A/B]_{U}} = \frac{[C/D]_{I}}{[C/D]_{U}} = \text{Coop} \qquad (4.26)$$

The factor *Coop* is the degree of cooperativity between the two interactions. Consequently, each interaction can be more stable in the presence of the other interaction than when it takes place alone.

If additional groups that may also interact simultaneously are present on the polypeptide chain, these equilibria are extended in a similar way. The overall equilibrium constant between the final state (with all the interactions present) and the unfolded state (with none) is the product of the individual equilibrium constants along any of the conceivable reaction paths; for example,

$$K_{\text{net}} = (K_{AB}[A/B]_{U})(K_{CD}[C/D]_{I})(K_{EF}[E/F]_{III}) \cdots \qquad (4.27)$$

The value of K_{net} is independent of the reaction path, so we need not know or propose a specific "folding pathway."

The final conformation is stable — that is, populated by most of the molecules — only if the value of K_{net} is greater than unity. Consider a series of weak interactions. The first will be very weak, with an equilibrium

constant of 10^{-3}–10^{-7}. But the presence of the first interaction can increase the effective concentration of the second pair of groups, so the equilibrium constant for the second interaction may be somewhat larger than that of the first, by the factor Coop. If the second equilibrium constant is also less than unity, however, the product of the two equilibrium constants is even smaller than the first (Fig. 4.14). Similarly, the net stabilities of conformations with additional weak interactions are even lower than that of the conformation with a single interaction. This process continues until the effective concentrations of additional interacting groups are suf-
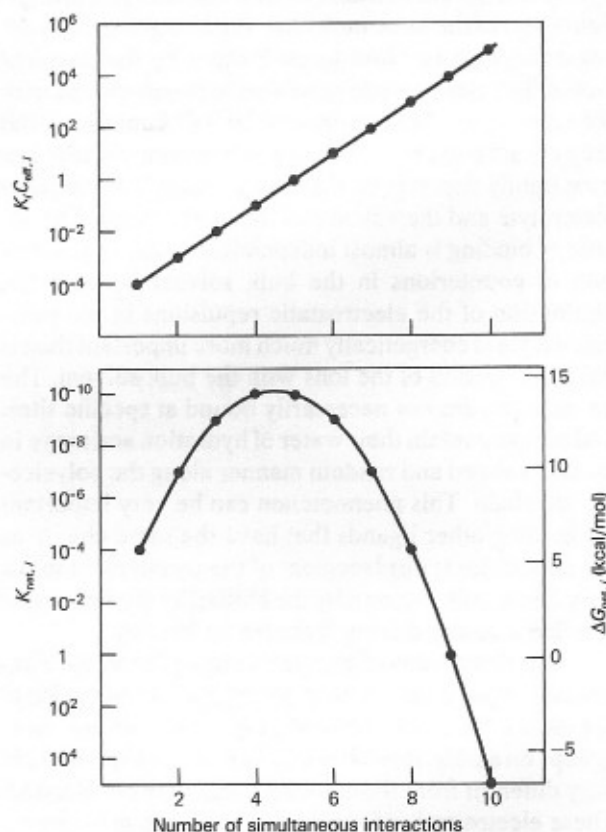

FIGURE 4.14

Hypothetical illustration of the cooperativity produced by multiple weak interactions. Up to 10 interactions are possible simultaneously, and the contribution of the *i*th interaction is given at the *top*. The initial interaction has an equilibrium constant of 10^{-4}, and each additional interaction is 10 times stronger than the previous one as a result of an increase in the effective concentration. The overall equilibrium constant $K_{\text{net},i}$ (*bottom*) is the product of the contributions of the i interactions present (see Eq. 4.27). Only with 10 such interactions is $K_{\text{net},i} > 1$, implying stability of the folded structure. The free energy of each state relative to U is given by $\Delta G^{\circ}_{\text{net},i} = -RT \ln K_{\text{net},i}$, with the scale on the right pertaining to 25°C.
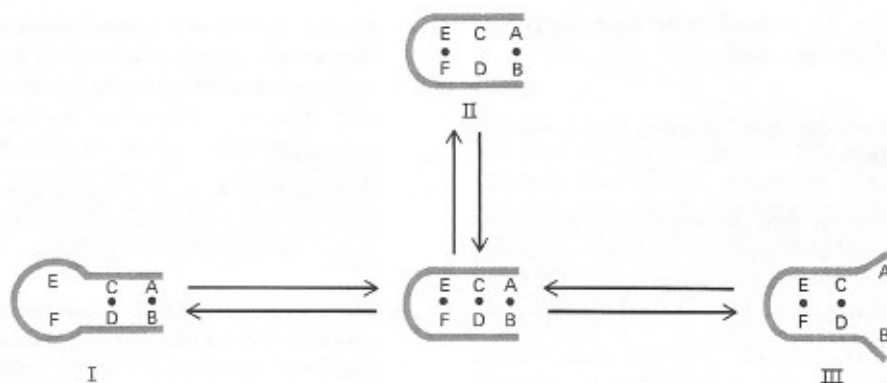
FIGURE 4.15

Simple schematic diagram of cooperativity among three simultaneous interactions occurring between groups A and B, C and D, and E and F. The strength of each interaction is determined by the effective concentration of the two groups when they are not interacting, as in conformations I, II, and III. Assuming there are no other considerations, the value of the effective concentration will be inversely proportional to the degree of flexibility permitted. Therefore, the most stable interactions should be those between groups that are held most rigidly by the other interactions, in this case C and D, and the stability of each interaction should depend on the stabilities of all the others.

ficiently increased to make the equilibrium constant for each additional interaction greater than unity. The value of K_{net} then increases in magnitude with each additional interaction. A sufficient number of simultaneous weak interactions can make the value of K_{net} greater than unity and the folded conformation stable.

An example with Coop = 10 for each additional interaction is given in Figure 4.14. Partially folded structures, those with incomplete stabilizing interactions, are unstable relative to the initial and final states, which means that the transition is cooperative. The degree of cooperativity will be even greater if the intermediate structures have nonbonded groups in unfavorable environments, such as polar groups present in nonpolar environments without being hydrogen bonded. Such situations have been ignored here.

In summary, weak interactions are expected to stabilize a particular folded conformation only when they cooperate so that the interacting groups have very high effective concentrations in that structure. The effective concentration of two groups in a folded structure depends on the extent to which the groups are held in proximity when not interacting (Fig. 4.15), which in turn depends on the stability of all the surrounding interactions. All parts of such a structure, therefore, are expected to be mutually dependent to varying degrees.

As just described, the contribution of each interaction to net stability of the folded structure should depend on the effective concentration of the interacting groups in that folded structure. If the groups are on the surface or in a flexible part of the folded structure, their effective concentration will be low and the interaction will provide little, if any, net stability. Breaking that interaction will have little effect on the folded state. On the other hand, groups in relatively rigid parts of the folded structure will have high effective concentrations, and their interaction will provide a substantial contribution to the net stability of the conformation; removing or altering such an interaction would have a large effect on the stability of the folded conformation.

References

An empirical approach to protein conformation stability and flexibility. T. E. Creighton. *Biopolymers* 22:49–58 (1983).

Dissecting the roles of individual interactions in protein stability: lessons from a circularized protein. D. P. Goldenberg. *J. Cell. Biochem.* 29:321–335 (1985).

Strategy for analysing the co-operativity of intramolecular interactions in peptides and proteins. A. Horovitz and A. R. Fersht. *J. Mol. Biol.* 214:613–617 (1990).

Exercises

1. Ion pairs in proteins involving Arg residues have been observed to be energetically stronger than those involving Lys residues. If this were the case, in what ways might

Arg and Lys residues be used differently in proteins? How could this hypothesis be tested?

ANSWER
D. B. Wigley et al., *Biochem. Biophys. Res. Commun.* 149:927–929 (1987).

2. The hydrogen bond between two water molecules in isolation is about ten times stronger energetically than a van der Waals contact between two xenon atoms, yet water molecules dimerize in the gas phase only about 30% more frequently than xenon atoms. Why?

ANSWER
Section 4.4.

3. The ammonia molecule (NH_3) might be considered analogous to H_2O in its hydrogen-bonding properties. In what way is it different?

ANSWER
D. D. Nelson, Jr., et al., *Science* 238:1670–1674 (1987).

4. Early models of liquid water envisaged mixtures of icelike and disordered collections of H_2O molecules. What are the implications for such models of the observation that liquid water can be supercooled to −41°C?

ANSWER
C. A. Angell, *Ann. Rev. Phys. Chem.* 36:321–346 (1985).

5. Many measurements of hydrophobicity measure the partition of a molecule between aqueous and octanol phases (e.g., C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, John Wiley, New York, 1979). What is the likely significance of the presence of 2.3 M water in the octanol phase? Water-saturated cyclohexane contains only 2.5 mM water, just slightly more than the vapor phase of water. Why is the partitioning of hydrophobic molecules between the aqueous and cyclohexane phases similar to that between the aqueous and vapor phases (Fig. 4.12)?

ANSWER
A. Radzicka and R. Wolfenden, *Biochemistry* 27:1664–1670 (1988).

6. Some measurements of hydrophobicity indicate that the Trp side chain is less hydrophobic than the Phe side chain; others indicate just the opposite (J. L. Fauchere, *Trends Biochem. Sci.* 10:268, 1985). What is the most likely explanation?

ANSWER
R. Wolfenden and A. Radzicka, *Trends Biochem. Sci.* 11:69–70 (1986); *Biochemistry* 27:1664–1670 (1988).

7. Many properties of larger molecules can be calculated from the properties—such as volumes (Table 4.2) and

hydrophilicities and hydrophobicities (Table 4.10)—of the smaller groups from which they are constructed. Under what conditions can this be valid, and when is it not?

ANSWER
R. Wolfenden et al., *J. Amer. Chem. Soc.* 109:463–466 (1987); *Science* 222:1087–1093 (1983). M. A. Roseman, *J. Mol. Biol.* 200:513–522 (1988).

8. Some small proteins are observed to act as antifreeze agents, depressing the freezing point of water in polar fish that live in sea water at −1.9°C. Other proteins promote the crystallization of water. How might proteins have these effects?

ANSWER
R. E. Feeney et al., *Ann. Rev. Biophys. Biophys. Chem.* 15:59–78 (1986); D. S. C. Yang et al., *Nature* 333:232–237 (1988); P. Wolber and G. Warren, *Trends Biochem. Sci.* 14:179–182 (1989).

9. The hydrophobicity of a molecule measured by its ΔG_{tr} between aqueous and nonpolar solvents is increased if it has multiple polar groups that can interact in the nonpolar solvent ("self-solvation"), for example by hydrogen bonding (M. A. Roseman, *J. Mol. Biol.* 200:513–522, 1988). In this case, is it better to measure the hydrophobicities of amino acid side chains using the free amino acids (Y. Nozaki and C. Tanford, *J. Biol. Chem.* 246:2211–2217, 1971), N-acetyl amino acid amides (J. L. Fauchére and V. Pliska, *Eur. J. Med. Chem.* 18:369–375, 1983), analogues of the side chain only (A. Radzicka and R. Wolfenden, *Biochemistry* 27:1664–1670, 1988), or some other model compound?
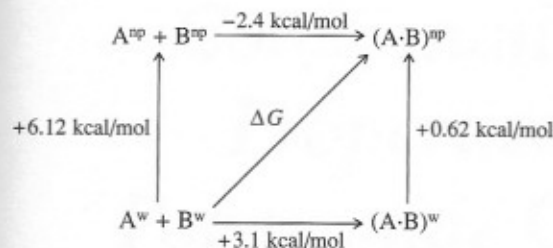
10. The carbonyl and amide —NH— groups of the protein backbone are very hydrophilic and readily form hydrogen bonds with water. However, it has been suggested that when they are hydrogen-bonded to each other, they become essentially hydrophobic and their surface area can be considered nonpolar (C. Chothia, *Nature* 248:338–339, 1974). How could this idea be tested?

ANSWER
M. Roseman, *J. Mol. Biol.* 201:621–623 (1988).

11. Individual model molecules (say, acceptor A and donor B) do not often form hydrogen bonds in aqueous solution (Table 4.9), presumably because of the loss of entropy involved in bringing two molecules together and because the molecules form comparable hydrogen bonds with water. The hydrogen bonds in a folded protein often occur in the nonpolar interior, where hydrogen bonds might be expected to be much more stable than in aqueous solution. The question of the energetics of hydrogen bonding in folded relative to unfolded proteins then becomes that of the energetics of hydrogen bonding in a nonpolar envi-

ronment, designated by the superscript np, relative to the individual polar groups in water, designated by superscript w. This is often analyzed using the free energies of transfer of the bonded and nonbonded molecules from water to nonpolar solvent, and by the extent of their hydrogen bonding in the two solvents. The data for dimerization of N-methylacetamide designated as $A \cdot B$ (Eq. 4.14) in water and in CCl_4 as the nonpolar solvent give the following picture (M. Roseman, *J. Mol. Biol.* 201:621–623, 1988):

$$A^{np} + B^{np} \xrightarrow{-2.4 \text{ kcal/mol}} (A \cdot B)^{np}$$

$$+6.12 \text{ kcal/mol} \qquad \Delta G \qquad +0.62 \text{ kcal/mol}$$

$$A^{w} + B^{w} \xrightarrow{+3.1 \text{ kcal/mol}} (A \cdot B)^{w}$$

From the requirement that free-energy changes around any cycle be zero, the free energy of forming a hydrogen bond in a nonaqueous environment from donors and acceptors in water, ΔG, has the value $+3.72$ kcal/mol; that is, it appears to be unfavorable energetically. Does this indicate that "hydrogen bonding opposes folding" of proteins (K. Dill, *Biochemistry* 29:7123–7155, 1990)?

ANSWER
T. E. Creighton, *Curr. Opinion Struct. Biol.* 1:5–16 (1991).

12. The nature of the genetic code (Fig. 2.4) is such that an amino acid specified by a gene sequence in the normal way has opposite hydropathic characteristics to the amino acid that would be specified by the complementary DNA strand when read in either the 3′ to 5′ or 5′ to 3′ directions (J. E. Blalock and E. M. Smith, *Biochem. Biophys. Res. Commun.* 121:203–207, 1984; J. E. Blalock and K. L. Bost, *Biochem J.* 234:679–683, 1986; J. E. Blalock, *Trends Biotechnol.* 8:140–144, 1990). Such "antisense" peptides are not normally synthesized in biological systems (Chap. 2), but when they are made in the laboratory, they have been claimed to form specific complexes with the normal "sense" peptide (e.g., K. L. Bost et al., *Proc. Natl. Acad. Sci. USA* 82:1372–1375, 1985; R. R. Brentani et al., *Proc. Natl. Acad. Sci. USA* 85:1364–1367, 1988). What are the implications of this observation?

ANSWER
Y. Shai et al., *Biochemistry* 28:8804–8811 (1989); G. Markus et al., *Arch. Biochem. Biophys.* 272:433–439 (1989); E. S. Najem et al., *FEBS Letters* 250:405–410 (1989); J. E. Zull and S. K. Smith, *TIBS* 15:257–261 (1990); E. Roubos, *Trends Biotechnol.* 8:279–281 (1990); R. R. Brentani, *Trends Biotechnol.* 8:281 (1990).