Role of Constraint in Catalysis and High-Affinity Binding by Proteins

Donald G. Vanselow
Glen Waverley, Victoria 3150, Australia

ABSTRACT Using a model for catalysis of a dynamic equilibrium, the role of constraint in catalysis is quantified. The intrinsic rigidity of proteins is shown to be insufficient to constrain the activated complexes of enzymes, irrespective of the mechanism. However, when minimization of the surface excess free energy of water surrounding a protein is considered, model proteins can be designed with regions of sufficient rigidity. Structures can be designed to focus surface tension or hydrophobic attraction as compressive stress. A monomeric structure has a limited ability to concentrate compressive stress and constrain activated complexes. Oligomeric or multidomain proteins, with domains surrounding a rigid core, have unlimited ability to concentrate stress, provided there are at least four domains. Under some circumstances, four is the optimum number, which could explain the frequency of tetrameric enzymes in nature. The minimum compressive stress in oligomers increases with the square of the radius. For tetramers of similar size to natural enzymes, this stress agrees reasonably well with that needed to constrain the activated complex. A similar principle applies to high affinity binding proteins. The models explain the trigonal pyramidal shape of fibroblast growth factor and provide a basis for interpretation of protein crystal structures.

INTRODUCTION

The aim of this work was to elucidate the role of constraint in catalysis and high affinity binding by a soft polymer such as protein and to design model proteins with sufficient constraint to support catalytic or binding functions. The characteristics of the model proteins were then compared with real functional proteins.

Background

An early proposed mechanism for enzymic catalysis was that the enzyme was able to mechanically strain a particular bond of a target molecule and thus provide the activation energy needed for rapid reaction. However Levitt (1974) used molecular modeling techniques to simulate the action of lysozyme and showed that it was mechanically impossible for a model protein structure to apply any significant strain to a model covalent bond because the bonds thought to govern protein structure were much softer than the target covalent bond. In other words, according to the simulation of Levitt (1974), protein was not sufficiently rigid to constrain a highly strained covalent bond.

Hackney (1990), in a thorough review of theories of enzyme function, drew attention to the inability of an enzyme to support a strain mechanism, as follows: ". . . little force is required to deform the enzyme slightly . . . a strained (enzyme-substrate complex) could readily relax to an unstrained conformation by small, low-energy movements." Hackney (1990) went on to describe entropic mechanisms of catalysis including proximity and orientational (orbital steering) effects, noting that these effects depend on restraint or constraint of the atoms involved. However, Hackney (1990) did not discuss the restraining or constraining forces in the same way that he did for the strain mechanism. Perhaps it was mistakenly thought that the restraint or constraint of a low entropy state did not involve forces.

Glennon and Warshel (1998) used computational methods to simulate the action of ribonuclease. By assuming that a number of atoms of the active site and substrate held constrained positions relative to each other, they showed that the simulated interactions between the substrate and those atoms could account fairly well for the experimentally observed catalytic effect. These authors found an approximate value for the necessary constraints but did not show the source of the constraints. They speculated that if they were able to model the protein more completely, the need to apply artificial constraints would disappear. Finally, Glennon and Warshel (1998) seemed to acknowledge the problem of constraint in a catalytic protein when they wrote, "Although the electrostatic model requires much less precise orientation than the orbital steering model, one may wonder how what looks as a rather small structural rearrangement can result in large free energy changes (the energy of flexible systems is not strongly dependent on small structural changes)." They suggested that this difficulty should be the subject of further studies. However, to date there has been no generalized theoretical study of the magnitude of the forces acting on the atoms of the active site, nor of structural features needed to maintain the atoms in constrained positions during catalysis.

Therefore the problem first encountered by Levitt (1974) may still remain. Are protein structures, as they are currently understood, rigid enough to hold the atoms of the active site so that they can impart a large amount of activation energy through a short motion of the reacting species? As shown in the Appendix, the question should be asked of any proposed mechanism of enzyme action and not

Submitted April 4, 2001, and accepted for publication January 28, 2002.

Address reprint requests to Dr. Donald G. Vanselow, 54 Greenways Road, Glen Waverley, Victoria 3150, Australia. Tel.: 61-3-8802-9898; Fax: 61-3-8802-9898; E-mail: dvanselow@hotmail.com.

0006-3495/02/05/2293/11 $2.00

just those that are largely enthalpic, such as bond strain and electrostatic mechanisms.

The model described in the Appendix shows that, irrespective of the mechanism of catalysis, the active part of a catalytic material must have an effective modulus of elasticity not less than the stress or pressure exerted by the activated complex, or intermediate states, on which it does reversible work. If this is not so, the work is distributed throughout the material, and very little is applied to the activation of the reactants.

In the case of an enzyme-catalyzed reaction, the force exerted by the activated complex or intermediate states can be estimated by assuming that the energy of activation is reduced linearly over the distance of interaction of the reactants, i.e., that there is a constant force applied during activation. This assumption may underestimate the maximum force but is adequate for the present purpose. The stress is then calculated by dividing by the assumed cross-sectional area of the activated state.

$$\text{Stress} = \Delta E_a/(A \times \Delta x) \quad (1)$$

in which ΔE_a is the reduction in energy of activation, Δx is the distance over which activation occurs, and A is the cross-sectional area perpendicular to Δx. A maximum value for ΔE_a was estimated by Ji (1974) to be 40 kJ/mol, which is equivalent to 6.6×10^{-20} J/molecule. Glennon and Warshel (1998) estimated ΔE_a for ribonuclease, a particularly effective enzyme, to be of the order of 80 kJ/mol. The value estimated by Ji (1974) is used in the present model. Δx is estimated to be of atomic dimensions with a value of 10^{-10} m, and A is similarly estimated to be 4×10^{-20} m^2. Using these values, the force of activation is 6.6×10^{-10} newton and the stress is calculated to be 1.7×10^{10} Pa or 17 GPa.

The same argument can be applied to any other atoms or orbitals in the active site whose position is critical to maintenance of the activated complex. In the models of Glennon and Warshel (1998), the activated state was reached in two steps, and up to five atoms were simultaneously involved. Consequently, the forces were spread over a larger cross-sectional area, reducing the stress. Even so, the modeled catalytic effect depended on certain atoms being constrained by harmonics of the order of 12 kJ mol^{-1} Å^{-2}. If it is assumed that each of the constraining harmonics was provided by an interaction of atomic dimensions, i.e., by an interaction with an unstrained length of 10^{-10} m and cross-section of 4×10^{-20} m^2, this value for the harmonic is equivalent to an elastic modulus of 5 GPa. Measurements on macroscopic proteinaceous materials show that the elastic modulus of wet structural materials (von Recum, 1986) ranges from 0.05 GPa (heart valve) to 1.0 GPa (tendon). Measurements on flat arrays of wet hexagonally packed intermediate (HPI) protein molecules using the atomic force microscope (Müller et al., 1999) suggest that a force of the order of 10 pN applied over an area of ~2 nm radius deforms the protein by ~0.1 nm. According to Baumeister et al. (1986), the hexagonally packed intermediate layer is 6.5 nm thick. From these measurements, the modulus of elasticity of these protein molecules can be estimated to be of the order of 0.05 GPa and certainly not more than 0.5 GPa. Recently Kharakoz (2000) quoted a value of 5 GPa for the Young's modulus of protein crystals. However, as this value is much higher than the value for bone (1.7 GPa according to von Recum (1986)), there must be some doubt about it. If it were possible for protein to have such a high modulus of elasticity there would be no reason for bone to exist in nature. Therefore, the balance of evidence indicates that the intrinsic rigidity of typical protein material is characterized by a modulus of elasticity of the order of 0.1 GPa or less, whether it is measured in macroscopic specimens or on a molecular scale. This is not surprising because, in both cases, the internal noncovalent interactions allow strain to occur with relatively little energy requirement. These calculations and measurements lead to the conclusion that protein structure appears to be not stiff enough to support strong catalysis.

Reference to a table of elastic moduli (Kaye and Laby, 1995) indicates that only rigid inorganic materials have elastic moduli of sufficient magnitude.

Source of reversible work

The model described in the Appendix shows that it is a requirement of catalysts, that they do reversible work on the reaction they catalyze. In the case of inorganic catalysts, the reduction in the energy of activation is relatively small and the source of reversible work is most probably the energy of binding of the reactants/products to the catalyst. For enzymic catalysis, this source of reversible work is unlikely because the amount of work required is much greater, while the participants in the reaction, including the enzyme, are usually chemically unreactive.

Hypothesis

It is proposed that the only circumstance under which a coiled protein macromolecule could acquire the rigidity to constrain the activated complex or intermediate states of a typical biochemical reaction would be under extreme pressure imposed by structure beyond the active site. The source of this pressure is proposed to be the surface excess free energy arising from the polar component of the cohesive energy of water (Van Oss, 1994). Its value is half the polar component of the cohesive energy of water and is estimated to be 51 mJ m^{-2} (Van Oss, 1994). Sharp et al. (1991) used a value of 72 cal mol^{-1} Å^{-2}, which is equal to 50 mJ m^{-2}. Surface excess free energy around a liquid sphere produces an internal hydrostatic pressure. To the extent that the material in a small protein sphere has solid properties, the

internal forces are better described as stresses. It is postulated that if the geometry of the solid region is well designed, much of the pressure, or internal stress, can be focused on the active site during activation and deactivation. By this means, sufficient rigidity may be generated to constrain the activated complex of a higher order reaction mechanism. (The model described in the Appendix shows why higher order reaction mechanisms require less constraint than lower order reaction mechanisms.)

Furthermore, it is suggested that, by means of a tetrahedral arrangement of four separate protein molecules or structural domains, the stresses arising from the surface excess free energy of water can be further concentrated into the center of the cluster, generating sufficient rigidity to constrain the activated complexes of even the lower order biochemical reactions.

As in the model catalytic device described in the Appendix, it is proposed that these states of high compression are imposed reversibly, transiently, and randomly.

Because the constraining force is constantly applied throughout the movement of the reacting atoms, it can be seen to be the ultimate source of the work of activation.

The constraint generated by the surface excess free energy of water is thus an important part of the function of an enzyme. Whereas the mechanism determines the nature of reactants and products as well as the energy of activation to be overcome, it is the constraint that allows the mechanism to function and ultimately supplies the energy of activation. The relationship between mechanism and constraint is in some ways like the relationship between the gears of a clock and the frame that holds them in place. Both are necessary for function.

DESIGN OF THE STRUCTURE OF A MONOMERIC PROTEIN ABLE TO CONSTRAIN AN ACTIVATED COMPLEX

Protein material is not uniformly soft. Its softness relative to some other materials arises from the lack of a three-dimensional network of strong bonds and the freedom of rotation about those strong bonds that are present. The largest structures not subject to distortion by bond rotation would be tetrahedral groups of directly covalently bound atoms such as the α-carbon of each amino acid, together with the atoms to which it is directly bound. This group would constitute a rigid body of ~0.21 nm radius. Within a solid globular protein, concentration of the compressive stress from the surface is possible because of the finite size of such rigid groups of atoms in the solid lattice. Fig. 1 illustrates this concept. Within an otherwise closely packed compressed lattice the presence of an oversized subunit has the effect of redirecting the forces between nearby lattice members onto itself. If the lattice is designed so that the bound reactants form an oversized subunit, they will be subjected to a large part of the compressive stress in the vicinity. Mechanical

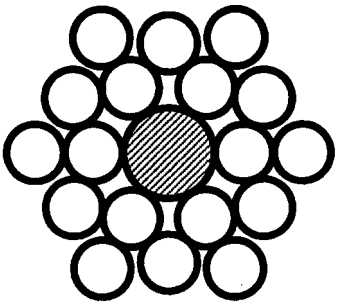

FIGURE 1 Two-dimensional illustration of the concentration of compressive forces onto an oversized sphere in a lattice. Most circumferential contacts are broken, resulting in most of the compressive forces being transmitted through the oversized sphere.

analysis of a sector of a close-packed three-dimensional model of hard spheres shows that as the size of the array of spheres increases, the potential for stress amplification increases as the square of the radius of the array, but the amount of stress bypassing the center also increases with the size of the array. Furthermore, it is well known that the pressure generated inside a spherical object by surface excess free energy is inversely proportional to its radius. The overall result of these mechanical and surface effects is that the central compressive stress in such an array does not change much as subsequent layers of hard spheres are added beyond the first layer. To be more accurate, the model of Sharp et al. (1991) for the magnitude of the surface excess free energy around very small spheres was used, taking care not to confuse surface excess free energy with surface tension. Under the model of Sharp et al. (1991), surface excess free energy and surface tension are no longer numerically equal. The surface generated stress, or pressure, was calculated as dG/dV of the system, in which G is the total free energy and V is the volume of the spherical array. Defining $\Delta G_{S\infty}$ as the surface excess free energy per unit surface area with infinite radius of curvature, the model of Sharp et al. (1991) was written as:

$$\Delta G_{SR}/\Delta G_{S\infty} = 1/(1 + a/R) \quad (2)$$

in which ΔG_{SR} is the surface excess free energy per unit surface area of a sphere of radius R and a is the radius of a water molecule. By differentiating, it can be shown that:

$$dG/dV = \Delta G_{S\infty}(2R + 3a)/(R + a)^2 \quad (3)$$

The result was that a single layer of 12 hard spheres around the central, slightly oversized sphere would experience a surface pressure approaching 0.14 GPa and concentrate the compressive stress by a factor approaching nine to a value of up to 1.3 GPa on the central sphere. Addition of further layers would produce a slight decrease in the central compressive stress. Note that as the central sphere approaches the same size as the surrounding spheres, a number of new contacts begin to form between the surrounding spheres.

This would allow Van der Waals attraction to also contribute a small amount to the compressive stress. For simplicity, this possibility has not been included in the calculations.

In a real solid lattice there is some atomic vibration, so that the time-averaged behavior of the clusters of atoms would be softer than the hard sphere model. A hard sphere model probably overestimates the ability of real structures to concentrate stress. However, this model suggests that there is the potential for a polypeptide molecule in water to catalyze a high order reaction mechanism weakly. Remembering that the hard spheres of the model corresponded to the α-carbon and nearest neighbor atoms of amino acids, and ignoring the possibility that some side-chain structures could play a role, this model suggests that a catalytic polypeptide must contain at least 12 amino acids. Such a cluster would have a radius of at least 0.63 nm.

Higher level structures may assist stress concentration

The above model of a monomeric polypeptide catalyst produced a maximum internal compressive stress well short of the stress required to constrain the activated complexes discussed earlier. Possibly this shortcoming of the model could be partially overcome by considering larger groups of atoms as sufficiently hard bodies in the outer layers of larger arrays. This would be justified by the lower stress present in the outer layers of large arrays. However, a detailed consideration of the surface excess free energy of water shows that the above shortcoming in compressive stress can be totally overcome by forming oligomeric clusters of protein molecules or structural domains.

Distribution of the surface excess free energy of water

It is assumed in computational chemistry that the surface excess free energy of the water/organic interface (Sharp et al., 1991) resides in the water molecules at the interface. According to such a view, the surface excess free energy of two organic surfaces coming together would be released only when the water molecules at the interfaces are displaced by the meeting of the surfaces. However, experimentally, it is found that the hydrophobic attraction between surfaces is still measurable at quite large separations (Israelachvili and Pashley, 1982). Therefore, the surface excess free energy of water must be distributed to a considerable distance from an interface. Applying the Derjaguin approximation (Israelachvili, 1991) to the attractive force data, the free energy of separation of infinite flat surfaces can be shown to be of the form (Van Oss, 1994):

$$\Delta G/A = E_c e^{-D/Do} \tag{4}$$

in which A is the unit area of each flat surface, E_c is the polar component of the cohesive energy of water with a value of 102 mJ m^{-2}, D is the separation between the plates, and D_0 is a constant decay length of $\sim$1 nm.

This energy-separation relationship can be converted into a distribution of surface excess chemical potential as follows. The water near an interface can be considered to have a surface excess chemical potential related to the amount of surface nearby as well as the distance from the surface. A model having these properties, as well as exponential decay, can be expressed mathematically as:

$$\mu = \frac{\mu_0}{2\pi}\int_0^{4\pi} e^{-X_i/X_0}\, d\theta \tag{5}$$

in which μ is the excess chemical potential of an element of water located at distances X_i from incremental elements of interface subtending solid angles of $d\theta$ steradians. μ_0 is a constant equal to the chemical potential at the surface of a flat interface in the absence of other interfaces, and X_0 is the decay length of the exponential function.

By summing the chemical potentials of finite elements of water between two infinite flat plates at various separations, one can find a relationship between separation and free energy that very closely approximates to the energy-separation relationship observed experimentally.

For example, with E_c equal to 102 mJ/m^2 and D_0 assumed to be 1.00 nm in Eq. 4, substitution of $\mu_0 = 6.53 \times 10^7$ J/m^3 and $X_0 = 1.55$ nm into Eq. 5 produces an almost identical relationship between energy and separation over the range 0.5 to 40 nm. In fact, the decay length (D_0) of hydrophobic attraction cannot be measured with such accuracy and may well depend on other factors (Van Oss, 1994), but for the present purpose it is assumed to be known accurately. It must also be pointed out that the chemical potential model only matches experimental results if changes in the amount of water between the two flat plates are taken from, or added to, a pool of water with very small surface excess chemical potential, i.e., bulk water. The model predicts that, in very concentrated solutions, in which all water has a substantial surface excess chemical potential, the hydrophobic attraction would be reduced. In extremely concentrated solutions, the model predicts reversal of hydrophobic attraction. Perhaps this dependence of hydrophobic attraction on solute concentration is reflected in the phenomenon of aggregation on dilution, recently reported by Samal and Geckeler (2001).

The advantage of this model is that it describes well the interfacial forces and energies of the simple geometry for which experimental data exist and can be used to estimate the interfacial effects of more complex geometries, such as edges.

The curvature effect of Sharp et al. (1991) was not included in this model because it is a relatively small correction for the larger structures considered below.

Inspection of the model reveals that the contribution of a unit of surface area to the total free energy of surrounding

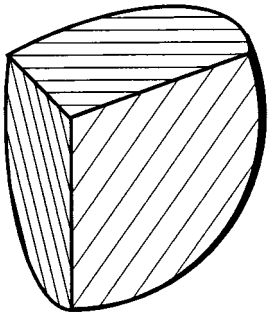

FIGURE 2 Trigonal quartersphere.

water is the same for convex and flat surfaces but is reduced for concave surfaces or surfaces facing one another.

Attraction between very small flat surfaces at very small separations

The above model was used to estimate the edge effect of two flat surfaces 4 nm across and separated by 0.3 to 0.8 nm. The edge effect was found to be only a few percent of the total attraction and was therefore neglected in the following calculations.

Design of the structure of a tetrameric protein able to constrain an activated complex

A segmented sphere has been used in the field of high pressure chemistry to constrain a reaction by concentrating pressure or compressive stress. Such a device was designed in the early 1940s by Baltzar von Platen to concentrate hydrostatic pressure for the synthesis of diamonds (Davies, 1984). The device consisted of six square pyramidal segments of a steel sphere enclosed in a watertight copper skin. The whole assembly was immersed in a pressurized tank of water. The device produced the first synthetic diamonds in 1953. An equivalent tetrahedral arrangement of compressive pistons was conceptualized by P. W. Bridgman and constructed by H. T. Hall in 1957 for the synthesis of diamonds (Davies, 1984).

It is proposed that a similar arrangement of protein domains could concentrate the compressive stress arising from surface tension. Unlike von Platen's device (Davies, 1984), such an arrangement would not have to be watertight, as it would not concentrate external hydrostatic pressure.

The basic architecture proposed can be described as a sphere composed of four identical trigonal quarterspheres as shown in Fig. 2. The active site of each quartersphere is located near the center of the sphere (i.e., near the trigonal apex of each quartersphere), and each site is occupied either by the reactants in some stage of activation or by other components of the solvent environment. Each quartersphere is compressed between its apex and the curved surface but also subject to hydrophobic attraction to adjacent quarterspheres, to produce a resultant compressive stress directed toward the center of the sphere. Apart from the hydrophobic attraction, each quartersphere has no substantial mechanical interaction with the other quarterspheres except near the trigonal apex at the center of the tetrameric sphere. The thickness of the crevice between adjacent faces of quarterspheres is sufficient to admit a layer of water, for example, 0.6 nm thick. As the cross-sectional area of the quartersphere becomes smaller toward the apex, the compressive stress is concentrated and the properties of the protein material become increasingly those of a solid. Near the apex of the quartersphere, mechanical contact occurs with the other quarterspheres and a region of very high rigidity is formed. Because the compressive stresses converge from the four tetrahedral directions, there is no direction in which there is a zero resultant compressive stress. There is, however, a variation in the compressive stress with direction. Geometric calculations show that the compression reaches a maximum along the three twofold symmetry axes and a minimum in directions perpendicular to the six planes of symmetry, the maximum exceeding the minimum by a factor of the square root of two. Table 1 lists the relative minimum compressive stresses obtainable using various numbers of subunits or domains in the constraining cluster. It can be seen that a tetrameric cluster is more effective than other clusters of small numbers of subunits, for a given radius. However, octamers and higher polymers would be still more effective by this geometric criterion. In real catalytic proteins there may well be other criteria determining the most effective number of subunits. For example, some flattening of the curved surfaces may be allowable. These other criteria will be discussed later.

If the cluster is designed such that the central region of high rigidity corresponds to the array of rigid spheres described for a monomeric catalyst, then the scope for focusing compressive stress onto a central catalytic site is unlimited.

The final part of the design is the construction of a catalytic site with suitable geometric and chemical properties, i.e., the provision of a mechanism. Mechanism is beyond the scope of this paper.

Electrostatic effects to stabilize subunit separation in oligomeric clusters

The hydrophobic attraction between surfaces in water grows stronger as the separation decreases (Israelachvili and Pashley, 1982). If this were the only force between faces of adjacent subunits, the arrangement of minimum energy would be asymmetric. Most of the crevices between adjacent subunits would close, whereas one would open up as required by the geometry. The minimum compressive stress would be greatly reduced. To overcome this instability there would have to be a very short-range electrostatic repulsion

TABLE 1 Minimum compressive stresses generated by optimally segmented spheres

Number of segments	Symmetry of force vectors	Minimum compressive stress	Explanation
2	Linear	0	Because the two vectors are colinear there is a plane through the center of the sphere and normal to the vectors, in which the net force is zero
3	Triangular	0	The three vectors are necessarily coplanar. There is an axis through the center of the sphere and normal to the plane, along which the net force is zero
4	Tetrahedral	$0.608P(r_1^2 - r_2^2)/r_2^2$	Minima are along axes normal to planes containing any two vectors
5	Trigonal bipyramidal	$0.578P(r_1^2 - r_2^2)/r_2^2$	Minima are along axes normal to planes containing any three vectors, either all equatorial or two polar and one equatorial. These minima are simultaneously maximized when the angle at the apex of each polar segment face is 100.3 degrees
6	Octahedral	$0.554P(r_1^2 - r_2^2)/r_2^2$	Minima are along force vectors
8	Cubic	$0.707P(r_1^2 - r_2^2)/r_2^2$	Minima are along axes normal to planes containing any four vectors
Infinite		$P(r_1^2 - r_2^2)/r_2^2$	

Note: r_1, outer radius of sphere; r_2, radius of inner core; P, attractive force per unit area of interfacial surface.

attenuating the hydrophobic attraction at low separations, less than, e.g., 0.5 nm. With the net attractive force not increasing below this separation, perhaps even slightly decreasing, the subunits would maintain equal separation averaged over time from all neighboring subunits. Such an electrostatic repulsion could be achieved by having sufficient dipoles or charges at the protein-water interfaces. The increasing exclusion of the high dielectric medium, water, from the interfaces as they approached each other would produce electrostatic repulsion (Israelachvili, 1991).

Model energy-displacement curves

Fig. 3 depicts the proposed Gibbs free energy-displacement relationships in each tetrameric catalytic protein during reaction at equilibrium. The use of straight lines is schematic, but the energy levels depicted are realistic. At the molecular scale the straight lines represent series of minute quantum steps in energy and configuration. Lines ABC represent the free energy of the interfacial water between the faces of adjacent quarterspheres. It is assumed that the curved surface of each quartersphere does not change during catalysis, and hence there is no change in the free energy attributable to the outer surfaces. Lines ADC represent the free energy of activation of the reactants and catalytic site. The total free energy of each reacting system (curve AC) is the sum of the two free energy-displacement relationships and lies within the hatched zone, which is 2 kT wide. (As outlined in the Appendix, restricting total free energy changes to one or two times kT allows frequent random transitions from one state to another.) All free energies are relative to unbound catalyst and reactants. Curve ABC is symmetrical with slope AB equal to slope BC. The slope is ultimately defined by the radius of the protein cluster. The sloping parts of curve ADC are defined by the energy-displacement properties of the reactants and products interacting with the catalytic site. That is, they are determined by the mechanism. In general, curve ADC would not be symmetrical because the slope for the reactants would differ from the slope for the products. However, appropriate design of the mechanism could minimize the asymmetry and hence the residual energy of activation represented by the difference between the highest and lowest points of curve AC.

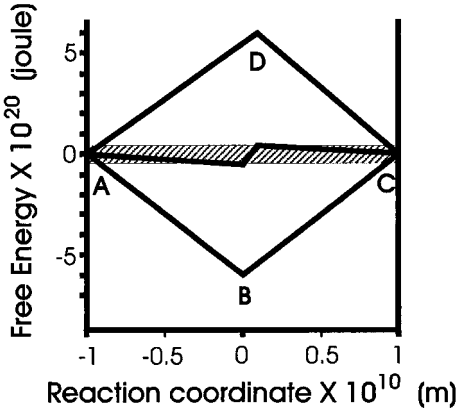

FIGURE 3 Free energy/displacement relationships in a model tetrameric catalytic protein. ABC is the free energy of the interfacial water as subunits approach and recede from each other. ADC is the free energy of activation of the reactants and catalytic site. AC is the total free energy of the assembly. AC lies entirely within a hatched band, which is 2 kT high. All free energies are relative to unbound catalyst and reactants.

It is unlikely that the constraint and/or stressing of molecules other than those to which the site is suited would result in a reaction. Selectivity would be based on the mechanism of interaction between precisely located atoms and orbitals, constrained by the steep energy-displacement curves of the compressed solid protein.

Equilibration of the active site with the bulk solution

The simple model above describes only the structure necessary to constrain the activated complex. It would need

considerable elaboration at atomic detail to also describe a series of states allowing rapid exchange of reactants and products with the bulk solution. Presumably such transient states would include relaxation of the compressive stress. Features that could allow such states to exist might include electrostatic repulsion between subunits at separations just beyond the separation needed for constraint of the activated complex and provision of alternative sites for stress concentration. It is particularly important that the catalytic site is not compressed when there is insufficient matter (reactant, product, or solvent) within it to reverse the compression in a reasonable time. The action of Van der Waals forces in maintaining essentially close packing of atoms in the liquid or solid phases could possibly ensure that the catalytic site is very seldom empty of all atoms and could make "irreversible" compression a rare event. However, most real enzymes appear to gradually "fall" into an inactive state, especially in the absence of the reactant or product (Schmid, 1979).

As the atomic structures of real enzymes become better understood, the nature of the states allowing exchange with the bulk solution should become clearer.

Comparison with real enzymes

The above model catalytic proteins resemble real enzymes in many qualitative aspects. They depend on the presence of at least a film of water around them for catalytic function, as appears to be the case for real enzymes (Marty et al., 1994). By virtue of their (transient) rigidity, they are expected to be extremely selective with respect to reactant and product, as is observed for enzymes.

Quantitatively, a comparison is possible between the compressive stress that could be generated by model proteins of the same size as real enzymes and the compressive stress required to constrain the activated complex of an enzyme. For monomeric enzymes this comparison is not very instructive because the model predicts that compressive stress does not greatly depend on size beyond a very small minimum. Furthermore, in the case of monomeric enzymes with macromolecular substrates, such as the ribonuclease modeled by Glennon and Warshel (1998), it is possible that the substrate has an important structural role and serves as an extra member of a constraining cluster.

However, comparison of oligomeric enzymes with their model counterparts does reveal some similarity. The minimum compressive stress focused by a cluster of quarterspheres is listed in Table 1 as $0.608P(r_1^2 - r_2^2)/r_2^2)$ in which r_1 is the outer radius of the sphere, r_2 is the radius of the central zone onto which the stress is focused, and P is the attractive force per unit area between adjacent faces of quarterspheres. The force (F) per unit area (A) generated by surfaces, for instance, 0.6 nm apart is calculated from the force law equivalent to Eq. 4 (Van Oss, 1994).

$$F/A = (E_c/D_0)e^{-D/Do} = P \quad (6)$$

in which F is the force of attraction, equal to $d\Delta G/dD$, and other symbols are defined as for Eq. 4.

Taking the outer radius of real oligomeric enzymes to be in the range 3 to 7 nm and the radius of the most rigid solid core to be 0.63 nm (as used in the model of a monomeric catalyst), the focused compressive stress on the outside of the solid core was calculated to be from 0.7 to 4 GPa. As was calculated for the model of a monomeric constraining protein, the solid core can concentrate compressive stress up to a further ninefold at the active site. Therefore, the compressive stress at the active site could range from 6 to 36 GPa. This is sufficient to oppose the stress generated by the model activated complex and thus constrain and stabilize it. The possibility that the large size of enzymes is connected with the need for precise alignment of functional groups was suggested by Knowles (1991). In the presence of thermal motion, precision requires rigidity, and the precise (pre)arrangement of atoms and their orbitals, opposed by entropic and nonentropic forces, requires thermodynamic work to be done.

Another comparison with real enzymes is possible by surveying known enzymes for the occurrence of tetrameric structures. Table 2 lists the occurrence of various numbers of subunits in the hydrolase class of enzymes (Schomburg and Salzmann, 1991). Table 3 lists similar data for whole enzymes of any type, compiled from a CD-ROM database. Tetrameric forms clearly predominate over trimeric and higher forms. Monomeric and dimeric forms appear to outnumber all others in literature reports. However, they include many enzymes with more than one domain (Janin and Wodak, 1983). Using the same database from which Table 3 was compiled, it was found that where the domain structure of homodimeric enzymes was reported most subunits contained two domains. These dimeric enzymes were therefore mechanically equivalent to tetramers. Table 1 compares numbers of model subunits according to the minimum compressive stress they can generate. A tetrameric structure produces a greater minimum stress than other oligomeric structures up to octamers. Therefore, to produce a given minimum compressive stress, the size of a tetrameric cluster can be smaller than other small oligomers. By the criterion of efficient use of protein per spherical

TABLE 2 Numbers of subunits in the hydrolase class of enzymes (Schomburg and Salzmann, 1991)

Number of subunits	1	2	3	4	5	6	7	8	Other	Total reports
Number of reports	91	106	6	70	2	13	1	15	27	331
Percentage of reports	27.5	32.0	1.8	21.1	0.6	3.9	0.3	4.5	8.2	100.0

TABLE 3 Numbers of oligomers recorded in a database* of biochemical literature

Keyword (including adjectival and plural forms)	Homo- or heterotrimer†	Homo- or heterotetramer	Homo- or heteropentamer	Homo- or heterohexamer	Homo- or heteroheptamer	Homo- or heterooctamer
Number of reports referring to whole enzymes	7	48	0	3	0	1

*Cambridge Scientific Abstracts, Life Sciences, January 1990 to June 1993. Silver Platter International N.V.
†Most reports of trimers describe parts of functional units and have been excluded from the count.

cluster, tetrameric clusters appear to be better structures than other oligomers up to and including hexamers but less efficient than octamers and higher polymers. However, if it is assumed that the protein is strong enough to withstand some flattening of the curved surfaces, calculations show that it is possible to reduce the volume of material in the clusters without reducing the minimum compressive stress. This is particularly true for the tetramer. If surface flattening is allowed, calculations show that tetramers become more efficient than octamers. To further complicate the comparison between model and real enzymes, it is unclear which part of the catalytic process is critical in the natural selection of enzymes. If it is the events at the active site, then the higher polymers, with more sites per cluster, would be favored. Alternatively, the rate of encounter of reactants with the surface of the cluster could be rate limiting under physiological conditions, rendering the number of active sites per cluster less important.

A further comparison might be provided by x-ray crystallography of enzymes. The model described above is dynamic. To function, its parts must be rapidly moving distances of 0.3 nm or more relative to each other. Any enzyme resembling the model might therefore be expected to lose its activity on crystallization. However, with the inevitable imperfections in crystal structure, some enzyme molecules might be incorporated at sites where sufficient flexibility remained to allow catalytic activity. Unfortunately, such catalytically active enzymes would probably not produce a detectable x-ray diffraction pattern, either because of their free thermal motion or because they are a minor or randomly oriented component of the crystal. For enzymes with dynamic mechanisms, the best one could hope to find, through x-ray crystallography, would be constrained structures that do not differ too greatly from the active forms. In fact, structures of crystallized enzymes, on initial inspection, show no evidence of spherical clusters with centrally located active sites. Perhaps the high concentrations of protein or other solutes at crystallization weakens hydrophobic attraction, as discussed above, resulting in a tendency toward minor rearrangements of subunits to produce packings more compatible with crystal lattices.

It is frequently reported that a crystalline enzyme retains some catalytic activity. As discussed above, this does not necessarily mean that the reported crystal structure resembles a catalytically active state. The reported crystal structure represents only the most ordered, orientated, and immobile components of the crystal. Perhaps the theory developed here provides a basis for the reinterpretation of protein crystal structures, e.g., it provides an explanation for the structure of fibroblast growth factor, described by many as "trigonal pyramidal" (Zhu et al., 1991; Eriksson et al., 1991; Nugent and Iozzo, 2000) or as a "trefoil" (Pellegrini et al., 2000). It suggests that this protein hormone could act by concentrating stress in association with the binding domains of a receptor protein. However, x-ray crystallography of fibroblast growth factor in association with its receptor domains (Pellegrini et al., 2000) does not show, on first inspection, evidence of a stress-concentrating cluster of domains.

Finally, the observed volume of activation of a tetrameric enzyme (Somero et al., 1977) can be compared with the model. The Gibbs free energy of activation (curve ADC of Fig. 3) includes a pressure $\times$ volume term. In the model, the region of convergent stresses is the only part of the structure that changes volume during activation and deactivation of the reactants. At elevated hydrostatic pressure, the pressure $\times$ volume term would become significant and curve ADC would be lowered. The difference between the highest and lowest points of curve AC would therefore be reduced. The volume change of full activation at the site of reaction can be estimated from the figures used above for calculation of the stress generated by activation. The volume change is equal to $-A\Delta x$ and has a value of -4×10^{-30} m^3. In the model, this volume change would be magnified by the region of convergent forces by a factor of up to 9, one lattice unit away from the site of reaction, to a value of approximately -4×10^{-29} m^3. However, only part of this volume change occurs during the rate-limiting step of the catalyzed process (the small step on curve AC). For the left to right reaction in Fig. 3, the measurable volume of activation could be a fraction of the full volume change. The magnitude of the measured volume of activation of the tetrameric enzyme lactate dehydrogenase at low salt concentrations is 1.3×10^{-29} m^3/molecule (Somero et al., 1977), which is consistent with the above. The sign of the volume of activation reported by Somero et al. (1977) was positive, suggesting that the rate limiting process at high pressures occurred during reexpansion of the compressed complex, as shown for the left to right reaction of Fig. 3. For other reactions it may well be negative, depending on the relative slopes of AD and AB.

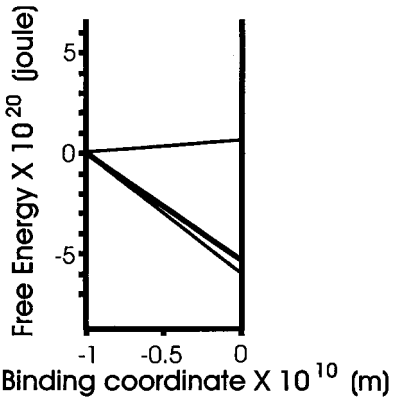

FIGURE 4 Free energy/displacement relationships in a model tetrameric binding protein. The upper thin line is the free energy of the bound molecule occluded in the binding site as a function of the binding coordinate (compression of the binding site), the lower thin line is the free energy of interfacial water between subunits as they move closer together, and the thick line is the total free energy of the system. All free energies are relative to the unbound state of protein and molecule.

Design of a high affinity binding protein

A tetramer constructed like the model constraining protein, but enclosing a molecule in such a way that there is little resistance to compression, would act as a high affinity binding protein. Without the need to impart activation energy (curve ADC of Fig. 3), more of the work done by the interfacial water between adjacent faces of subunits (curve ABC) could be used to offset the free energy of binding from very low concentrations. That is, the total free energy of each bound system would be very much lower than the free energy of unbound molecule and protein. This is illustrated in Fig. 4, where the upper thin line is the free energy of the bound molecule occluded in the binding site, as a function of the binding coordinate (compression of the binding site). The lower thin line is the free energy of interfacial water between subunits as they move closer together, and the thick line is the total free energy of the system.

As discussed in the Appendix, this lowering of free energy of the bound system in the compressed bound state would shift the binding equilibrium strongly in favor of the bound state. Binding specificity would depend on the ligand in the binding site having a smaller volume than would be possible when other constituents of the medium occupied the binding site.

The occurrence of tetrameric or four-domain structures in binding proteins is common. A notable example is the binding fragment of an antibody, which is made up of two protein chains, each comprising two globular domains of ~100 amino acids, as described by Colman et al. (1987). The present "binding by occlusion" model cannot describe binding of intact macromolecular antigens—they can be described instead by the simple interfacial effects of hydrophobic attraction and electrostatic repulsion (work in progress). Suitable antibodies do, however, bind low molecular weight molecules with high affinity and specificity (Janin and Wodak, 1983). Such binding seems difficult to explain except by the present model.

DISCUSSION

It is possible to account for the constraint of the activated complexes of enzymes and high binding affinity of certain proteins using known physical phenomena. The laws of thermodynamics and classical mechanics, combined with the principle of efficient use of protein material, lead to certain structural and dimensional requirements in a catalytic molecular assembly or a high affinity binding protein.

Some of the concepts developed in this work have appeared in slightly different forms in earlier discussions of the functions of enzymes. Some were foreshadowed in the "Conference on the Mechanism of Energy Transduction in Biological Systems" held in 1973 and published as Volume 227 of "Annals of the New York Academy of Science" in 1974. Notably, Ji (1974) proposed a role for compressive stress and developed a mechanical model of catalysis. Unfortunately the focus was on the kinetic process rather than on equilibrium states along the reaction coordinate, and the model Ji described appeared to be capable of concentrating thermal energy without a source of free energy, which is forbidden by thermodynamics. Lumry (1974a) noted the experimental evidence that an enzyme displays properties of a rigid solid. In another contribution, Lumry (1974b) discussed evidence for linkage between protein function and a surrounding shell of water with altered physical properties. Earlier, Lumry (1973) had speculated about the role of surface free energy in generating the internal rigidity in globular proteins.

The common requirement of all catalytic mechanisms is for adequate mechanical constraint. It appears that it is this requirement that determines the overall architecture of enzyme molecules and their assemblies and limits their activity to aqueous environments.

A fuller understanding of protein structure and function should result from using these model catalytic and binding proteins, together with the improved description of the surface excess free energy of water, as a theoretical basis for interpreting x-ray crystal structures.

APPENDIX: A MODEL FOR CATALYSIS

The essential features of a catalytic system can be quantified by consideration of a simple model from a thermodynamic perspective. The system modeled is in dynamic equilibrium, so that all states of the model are in equilibrium with all other states. Thermodynamics provides descriptions of states at equilibrium and does not, in itself, describe kinetic effects. However, by describing a large number of intermediate states along the reversible reaction coordinate, thermodynamics can provide a very full description of the process. If the intermediate states are sufficiently close

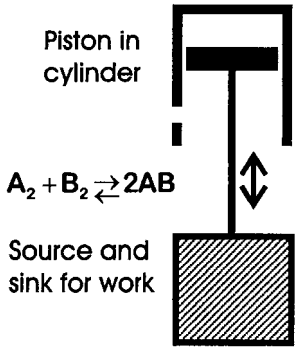

FIGURE 5 Model catalytic device catalyzing equilibration between the gas molecules A_2, B_2, and AB. Up and down motion of the piston is random. Pressure in the cylinder is approximately balanced by the force exerted by the source of work.

together in energy and spatial relationships, the transition from one state to another can be described by simple thermal motion. In the following discussion, and in the main body of the paper, some of the concepts of classical thermodynamics, mechanics, and statistical thermodynamics are applied to a system so small that it is subject to thermal fluctuations in its properties. It should be understood that where a property of the system is discussed, it is the time-averaged or ensemble-averaged value of the property (McQuarrie, 1973).

Description of the model

The model will be described by considering the balance of forces as well as the energetics. The model system is a piston in an isothermal cylinder that has an inlet/outlet port (Fig. 5). The piston is connected to a source of reversible work. The environment of the cylinder is filled with a perfect gas mixture of the molecules A_2, B_2, and AB that can react according to the equation,

$$A_2 + B_2 \rightleftharpoons 2AB \tag{7}$$

The rate of the forward reaction is proportional to the partial pressures of A_2 and B_2. If the force imparted on the piston by the gas in the cylinder is sufficiently close to the force exerted by the source of reversible work at all stages of compression of the gas, then the free energy of all the states of the device will be approximately the same and thermal energy will move the piston randomly through all stages of compression and decompression of the gas mixture. More specifically, provided the free energy of the device, including the source of work, does not vary by an amount much greater than one or two times kT (the product of the Boltzmann constant and the absolute temperature; with a value of 4.14×10^{-21} J at 300 K) over the whole range of states, and the mass of the device is sufficiently small, the contained gas will be randomly subjected to compression over a very short period of time.

The relative probability that the device will be in any given state, i.e., at any particular position, at any time is determined by a Boltzmann expression, $e^{-\Delta G/kT}$ in which ΔG is the difference in free energy content of that state relative to a state with average free energy content. For any macroscopic device, ΔG of many states (positions) will be negative and large relative to kT. Once the device is in such a position there is no realistic possibility of it changing spontaneously to another position during the lifetime of the device. In a macroscopic device this would be described as static friction.

However, for very small devices in which ΔG is a small multiple of $-kT$ there is always a realistic probability of the device moving from one state to another. The effect of ΔG is to determine the relative probabilities of the states. In a large ensemble of similar devices the effect is to determine the relative populations in the states. Thus, any mismatch between the force of the gas on the piston and the force exerted by the source of reversible work becomes manifest as the entropic work of changing the relative populations in an ensemble of similar devices. This concept is important to the discussion of binding in the main body of this paper.

The chemical reaction in the gas phase is now described using kinetic theory. The rate equation of the forward reaction can be written using an Arrhenius expression:

$$r_1 = A\, e^{-\mathrm{Ea/RT}} \times P_{\mathrm{A1}} P_{\mathrm{B1}} \tag{8}$$

in which r_1 is the rate of forward reaction at the environmental pressure, R is the gas constant, T is the absolute temperature, P_{A1} and P_{B1} are the partial pressures of A_2 and B_2, and A and E_a are constants.

The rate of forward reaction in a compressed state, can be written as:

$$r_2 = A\, e^{-\mathrm{Ea/RT}} \times P_{\mathrm{A2}} P_{\mathrm{B2}} \tag{9}$$

in which the subscripts 2 refer to values in the compressed state.

However, the observer having no knowledge of the internal conditions of the device will attribute the increased reaction rate, r_2, to a lower value of E_a, the activation energy, rather than to increased partial pressures. This apparent reduction in the activation energy can be shown to be equal to $RT \ln(P_{\mathrm{A2}}\, P_{\mathrm{B2}}/P_{\mathrm{A1}}\, P_{\mathrm{B1}})$ or $2RT \ln(P_2/P_1)$, in which P_2 and P_1 are total pressures in the cylinder. This is also the classical thermodynamic expression for the work done in isothermal compression of 2 mol of perfect gas molecules (Denbigh, 1966). Thus, the work done reversibly by the catalytic device on each pair of reactant molecules moving from one state of activation to another results in a corresponding reduction in the apparent energy of activation. For higher order reaction mechanisms the reduction in the apparent energy of activation would be greater than for the second order reaction shown above. For example, the apparent activation energy of a third order reaction would be reduced by $3RT \ln(P_2/P_1)$ and so on. Therefore, the dominant reaction mechanism in the catalytic device may be a higher order mechanism than observed in the uncatalysed process.

If the above discussion has shown that it is possible to raise a group of atoms to a higher state of activation by doing reversible work on them, then the laws of thermodynamics require that any path to the higher state of activation requires at least the same amount of work to be done. Otherwise a cycle could be constructed to convert heat to work. It would seem to be a universal requirement of catalysts that they do reversible work on the reaction they catalyze.

The catalytic device can only function if the material of which it is composed is sufficiently rigid. It must be able to constrain the reactants in the activated complex and in the various intermediate states without excessive stretching of its own dimensions. On reflection, the requirement to constrain can be seen to apply to all catalytic systems and mechanisms but most obviously to solid catalysts and enzymes. In the case of enzymes, constraint requires either that the protein has an intrinsic modulus of elasticity greater than the stress needed for constraint or that there is some external source of compressive stress.

The catalytic device described above catalyzes by the proximity effect, which is entirely entropic. It shows that the need for constraint during catalysis applies not only to bond straining (enthalpic) effects but equally to entropic effects. A further illustration of the forces generated by entropy changes is provided by the behavior of rubber, where most of the elastic properties can be attributed to changes in conformational entropy (McQuarrie, 1973).

I thank Robert Gani for pointing out the similarity between these model proteins and the diamond synthesizing apparatus, and Mike Coning for preparing a number of illustrations of quarterspheres.

REFERENCES

Baumeister, W., M. Barth, R. Hegerl, R. Guckenberger, M. Hahn, and W. O. Saxton. 1986. Three-dimensional structure of the regular surface layer (HPI layer) of *Deinococcus radiodurans. J. Mol. Biol.* 187: 241–253.

Colman, P. M., W. G. Laver, J. N. Varghese, A. T. Baker, P. A. Tulloch, G. M. Air, and R. G. Webster. 1987. Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature.* 326: 358–363.

Davies, G. 1984. Diamond. Adam Hilger Ltd, Bristol, UK.

Denbigh, K. 1966. The Principles of Chemical Equilibrium. Cambridge University Press, London.

Eriksson, A. E., L. S. Cousens, L. H. Weaver, and B. W. Matthews. 1991. Three-dimensional structure of human basic fibroblast growth factor. *Proc. Natl. Acad. Sci. U. S. A.* 88:3441–3445.

Glennon, T. M., and A. Warshel. 1998. Energetics of the catalytic reaction of ribonuclease A: a computational study of alternative mechanisms. *J. Am. Chem. Soc.* 120:10234–10247.

Hackney, D. D. 1990. Binding energy and catalysis. *In* The Enzymes. Vol. 19. D. S. Sigman and P. D. Boyer, editors. Academic Press, San Diego, CA. 1–36.

Israelachvili, J., and R. Pashley. 1982. The hydrophobic interaction is long range, decaying exponentially with distance. *Nature.* 300:341–342.

Israelachvili, J. N. 1991. Intermolecular and Surface Forces, 2nd Ed. Academic Press, London.

Janin, J., and S. J. Wodak. 1983. Structural domains in proteins and their role in the dynamics of protein function. *Prog. Biophys. Mol. Biol.* 42:21–78.

Ji, S. 1974. Energy and negentropy in enzymic catalysis. *Ann. N.Y. Acad. Sci.* 227:419–437.

Kaye, G. W. C., and T. H. Laby. 1995. Tables of Physical and Chemical Constants, 16th Ed. Longman, Harlow, UK.

Kharakoz, D. P. 2000. Protein compressibility, dynamics, and pressure. *Biophys. J.* 79:511–525.

Knowles, J. R. 1991. Enzyme catalysis: not different, just better. *Nature.* 350:121–124.

Levitt, M. 1974. On the nature of the binding of hexa-*N*-acetylglucosamine substrate to lysozyme. *In* Peptides, Polypeptides and Proteins. E. R. Blout, F. A. Bovey, M. Goodman, N. Lotan, editors. Wiley, New York. 99–113.

Lumry, R. 1973. Some recent ideas about the nature of the interactions between proteins and liquid water. *J. Food Sci.* 38:744–755.

Lumry, R. 1974a. Conformational mechanisms for free energy transduction in protein systems: old ideas and new facts. *Ann. N.Y. Acad. Sci.* 227:46–73.

Lumry, R. 1974b. Participation of water in protein reactions. *Ann. N.Y. Acad. Sci.* 227:471–485.

Marty, A., D. Combes, and J.-S. Condoret. 1994. Continuous reaction-separation process for enzymatic esterification in supercritical carbon dioxide. *Biotech. Bioeng.* 43:497–504.

McQuarrie, D. A. 1973. Statistical Thermodynamics. Harper and Row, New York.

Müller, D. J., D. Fotiadis, S. Scheuring, S. A. Müller, and A. Engel. 1999. Electrostatically balanced subnanometer imaging of biological specimens by atomic force microscope. *Biophys. J.* 76:1101–1111.

Nugent, M. A., and R. V. Iozzo. 2000. Fibroblast growth factor-2. *Int. J. Biochem. Cell Biol.* 32:115–120.

Pellegrini, L., D. F. Burke, F. von Delft, B. Mulloy, and T. L. Blundell. 2000. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature.* 407:1029–1034.

Samal, S., and K. E. Geckeler. 2001. Unexpected solute aggregation in water on dilution. *Chem. Commun.* 2001:2224–2225.

Schmid, R. D. 1979. Stabilized soluble enzymes. *Adv. Biochem. Eng.* 12:41–118.

Schomburg, D., and M. Salzmann, editors. 1991. Enzyme Handbook 4, Class 3: Hydrolases. Springer-Verlag, Berlin.

Sharp, K. A., A. Nicholls, R. F. Fine, and B. Honig. 1991. Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. *Science.* 252:106–109.

Somero, G. N., M. Neubauer, and P. S. Low. 1977. Neutral salt effects on the velocity and activation volume of the lactate dehydrogenase reaction: evidence for enzyme hydration changes during catalysis. *Arch. Biochem. Biophys.* 181:438–446.

Van Oss, C. J. 1994. Interfacial Forces in Aqueous Media. Marcel Dekker, New York.

von Recum, A. F., editor. 1986. Handbook of Biomaterials Evaluation: Scientific, Technical and Clinical Testing of Implant Materials. Macmillan, New York.

Zhu, X., H. Komiya, A. Chirino, S. Faham, G. M. Fox, T. Arakawa, B. T. Hsu, and D. C. Rees. 1991. Three-dimensional structures of acidic and basic fibroblast growth factors. *Science.* 251:90–93.