The Interactions of Water and Proteins in Cellular Function

J.G. WATTERSON[1]

1 Introduction

Our understanding of the role of water in biology is extremely limited; this most abundant component of living cells is traditionally viewed as structureless, space-filling, background medium in which biochemical events occur. The biochemical and biophysical reactions occurring in aqueous regions of the cell are viewed as occurring in aqueous solution. However, this view disregards the predominantly gelled state of the cell interior. Kempner and Miller (1968) showed by centrifuging intact cells that the fluid aqueous portion of the cytoplasm is devoid of macromolecules. Every biochemist knows that spinning a protein solution under native conditions at 300 000 *g*, corresponding to many hundreds of atmospheres pressure, does not yield a protein concentration as high as 5%. These observations demonstrate that intracellular proteins in their native conformations are able to gel a large proportion of cellular water and prevent its flow.

That extracted enzymes function in solution in vitro is a fact that has traditionally been extrapolated into the concept of the living cell by most biochemists. However, there is no justification for this view, indeed there is a great deal of evidence that speaks against it, as thoroughly reviewed by Welch (1977). In the alternative view, enzymes are associated into complex protein networks permeating the space of the cytoplasm. The existence of such superstructures means that the sequence of chemical steps in metabolic pathways has a corresponding physical state in the cell, wherein enzymes are organized in that sequence in large subcellular aggregates (Srere 1987; Srivastrava and Bernhard 1987). It further means that old concepts of enzyme kinetics and random diffusion of substrate molecules are not appropriate in the reality of the cell (Masters et al. 1987) and implies flexible but highly regulated association-dissociation equilibria and translocation of enzymes themselves (Kaprelyants 1988). This picture of extensive three-dimensional order is strongly supported by the cytoplasmic matrix, as revealed by the work of Porter and his group (1983). But it also poses a problem: such an intricate network of protein scaffolding would hinder, rather than help, intracellular streaming and the free diffusion of metabolites. In other words, the concepts of efficient cytoplasmic movement and structure appear to be mutually exclusive. This question is at present under investi-

[1]Department of Science, Gold Coast University College, Private Mail Bag 50, Gold Coast, Queensland 4217, Australia

P. Jeanteur et al. (eds.), *Progress in Molecular and Subcellular Biology*

gation (Gershon et al. 1985; Luby-Phelps et al. 1988). As argued by Clegg (1984), all the components of the cell, including water, "should be considered as a single system if we are to understand the whole." Thus, the matrix and aqueous elements must work together, as opposed to independently, to produce cytoplasmic movements.

It is now well established that the cytoskeleton is composed of long, linear filaments involved in the production of mechanical force and direction of movement within the cell. Thus, contraction (Huxley 1973), cytoplasmic gel-sol transition and streaming (Taylor and Fechheimer 1987) and axonal transport (Amos and Amos 1985) occur via energetic protein interactions with these filaments. These models are today widely accepted, but in none of them is any significant role assigned to water. If we fill the subcellular space with a simple, liquid solvent, we introduce difficulties for these models from the point of view of the traditional concept of the essential randomness of molecular events. For these processes to function properly, the enzymic machinery that converts chemical energy into work is required to cycle through a series of precise physical steps which cannot tolerate disruptive energetic bombardment from outside. The energy source, usually a phosphate bond of ATP, is equivalent to 10 or 20, or at the very maximum 30, hydrogen bonds and so could hardly be used to tame the violent surroundings of many hundreds of independently acting water molecules, let alone also then be used for the task at hand. A resolution of this problem is offered by a concept of co-operation and participation on the part of the solvent in the function of molecular machinery (Wiggins and MacClement 1987). Examples of physical aspects of such a new concept are the structured interfacial or vicinal water (Drost-Hansen 1985) and the liquid crystalline nature of protein gels (Buxbaum et al. 1987). Recent nuclear magnetic resonance (Lamanna and Cannistraro 1989) and neutron scattering (Giordano et al. 1990) studies on dilute protein solutions have revealed long-range solute-solvent interactions, supporting the idea of solvent involvement.

According to the principles of statistical thermodynamics, the probability that a polymer the size of a protein will spontaneously fold into a unique conformation is vanishingly small if the overall "stabilizing" energy is not very large. But this stabilizing energy is not at all large, being about 10 kcal/mol (Kauzmann 1959; Tanford 1968 and Privalov 1979), which is near that of a phosphate or 10 to 20 hydrogen bonds. In addition, the substitution of a single amino acid can reduce this value by as much as 3 kcal/mol (Goldenberg 1988). Such information has often led to the conclusion that proteins are not very stable, thermodynamically speaking, and raised the question whether they can really adopt a single conformation, when thermodynamics dictates the "kicking and screaming stochastic molecule" of the statistical world (Weber 1975). This problem was addressed by Kauzmann (1959), who proposed the "hydrophobic bond" as the mechanism that gives globular proteins their stability; this bond operating in the protein interior between nonpolar side groups which are repelled by water but attracted to one another. The idea was vigorously criticized, for example by Hildebrand (1968), and during the 1970's the concept was renamed with the terms "hydrophobic effect" and "hydrophobic interaction" (Tanford 1973; Franks 1975). These latter terms imply only a thermo-

dynamic preference shown by nonpolar residues for a hydrocarbon over an aqueous surrounding and do not portray a mechanical picture of protein stability.

The number of energy states readily available to a protein molecule is extremely large, 10^x where x lies between 30 and 70 (Frauenfelder (1983) suggests that x is around 50). Since a protein does not possess special conformational states which differ greatly in energy from one another, the question arises as to whether it exists in all these 10^x states. There is not yet agreement on the answer, as can be seen from directly contradictory opinions expressed in recent reviews (Frauenfelder et al. 1988; Goldenberg 1988). One strategy often used assumes that proteins adopt only those states which are virtually equal in energy to that of the native active state. However, the two principle conformations adopted by hemoglobin, tense and relaxed, are separated by 3 kcal/mol (Perutz 1979), a very sizable fraction of its overall stabilizing energy; so one cannot regard these conformations as energetically close, for otherwise so is the unfolded conformation! On the other hand, one might argue that the energy barriers (activation energies) to most of the 10^x states are too high to be surmounted and consequently only a few states can be occupied. This is a common view of workers interested in protein folding, since this process appears to occur via a strict kinetic sequence of a few conformational intermediates. But such barriers must truly be extraordinarily high, amounting to hundreds of kcal/mol, because the average fluctuation in energy due to random thermal motion in a representative molecule of 25 000 Da exceeds 30 kcal/mol (Cooper 1976). We would then expect that all proteins must have intramolecular bonding arrangements of a very specialized nature indeed to ensure the existence of such high barriers.

The subject of random thermal fluctuations in proteins has been under investigation for more than 20 years. Welch et al. (1982) published an excellent overview of prominent models based on the classical approach, to which must be added the newer computer simulations which are gaining much attention (Karplus and McCammon 1983). The number of three-dimensional protein structures deduced from X-ray crystallography exceeds 200 and the prediction of new ones from sequence data is beginning to show success (Crawford et al. 1987), yet there is no evidence so far of structural channelling mechanisms which could collect and divert the random energy to do useful work at the active site. A different approach proposes that this energy can be stored in rarely occurring, irreversible transitions (Kell 1988).

The idea that molecular fluctuations can perform work is reminiscent of the popular theory of osmosis, in which the random motion of molecules is proposed to provide a force. This mechanism is illustrated in Fig. 1, in which a semi-permeable membrane separates solvent and solution phases. There are fewer solvent molecules (circles) in the solution than in the pure solvent because some have been displaced by solute molecules (stars). Solvent molecules move randomly in all directions, and initially the pressure on both sides of the membrane is equal. Fewer molecules pass from the solution into the solvent than in the opposite direction because there are fewer of them to collide with the membrane of the solution side. This results in a net flow of solvent into the solution, increasing the pressure there above that on the solvent side. The flow against ever-increasing pressure continues until the osmotic

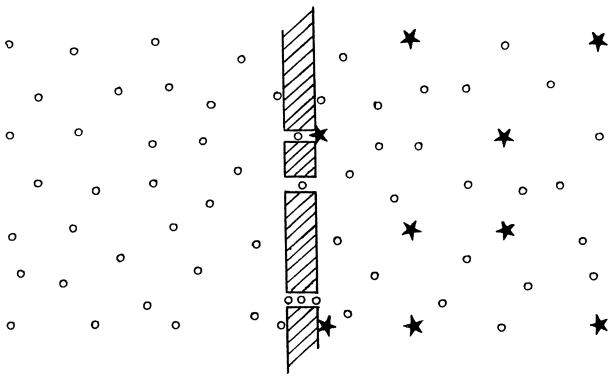

Fig. 1. Pictorial scheme showing a semi-permeable membrane separating solvent on the *left* from solution on the *right*, whereby the solvent molecules (*circles*) can move through the pores but the solutes (*stars*) cannot. Of course the circles should be everywhere in total contact, filling the whole space and not separated as shown here. According to the "molecular theory" of osmosis, the solvent molecules on the *right* are fewer in number, but are diffusing faster than their counterparts on the *left* because of the higher pressure in the solution

pressure is reached, i.e., the pressure difference across the membrane which stops this solvent net flow. The higher pressure in the solution increases the rate at which solvent molecules collide with the membrane, compensating for their fewer number, so the total number of solvent molecules passing from each side is now equal.

Not all physicists and chemists have supported the above theory. In their well-known text, Glasstone and Lewis (1963) give an extensive account of at least five mechanisms of osmosis proposed since Van't Hoff published his law a century ago, and it is of interest to note that none of them is deemed satisfactory by those authors. More recently, other "uphill" flow phenomena have also been detected (Gaeta and Mita 1979). As outlined above, the commonly accepted theory presents solvent flowing against pressure as a result of random molecular motion. However, the displacement of this mass by such a process would break Newton's Second Law of Motion, because the force acting on it during its acceleration is in the wrong direction. One must therefore conclude that the force underlying osmosis does not push the solvent into the solution, but must pull it. This conclusion demands a radical shift in our view of molecular events away from the old concepts of randomness.

2 The Cluster Model of Liquid Structure

2.1 Cluster Size

The idea that tension exists in liquids is not a new one. Its historical development can be found in the excellent review of Hammel and Scholander (1976). In fact, these authors themselves have proposed (Scholander et al. 1965) a modern version

of the theory, in which osmosis is caused by the enhancement of "solvent tension" by solute. However, their theory has been widely criticized (Plumb and Bridgman 1972; Andrews 1976; Hildebrand 1979). Their model is difficult to visualize in molecular terms, because, just as with the statistical theories of liquid structure, it relies also on random collisions. It is proposed that these interactions can give rise to tension. Thus, the basis of their mechanism appears to involve a contradiction, since collisions imply pressure and not tension.

This difficulty does not arise in the cluster model, which is based on the concept of dynamic, co-operative grouping of liquid molecules resulting from intermolecular bonding (Watterson 1982). In this picture tension can be exerted in any one direction as far as the molecules are interconnected in that direction. It is self-evident that tension cannot extend beyond a break in these connections and so, over spatial dimensions larger than those within which unbroken interconnections extend, pressure, and not tension, operates. In other words, at a given instant, tension is felt over a region of space that is as large as a cluster. This picture does not require that every possible bond is formed, as exemplified by regions with ice structure forming within liquid water. It means only that an unbroken interconnection, percolating through the cluster from one side to the other, exists at any instant.

An important aspect of the model is that the making and breaking of bonds are co-operative processes. This rules out the idea that clusters have flickering existence, appearing and disintegrating spontaneously at random. On the contrary, they move about because the change in the bonded state of a group of molecules affects that of its neighbors, and thus the making and breaking processes travel as on-going polymerization and depolymerization reactions through the liquid medium. These processes do not stop and restart independently at random, but are continuous and add together to give a wave motion. We have now the picture of a structure wave completely filling the liquid space, so there is no region where cluster formation is not occurring. The dimensions of a cluster are those than define the wave motion, i.e., the wavelengths. In the case of pure bulk liquid without boundaries this medium is isotropic and so can be idealized as a three-dimensional array of cubic wave units, each defining a cluster of volume u_0. At the corners of the units where the clusters meet are the point nodes in the wave motion. When a foreign solute molecule is introduced into the solvent it disrupts the solvent-solvent interactions which underlie the motion of the structure wave, whatever the particular nature of the solute-solvent interactions. As a result, an extra node forms at the position of the solute, just as an obstacle in any vibrating system generally produces a node. These extra nodes shorten the wavelength, and as a consequence decrease the volume of the wave units to u.

We can now extend this model to explain some fundamental properties of solutions by assuming that it is these wave units or clusters, rather than the single solvent molecules, that act as individual entities. In other words, a cluster can be viewed in some respects like a particle, because the molecules that comprise it at any given moment are bonded together. This picture leads directly to a simple explanation of the colligative properties of solutions (Watterson 1987a,b). For example, the assumption that the clusters evaporate as entire units gives a simple formula

for the reduction of vapor pressure of a solution compared to the pure solvent. Since clusters are smaller in the solution, fewer solvent molecules enter the vapor and so the pressure is correspondingly reduced by the ratio

$$\frac{P}{P_0} = \frac{u}{u_0}, \tag{1}$$

where P and P_0 are the vapor pressures of the solution and pure solvent respectively.

Figure 2 shows how different size clusters interact to produce osmotic equilibrium. In the solution, where the cluster size is reduced, shorter wavelength corresponds to higher tension, just as higher pitch of a vibrating string corresponds to higher tension. This difference has an effect in the region where clusters of different sizes are in contact, in that the material in the smaller clusters exerts a net pull on that in the larger. This is a force operating from within these structures. It is a mechanical force, and if equilibrium is not yet reached, there is transfer of solvent into the solution. In other words, the force results from collective molecular action, and not from independent random collisions.

Since the solvent medium pervades both phases, the wave is propagated in both directions across the membrane. In other words, clusters must also move back and forth across the membrane. They should not be viewed as static structures, as may be falsely interpreted from Fig. 2. Although they change size as they cross this boundary, they must not change energy when the system reaches equilibrium. Then there can be exchange, unit for unit, across the boundary without any further transfer of energy. This argument leads to the conclusion that the clusters obey the Gas Law

$$P_0 u_0 = kT, \tag{2}$$

where k is Boltzmann's constant or the gas constant expressed per molecule (Watterson 1987a). This means that at room temperature and pressure u_0 is about 40 nm^3,

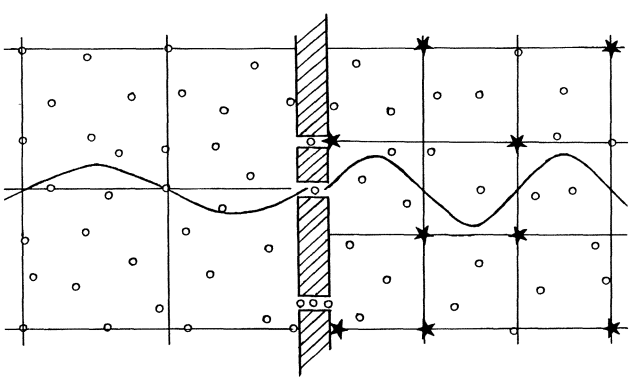

Fig. 2. Schematic representation of Fig. 1, now showing solvent clusters defined by the structure wave. In the solution, the solute molecules (*stars*) are located at nodes on the grid of the stationary wave pattern, and the higher tension in the smaller cluster is shown by the larger amplitude of the shorter wave. At equilibrium, the wave, i.e., clusters, passes smoothly in both directions accross the boundary unit for unit, changing its wavelength and amplitude but not its energy

i.e., a cube with an edge about 3.4 nm long. In the case of water this distance is spanned by 11 molecules roughly, so that a cluster contains 1300 to 1400 molecules and has a molecular mass of about 24 000 Da.

2.2 Cluster Energetics

We are all familiar with the ability of osmotic systems to do work. In terms of the cluster model, this happens while the two phases in contact, pure solvent and solution, move spontaneously towards equilibrium. The larger, energy-rich clusters carry energy into the solution phase until the smaller clusters acquire an equal amount of energy as given by Eq.(2). Clusters can exchange their energies and in doing so provide work, and this idea is expanded here to establish the quantitative relationship between cluster energy and work.

The upper panel in Fig. 3 represents the cycling of a piston in a cylinder just as in the classical Carnot cycle (Moore 1956). The wall of this cylinder contains an opening with a shutter, positioned opposite the piston, which allows the entry and

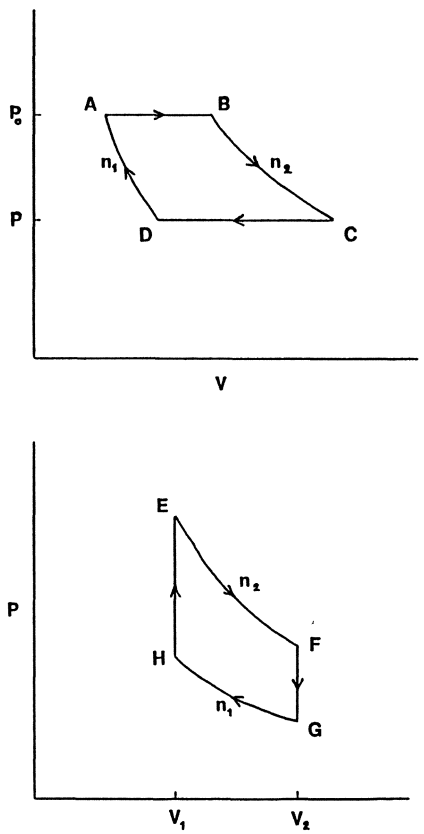

Fig. 3. Work cycles driven by a colligative potential. Heat alone is absorbed by the machine when it contains n_2 molecules of vapor and returned to the environment when it contains n_1 molecules, the difference being converted into work

removal of gases. Two large systems are available to act as source and sink. The first, the source, is a closed volume containing a pure liquid in equilibrium with its vapor at pressure P_0 for temperature T. The second, the sink, contains a dilute ideal solution of a nonvolatile solute in the liquid with the vapor now at the reduced pressure P for the same temperature. In addition, the environment, i.e., any large body also at this temperature, can act as source or sink for the transfer of heat alone. At A, the machine, already containing n_1 molecules of vapor, is placed in contact with the vapor above the pure solvent with the cylinder open. The piston does work reversibly by moving against pressure P_0 while $(n_2–n_1)$ molecules of gas enter from the vapor source. Then at B, the cylinder is closed and the piston now expands down the isotherm, i.e., the familiar PV hyperbolic curve, while absorbing the necessary heat from the environment. From C to D and D to A the steps are reversed with the machine open to the vapor at the lower pressure, P, above the solution (C to D) and then closed (D to A).

The work terms given by the areas under the isobars must be equal since the $(n_2–n_1)$ molecules of gas that entered from the source are all returned to the sink at the same temperature and thus take out with them the energy they brought into the machine. The work done by the machine is, therefore, given by the difference in areas under the isotherms

$$W = (n_2 - n_1)\, kT\, 1n\, (P_0/P). \tag{3}$$

The lower panel in Fig. 3 shows the symmetrical cycle obtained by interchanging the P,V axes. In this case the contents of the cylinder are diluted by a step between two set volumes instead of two set pressures. At the start, E, the machine already contains n_2 molecules of gas and, with the shutter closed, it expands down the isotherm to F while absorbing heat. It is then placed in contact with the vapor of a series of sources in succession and the shutter is opened each time to allow equilibration of the gas in the machine with each source in turn without further increase in volume. The vapor pressure of each source is maintained at a slightly lower value than that of the preceding source by having a corresponding incremental increase in solute concentration. At G there are n_1 molecules left in the machine, and from G to H and H to E the steps are reversed to complete the cycle. Here again, the work done is given by the area under the isotherms.

$$W = (n_2 - n_1)\, kT\, 1n\, (V_2/V_1). \tag{4}$$

These elementary thermodynamic cycles yield the result expected from classical thermodynamics: namely, that the work performed by the machine equals the free energy of dilution, whether the dilution step is an expansion between two given pressures or two given volumes. But, most importantly, the energy used to do this work is the difference in the heat exchanged by the machine while the shutter is closed, and does not originate from the solutions.

The machine can, however, operate in another cycle, shown in Fig. 4, using the steps at constant pressure and volume only, thereby eliminating any heat exchanges. At A the cylinder, already containing vapor at pressure P_0, is placed in contact with vapor above the pure solvent. The shutter is opened and the piston expands revers-

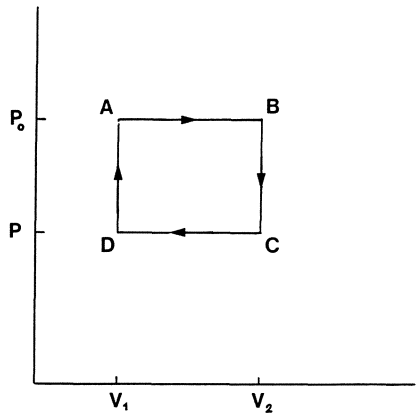

Fig. 4. Rectangular work cycle of the machine operating between the vapor pressure of the pure solvent P_0 and its colligatively reduced pressure P

ibly to B while the required number of molecules, $P_0(V_2–V_1)/kT$, enter the machine. From B to C, the contents are equilibrated with the sources of gradually increasing solute concentration in succession until the pressure falls to P. From C to D and D to A the steps are reversed to complete the cycle performing the work

$$W = (P_0 - P)\,(V_2 - V_1). \tag{5}$$

The work done per molecule is thus

$$w = \frac{P_o - P}{P_o}\,kT, \tag{6}$$

where $(P_0–P)/P_0$ is that fraction of molecules which leaves the machine during the drop in pressure, B to C, and so are not expelled by the return stroke, C to D, when work is done on the machine. All the molecules that enter during the expansion at P_0 are transferred to sources at lower pressures, but because they enter and leave at the same temperature, they take out with them all the energy they brought into the machine. However, in contrast to the cycles in Fig. 3, where work is obtained from heat, in this case the energy must be supplied by a nonthermal source within the solutions. This conclusion is a result of thermodynamic argument and is therefore quite independent of the cluster model.

To my knowledge there is no published interpretation of the rectangular cycle, either in terms of classical thermodynamics or presentday statistical theories of liquids. In terms of the cluster model, however, the reason for the availability of energy is clear. According to Eq.(1), the vapor pressures of the sources are directly proportional to the cluster sizes in their corresponding solutions, and so Eq.(6) becomes

$$w = \frac{u_0 - u}{u_0}\,kT. \tag{7}$$

Thus, the energy converted into work is equivalent to the drop in size of clusters between the first and the last vapor source. This result reinforces the notion of the cluster as a wave unit, i.e., the unit of a vibrating system, because in such a picture

clusters have something of the character of springs. In the pure solvent the spring is strongest, while in the solution it is weaker or less energetic, unit for unit.

In discussing cluster size, bulk liquid was depicted as being composed of a three-dimensional array of cubic clusters produced by an infinite isotropic stationary form of the structure wave. This picture is, of course, idealized, because in the absence of boundaries there are no given directions to guide wave propagation and reflection, and so the motion would not become stationary and form a pattern. One of the simplest examples of a boundary is an infinite two-dimensional surface, such as the air-liquid or solid-liquid interface. This surface would force a stationary two-dimensional planar node to take up position in the wave motion. The formation of such a node need not depend on the chemical nature of the solute surface, because its very existence would induce those molecules in the monolayer adjacent to it to have altered motions compared to those in the bulk. A nodal plane defines a face or side of a cluster, and so the presence of the surface would induce the side-by-side alignment of liquid clusters to spread over it. In this arrangement they form a layer of solvent, one cluster thick, adjacent to the interface.

Over the past two decades, surface hydration effects have been the subject of ever-increasing research interest. Their underlying cause is disputed, but the fact they are observed at clay, organic, and biological surfaces means that they must be a property of water, rather than the solute or the interface. An important effect is the presence of a repulsive "hydration force" exerted outward, normal to the solute surface, the molecular origin of which is an extremely controversial topic (Churaev and Derjaguin 1985). It has been measured in various experimental systems including clays (Norrish 1954; Van Olphen 1954), lipid bilayers (LeNeveu et al. (1976) and mica sheets (Israelachvilli and Adams 1976), and considering the diversity in the chemical nature of the surface investigated, it is astounding how often the distance of 3 to 4 nm is found to be their range of influence (Ninham 1980). The existence of a strong, lateral, inward tension, operating at right angles to the applied force, explains in a simple way the power exhibited by solute systems stacked in layers, such as hydrophilic clays or lamellar micelles, when they swell against imposed pressure (Watterson 1989).

Figure 5 illustrates a simplified view of cluster dynamics. It shows how boundaries may induce harmonic transitions to produce clusters of any size. It also shows how a stationary boundary forming a stationary wave results in a stationary cluster. This does not mean stationary molecules; they move just as in bulk water and maintain the wave motion. Only the monolayer of water molecules adsorbed directly onto the solute surface need be restricted in their movement.

On the other hand, stationary clusters mean no bulk flow. Thus in regions where stationary clusters are aligned together the medium forms a gel. There is no macroscopic flow of solvent because the clusters are fixed in space by being anchored onto the surfaces of fixed solutes. In this picture the packing of solutes, i.e., protein and lipid assemblies, together with stationary clusters is responsible for the gelled state of the cell interior. This is a different view from that offered by the popular theory of gelation. This latter theory, refined by Flory (1953), is based on the infinite degree of cross-linking of a polymeric solute present in the medium. Al-

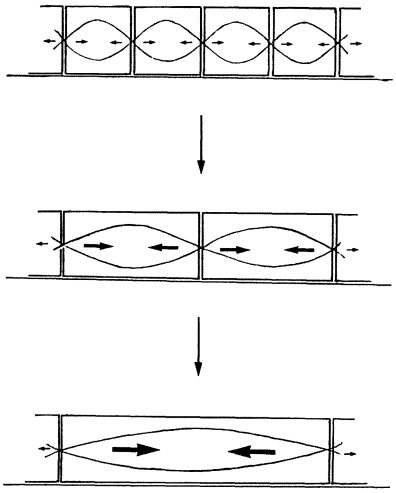

Fig. 5. Clusters become aligned in a row on a flat surface because the presence of the surface guides the direction of the wave motion parallel and normal to it, thus forming a fixed planar node in the wave pattern. Harmonic transitions can then convert many small clusters into a single large one. At the same time the internal lateral tension shown by the *arrows* is exerted over longer distances

though the theory explains why macromolecules stop flowing relative to one another, it has no role for the solvent, which in some systems can be as high as 99% of total content. But it is precisely the cessation of flow of solvent, and not solute, that is so surprising when a gel sets. Thus, to understand the phenomenon so pertinent to the subcellular world, we need a picture in which solvent interactions are seen as playing the central role. This model must explain how small solvent molecules with their individual liquid motions can become interconnected over distances that are long enough to prevent macroscopic flow. Since the popular theory is based solely on the statistical properties of polymeric solutes, it cannot be considered as an explanation of gelation in molecular terms.

3 The Domain Model of Protein Structure

3.1 Domain Size

Why are proteins the size they are? This question has become particularly intriguing since it has been recognized that globular proteins are folded into discrete domains (Richardson 1981; Rossmann and Argos 1981). Table 1 lists 16 proteins composed of a single chain folded into a single domain whose three-dimensional structures are known from X-ray crystallography. Each entry is a member of a large class of similar proteins in terms of tertiary structure. The average length of the 16 listed examples is 192 residues, and most fall within the range 150 to 250.

The citations are listed in chronological order so that there is no attempt at classification. However, prominence should be given to triose phosphate isomerase, as is often done in reviews on protein classification, because of the simplicity and beauty of its structure (Branden 1980; Richardson 1981; Rossmann and Argos 1981). Its folding pattern is referred to as the "TIM barrel", a compact hydrophobic

Table 1. Domain size in proteins of a single domain

Protein	Number of residues	Reference
Myoglobin	153	Kendrew et al. (1960)
Lysozyme	129	Blake et al. (1965)
Ribonuclease	124	Kartha et al. (1967)
Carboxypeptidase A	307	Lipscomb et al. (1970)
Carbonic anhydrase	258	Liljas et al. (1972)
Adenylate kinase	194	Schulz et al. (1974)
Soybean trypsin inhibitor	181	Sweet et al. (1974)
Triose phosphate isomerase	247	Banner et al. (1975)
Concanavalin A	237	Reeke et al. (1975)
Dihydrofolate reductase	162	Matthews et al. (1978)
Glutathione peroxidase	178	Ladenstein et al. (1979)
Ferritin	174	Heustersprente and Crichton (1981)
Aldolase	225	Mavridis et al. (1982)
Superoxide dismutase	151	Tainer et al. (1982)
Retinol binding protein	182	Newcomer et al. (1984)
α-Crystallin	174	Tardieu et al. (1986)

core of eight parallel β-strands surrounded by eight parallel α-helices. However, in this review, it heads the list for an additional reason. This molecule is a sphere with a diameter of 3.5 nm (Banner et al. 1975), and exemplifies a fundamental concept: the correspondence in size of a protein domain and a water cluster. It is thus the domain volumes, not the chain lengths, that I would prefer to list, but this information is not available.

Kendrew et al. (1958) give dimensions for the myoglobin molecule occupying a space of about 36 nm^3. On the other hand, Perutz et al. (1960) give dimensions for hemoglobin indicating about 44 nm^3 per subunit chain. Since structurally speaking these chains are virtually identical, we conclude that the general myoglobin domain occupies about 40 nm^3. From the papers cited in Table 1 we have the following data: along with triose phosphate isomerase, soybean trypsin inhibitor is 3.5 nm in diameter; crystallin is 3.7 nm; glutathione peroxidase is 3.8 nm; and retinol binding protein is 4.0 nm. The dimensions of aldolase and superoxide dismutase give volumes of just 40 nm^3 each, although these two chains differ in length by 74 residues. Similarly, the shapes of both concanavalin A and carboxypeptidase A, although 70 residues different, have outer dimensions of $4 \times 4 \times 4$ nm. So, despite the wide range of sizes judged from chain length, all these single-domain proteins are folded in such a way that they occupy roughly the same volume in the crystals, about 40 nm^3. Exceptions would appear to be lysozyme and ribonuclease; however, these small molecules must occupy volumes larger than expected from their chain lengths, because both have very open structures with wide clefts to accommodate their large macromolecular substrates, polysaccharides and polynucleotides, respectively.

Carboxypeptidase A is the largest single-domain protein so far reported, and one wonders whether it represents the maximum possible size. Matthews et al. (1972) suggested the size of about 16 000 Da (150 residues) is a convenient one for the optimum polar surface area to nonpolar core volume ratio. Edelman (1973) pro-

posed the "domain hypothesis" on the basis of the genetic control of antibody expression. Wetlaufer (1973) took nucleation of folding events as the underlying reason for domains requiring a size of 40 up to 150 residues, but this was criticized later by Reeke et al. (1975), who pointed out that concanavalin A is a single domain of 237 residues, and so these latter workers took a step further to ask if an upper limit exists at all. Lipscombe et al. (1970) give a molecular mass of 34 600 Da for the large carboxypepsidase A, and using a value of 0.74 for protein specific volume (Matthews 1977) gives a volume of 43 nm^3, corresponding to that of a water cluster.

The majority of proteins whose structures are known are composed of two or more separate domains, many of which are smaller than those listed in Table 1. The proteolytic enzymes (the first group in Table 2) are the most studies and discussed X-ray structures. Each is a monomeric molecule which is folded into two separate halves forming a duplex of two small domains. The middle group in Table 2 lists pairs of separate chains. In their functional state these chains are folded into small domains of about 100 residues each, which associate noncovalently to form dimers. Of course, association of native proteins is not restricted to these small domains. Association is a common and basic form of protein behavior, which does not depend on monomer size or type; for instance, the crystallins and ferritins are totally unrelated in their folded structure but form large multimer aggregates for their function. The small domains exemplified in Table 2 need to be associated in pairs to form an active molecule. They are thus named "semi-domains" to indicate that they are only half of a functional entity. The term "half-domain" is already in use in the

Table 2. Semi-domain size in proteins of one or two domains

Protein	Number of residues	Reference
Papain	111 + 101	Drenth et al. (1970)
Elastase	125 + 120	Shotton and Watson (1970)
Chymotrypsin	123 + 122	Birktoft and Blow (1972)
Thermolysin	156 + 160	Colman et al. (1972)
Trypsin	121 + 124	Stroud et al. (1974)
Penicillopepsin	125 + 117	James and Sielecki (1983)
Fab variable domain light, heavy chain	~110, ~110	Edelman (1973)
Fab constant domain light, heavy chain	~110, ~110	Edelman (1973)
Barnase, barstar	110, 89	Hartley and Smeaton (1973)
Prealbumin	127, 127	Blake et al. (1978)
Nerve growth factor	113, 113	Thoenen and Barde (1980)
Glycoprotein hormone α, β subunit	93, 113	Pierce and Parsons (1981)
Cytochrome C	104	Dickerson et al. (1971)
Thioredoxin	108	Eklund et al. (1984)
Calmodulin	148	Babu et al. (1985)

In the first 6 entries the "+" indicates that the protein is a single chain and so reads N-terminal + C-terminal residues. Edelman gives light and heavy chain fragments of Fab portions of immunoglobulins ranging from 102 to 114.

literature to describe a different concept, as referred to below. The last three entries, cytochrome C, thioredoxin, and calmodulin, belong to widely differing protein families. Although they are usually considered to be monomeric, they are invariably associated with another protein in their active forms.

Semi-domain subunits may be identical, as with prealbumin and nerve growth factor, although heterogeneous association is more common. Although their three-dimensional structures are not known, the subunits of the large family of glycoprotein hormones are tightly associated as dimers. In contrast, the even larger family of growth hormones appear to be active as monomers. However, these are larger molecules, being 191 residues long (Li et al. 1973), and therefore fall into the category of single whole domains.

Examples of possible arrangements of domains, some quite complicated, into which a single chain can fold, are given in Table 3. The first entry is a bifunctional enzyme composed of two similar domains which catalyze two successive steps in a metabolic reaction sequence. Alcohol dehydrogenase and actin are each composed of two dissimilar domains. However, the first belongs to an enormous superfamily which includes kinases, phosphorylases and isomerases, while the ubiquitous actin appears to be a unique molecular species. The serum albumins are composed of three very similar domains, while hydroxybenzoate hydroxylase is folded into one domain followed by two semi-domains. At present, little is available on the three-dimensional structures of molecules larger than 70 000 Da, although one example is the widely distributed transferrin, which is made of four roughly equal domains.

Many domains can be further subdivided on the basis of recurring elements of secondary structure, e.g., the semi-domains of the proteolytic enzymes can be further subdivided into half-domains (McLachlan 1979). These recurring folding units are a much-discussed topic (Matthews 1977; Andreeva and Gustchina 1979; Branden 1980; Ptitsyn and Finkelstein 1980; Rossman and Argos 1981; Richardson 1981; Janin and Wodak 1983; Chothia 1984). It is agreed that the repetitions are the outcome of gene duplications, which increased the sizes of earlier versions of functional proteins and so underlie the general process of their evolution. This theory implies a fundamental point: namely, that proteins are the size they are because of

Table 3. Domain sizes in multidomain proteins

Protein	Number of residues	Reference
Phosphoribose-anthranilate isomerase-indol-glycerol-phosphate synthase	255 + 197	Priestle et al. (1987)
Liver alcohol dehydrogenase	165 + 209	Branden et al. (1973)
Actin	150 + 255	Kabsch et al. (1985)
Human serum albumin	192 + 193 + 200	McLachlan and Walker (1977)
p-Hydroxybenzoate hydroxylase	180 + 108 + 103	Wierenga et al. (1979)
Transferrin	175 + 175 + 175 + 175	Baker et al. (1987)

random events at the level of the gene. Clearly, gene duplication is an efficient mechanism for producing a protein sequence containing elements of repeating structure; but this does not make it the cause. As far as I know, there is no evidence that a quarter of the trypsin molecule binds lysine, or that a quarter of the calmodulin molecule binds Ca ions, or that the βαβ-fold or supersecondary structure binds ATP. I think rather that the whole of each of these molecules was always required for their function; and further, that protein size is therefore determined by other forces which constitute the underlying cause of order in the biological world.

Clegg (1979) has pointed out that the absence of macromolecules smaller than 10 000 Da in the cell poses the question of the significance of size in molecular function. In the cluster-domain model, the shapes and sizes of globular proteins are spatially compatible with water clusters (Watterson 1987c, 1988a). Proteins possess dimensions which allow them to pack mutually together and build the large-scale integrated assembly of protein and solvent we know as the cytoplasmic gel. The gel is not infinitely cross-linked. On the contrary, the separate building blocks, water clusters, and protein domains fit together as replaceable parts. Thus the assembly is flexible, but it is also fragile. Water clusters are noncovalently linked internally and so are readily disrupted. As a consequence, the cytoplasmic superstructure can be destroyed by the mildest of ions or detergents, much less cell homogenization.

In the neighborhood of the cell membrane, water clusters are also ordered in the 3-nm-thick hydration layer. The lipid bilayer is itself 3 nm thick and this dimension no doubt gives stability to the mutual packing of layers of lipid and solvent, just as with the protein assemblies. In addition, many membrane-bound proteins have globular portions with enzyme function. Although anchored in the membrane, these enzymic portions are located 3 nm away from the membrane surface via a connecting stalk or rod (MacLennan et al. 1985; Semenza 1986). This mode of construction places the globular domains beyond the hydration layer and so is further demonstration that the dimensions of hydration fit those of protein domains. These observations lead us to a picture of the membrane, not as a single lipid bilayer, but as orderly stacked layers of lipid bilayers, water clusters, and protein domains.

Further out in the extracellular environment, away from the influence of cellular structures, there is no ordering of clusters because in bulk water the chaos of disruptive thermal fluctuations prevails. Proteins bound for this environment are strengthened in their folded conformations by internal covalent disulfide crossbridges. Thus the lack of the stable integrated superstructure is the reason for the presence of these internal crossbridges in certain proteins, not the thermodynamics of special folding mechanisms.

3.2 Domain Energetics

The subcellular medium is mechanically very weak, yet at the same time it is highly energized. These strongly contrasting physical properties are indeed thought-provoking since, without a strong containing framework, they are a recipe for chaotic disaster. In the cluster-domain model, the structural stability of the medium is due

to the packing of spatially compatible units. This is a co-operative process, so that the larger the superstructure, the more stable are its components, thus ensuring the stability of folded protein chains. This idea contrasts with the reductionist view of statistical thermodynamics, because it attributes protein stability to forces operating from above on the large scale, and not to the summation of independent contributions from separate small molecular effects (Watterson 1988b).

The binding of S-peptide, the first 20 N-terminal residues of ribonuclease A, to the rest of the protein illustrates this point. S-peptide is produced by cleaving the intact chain with subtilisin, but binds to the remainder of the molecule with high affinity under native conditions. This complex is active, while the truncated enzyme without S-peptide is inactive (Richards and Vithayathil 1959). The three-dimensional structure of the complex (Wyckoff et al. 1967) reveals only slight and superficial contact, whereby the reside Asp 14 at about the middle of the peptide is described as making a charge-charge interaction, or ionic bond, with the rest of the molecule. However, an ionic bond does not possess the strength of a covalent linkage, and so the complex could not survive energetic fluctuations such as expected during activity. In the cluster model, the binding forces originate from outside the domain. They travel through the water, into and through the protein molecule, so that all sections of the molecule are held in place by a force imposed on them by the large-scale structure of which they are part, and not by small-scale forces which are part of them.

That the cluster size is a fundamental energetic unit is further illustrated by the tight association of semi-domains. Small proteins dimerize often without an indication of what causes this ordering process. In their comparison of immunoglobulin-variable domains, Novotny and Haber (1985) examine the interaction across the interface between heavy and light chain semi-domains. One expects perhaps here, more than in any other case of protein association, to find obvious binding forces because of the finely tuned specificity shown by antibody function. Yet these workers found hydrogen bonding between just one residue on the light chain (Gln 38) and one on the heavy chain (Gln 39) to "extend the hydrogen bonded network across the domain-domain interface and anchor the interface β-sheets in their relative orientation." When one considers that there must be at least 100 hydrogen bonds within each separate chain of 100 residues (Baker and Hubbard 1984) which hold its conformation in place, then one must attribute very special binding powers indeed to this single interaction between the chains, if such small-scale polar interactions are seen as the only agents that can produce force. On the other hand, the cluster model provides a simpler explanation. Isolated molecules of the semi-domain size do not pack well geometrically into the solvent cluster network and consequently it is energetically favorable for them to double their size.

That the force underlying protein-protein attraction operates on a scale as large as the domain is a direct consequence of the cluster-domain model. The co-operative making and breaking of hydrogen bonds, that transmits the structure wave through water, operates equally well throughout protein domains. In Fig. 6, the domain, depicted as a barrel of twisted β-strands, is interpreted as a collection of hydrogen bonds holding the strands together. These rows of flexible bonds open and

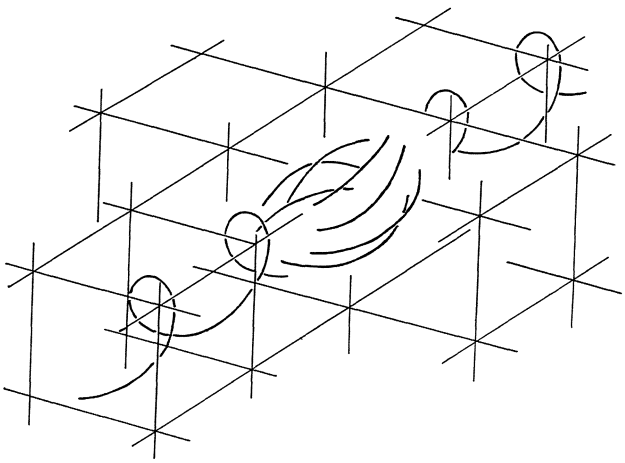

Fig. 6. Diagrammatic illustration of an array of clusters each occupying a volume of approximately 40 nm^3. The *central cube* contains a group of twisted strands representing a globular protein domain. The *spiral* represents the structure wave passing back and forth without interruption through the clusters and the domain

close in zipper-like fashion as the wave passes through the structure (Watterson 1988b). When two neighboring domains, whether belonging to the same molecule or not, adopt a mutual orientation so that the wave can cross their interface, the wavelength can double in the same way as depicted for clusters in Fig. 5. As a result, the internal tension that holds a domain together now operates through the network of bonds in both at once. As with the clusters, these harmonic transitions can produce multidomains of extensive proportions, able to transmit tension on a large scale. In this model, the apparent duplication seen in protein folding anatomy, as in half-domains, semi-domains, whole domains and so on, is intimately connected with the periodic nature of wave motion.

We can compare both approaches, that based on classical statistical theory and that of the cluster-domain model, by taking some well studied examples of protein interactions. Two very different systems, the soybean trypsin inhibitor-trypsin complex and the lysozyme-antilysozyme complex, are considered to be well understood on the basis of their crystal structures. In the first case, Sweet et al. (1974) give the dimensions of the inhibitor molecule as a sphere of 3.5 nm diameter and so in shape and size it resembles trypsin itself. The two spheres share a region of little contact in comparison to their overall sizes, and the authors single out 5 or 6 residues, out of a total of 400 or more, which are responsible for the high affinity interaction. The picture is a similar one in the case of the antigen-antibody complex. Amit et al. (1987) find that 16 residues of lysozyme make contact with 17 of its antibody across a "rather flat surface". However, these contacts are not described as bonds, and the authors list separately 12 hydrogen bonds which must hold together a sphere of 15 000 Da onto the end of an ellipsoid of 50 000 Da in a precise manner.

In the cluster approach, the solvent is an active participant in protein interaction. The attractive force holding the complex together could not operate without the presence of the surrounding medium. The force is transmitted through the whole of the space defined by the dimensions of the fused domains, whether the area of actual protein-protein contact is large or small. Thus, the explanation of the high affinity interactions lies in this large-scale force, and not in a few noncovalent bonds of extraordinary strength.

The dynamic properties mean that the domain-domain or domain-cluster interface has the potential of being a very active region. This follows because when the step doubling the wavelength across two domains occurs, the interface now experiences forces where previously there were none. For the transition to take place, the opposed surfaces need to be compatible, in the sense that the networks of hydrogen bonding existing within each domain separately can be joined, so that overall cooperativity remains ensured and the structure wave can pass smoothly across the interface. Adjacent surfaces can be made compatible by insertion of a small molecule, e.g., substrate, cofactor or metal ion, which makes bonds to both sides, thus bridging the interface. The transition can now take place, and as a result this small molecule will experience a tensile force pulling on it via these interconnecting bonds. It is the same force that pulls water across the membrane in osmosis. According to the amplitude of the ideal stationary wave depicted in Fig. 7, the tension becomes maximum at the center of the fused domain. It is now an easy step to propose that the large oscillations in tension precisely in this region can cause the making and breaking of covalent bonds within the small molecule. In other words, the cluster-domain model provides a mechanism in which chemical events at the enzymic active site are physically coupled to the mechanical events of larger scale in the surroundings. Furthermore, this coupling need not be localized to the region of one molecule. We have seen how clusters and domains can be co-operatively linked by transitions in the structure wave. The linkage of series of enzymic sites in an integrated superstructure opens the way for ordered interplay between the metabolism of the cell and the changing physical states of the cytoplasm.

At present, protein structure, metabolic sequences, and cytoplasmic streaming are regarded as disparate subjects, with the fact that they are all biological being their single common feature. It is, of course, possible that these different features of living systems are unconnected, that each operates independently according to its own principles, and that the cell functions by adopting average situations which result from summation of all the independent processes operating at a given moment. It is, however, unlikely that this mode of function could explain the ordered move-

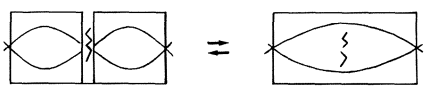

Fig. 7. The diagram represents in an extremely simplified way the functioning of an enzymic active site. When the substrate molecule, represented by the *zig-zag line*, is inserted between two domains, the harmonic transition becomes possible and the oscillating forces now acting in this region split it to form the products

ment exhibited by cells. Subcellular movement takes place as though directed by an underlying co-ordination, implying a unifying principle which links metabolic chemical energy reciprocally with macroscopic mechanical forces. This principle is clearly one of structure existing throughout subcellular space, and of all subcellular components, I think that water is the only one capable of fulfilling this role.

References

Amit AG, Mariuzza RA, Phillips SEV, Puljak RJ (1987) Three-dimensional structure of an antigen-antibody complex at 2.8A° resolution. Science 233:747–753

Amos LA, Amos WB (1987) Cytoplasmic transport in axons. J Cell Sci 87:1–2

Andreeva NS, Gustchina AE (1979) On the supersecondary structure of acid proteases. Biochem Biophys Res Commun 87:32–42

Andrews FC (1976) Colligative properties of simple solutions. Science 194:567–571

Babu YS, Sack JS, Greenhough TJ, Bugg CE, Means AR, Cook WJ (1985) Three-dimensional structure of calmodulin. Nature (Lond) 315:37–40

Baker EN, Hubbard RE (1984) Hydrogen bonding in globular proteins. Prog Biophys Mol Biol 44:97–179

Baker EN, Rumball SV, Anderson BF (1987) Transferrins. Trends Biochem Sci 12:350–353

Banner DW, Bloomer AC, Petsko GA, Phillips DC, Pogson CI, Wilson IA, Corran PH, Furth AJ, Milman JD, Offord RE, Priddle JD, Waley SG (1975) Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 A° resolution using amino acid sequence data. Nature (Lond) 255:609–703

Birktoft JJ, Blow DM (1972) Structure of crystalline α-chymotrypsin. J Mol Bio 68:187–240

Blake CCF, Koenig DF, Mair GA, North ACT, Phillips DC, Sarma VR (1965) Structure of hen egg white lysozyme. Nature (Lond) 206:757–761

Blake CCF, Geisow MJ, Oatley SJ, Rerat B, Rerat C (1978) Structure of prealbumin. J Mol Biol 121:339–356

Branden C (1980) Relation between structure and function of α/β proteins. Quart Rev Biophys 13:317–338

Branden C, Eklund H, Nordstrom B, Boiwe T, Soderlund G, Zeppezauer E, Ohlsson I, Akeson A (1973) Structure of liver alcohol dehydrogenase at 2.9 A° resolution. Proc Natl Acad Sci USA 70:2439–2442

Buxbaum RE, Dennerll T, Weiss S, Heidemann SR (1987) F-actin and microtubule suspensions as indeterminate fluids. Science 235:1511–1514

Chothia C (1984) Principles that determine the structures of proteins. Ann Rev Biochem 53:537–572

Churaev NV, Derjaguin BV (1985) Inclusion of structural forces in the theory of stability of colloids and films. J Coll Interface Sci 103:542–553

Clegg JS (1979) Metabolism and the intracellular environment, the vicinal water network model. In: Drost-Hansen W, Clegg JS (eds) Cell-associated water. Academic Press, NY, pp 363–380

Clegg JS (1984) Properties and metabolism of the aqueous cytoplasm and its boundaries. Am J Physiol 246:R133–R151

Colman PM, Jansonius JN, Matthews BW (1972) The structure of thermolysin. J Mol Biol 70:701–724

Cooper A (1976) Thermodynamic fluctuations in protein molecules. Proc Natl Acad Sci USA 73:2740–2741

Crawford IP, Niermann T, Kirschner K (1987) Prediction of secondary structure by evolutionary comparison. Prot Struct Funct Genet 2:118–129

Dickerson RE, Takano T, Eisenberg D, Kallai DB, Samson L, Cooper A, Margoliash E (1971) Ferricytochrome C. J Biol Chem 246:1511–1521

Drenth J, Jansonius JN, Kockoek R, Sluyterman LAA, Wolthers BG (1970) Cysteine proteinases. Phil Trans Roy Soc Lond B 257:231–236

Drost-Hansen W (1985) Role of vicinal water in cellular evolution. In: Pullman A, Vascilescu V, Packer L (eds) Water and ions in biological systems. Plenum, New York, pp 523–534
Edelman GM (1973) Antibody structure and molecular immunology Science 180:830–840
Eklund H, Cambillau C, Sjoberg BM, Holmgren A, Jornvall H, Hoog JO, Branden C (1984) Conformational and functional similarities between glutaredoxin and thioredoxin. EMBO J 3:1443–1449
Flory PJ (1953) Principles of polymer chemistry. Cornell Univ Press, Ithaca, pp 348–398
Franks F (1975) The hydrophobic interaction. In: Franks F (ed) Water. A comprehensive treatise, vol 4. Plenum Press, New York, pp 1–64
Frauenfelder H (1983) Summary and outlook. In: Porter R, O'Conner M, Wehlan J (eds) Mobility and function in proteins and nucleic acids. Pitman, London pp 329–339 (Ciba Foundation Symposium, vol 93).
Frauenfelder H, Parak F, Young RD (1988) Conformational substates in proteins. Ann Rev Biophys Biophys Chem 17:451–479
Gaeta FS, Mita DG (1979) Thermal diffusion across porous partitions. J Phys Chem 83:2276–2285
Gershon ND, Porter KR, Trus BL (1985) The cytoplasmic matrix. Proc Natl Acad Sci USA 82:5030–5034
Giordano R, Salvato G, Wanderlingh F, Wanderlingh U (1990) Quasielastic and inelastic neutron scattering in macromolecular solutions. Phys Rev A 41:689–696
Glasstone S, Lewis D (1963) Elements of physical chemistry. Macmillan, London pp 248–251
Goldenberg DP (1988) Genetic studies of protein stability and mechanisms of folding. Ann Rev Biophys Biophys Chem 17:481–507
Hammel HT, Scholander PF (1976) Osmosis and tensile solvent. Springer, Berlin Heidelberg New York
Hartley RW, Smeaton JR (1973) On the reaction between the extracellular ribonuclease of *Bacillus amyloliquefaciens*, barnase, and its inhibitor, barstar. J Biol Chem 248:5624–5626
Heustersprente M, Crichton RR (1981) Amino acid sequence of horse spleen apoferritin. FEBS Lett 129:322–330
Hildebrand JH (1968) A criticism of the term "hydrophobic bond." J Phys Chem 72:1841–1842
Hildebrand JH (1979) Forum on osmosis. A criticism of solvent tension in osmosis. Am J Physiol 237:R108–R109
Huxley HE (1973) Muscle contraction and cell motility. Nature (Lond) 243:445–449
Israelachvilli JN, Adams GE (1976) Direct measurements of long range forces between two mica surfaces in aqueous potassium nitrate solutions. Nature (Lond) 262:774–776
James MN, Sielecki AR (1983) Strucure and refinement of penicillopepsin at 1.8A° resolution. J Mol Biol 163:299–361
Janin J, Wodak SJ (1983) Structural domains in proteins and their role in the dynamics of protein function. Prog Biophys Mol Biol 42:21–78
Kabsch W, Mannherz HG, Suck D (1985) Three-dimensional structure of the complex of actin and DNase I at 4.5 A° resolution. EMBO J 4:2113–2118
Kaprelyants AS (1988) Dynamic spatial distribution of proteins in the cell. Trends Biochem Sci 13:43–46
Karplus M, McCammon JA (1983) Dynamics of proteins, elements and function. Ann Rev Biochem 52:263–300
Kartha G, Bello J, Harker D (1967) Tertiary structure of ribonuclease. Nature (Lond) 213:862–865
Kauzmann W (1959) Some factors in the interpretation of protein denaturation. Adv Prot Chem 14:1–63
Kell DB (1988) Coherent properties of energy coupling membrane systems. In: Fröhlich H (ed) Biological coherence and response to external stimuli. Springer, Berlin Heidelberg New York Tokyo, pp 233–241
Kempner ES, Miller JH (1968) The molecular biology of *Euglena gracilis* IV. Cellular stratification by centrifugation. Exp Cell Res 51:141–149
Kendrew JC, Bodo G, Dintizis HM, Parrish RG, Wyckoff H, Phillips DC (1958) A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. Nature (Lond) 181:662–666
Kendrew JC, Dickerson RE, Standberg BE, Hart RG, Davies DR, Phillips DC, Shore VC (1960) Structure of myoglobin. Nature (Lond) 185:422–427

Ladenstein R, Epp O, Bartels K, Jones A, Huber R, Wendel A (1979) Structure analysis and molecular model of the selenoenzyme glutathione peroxidase at 2.8 A° resolution. J Mol Biol 134:199–218

Lamanna R, Cannistraro S (1989) Water proton self-diffusion and hydrogen bonding in aqueous human albumin solutions. Chem Phys Lett 164:653–656

LeNeveu DM, Rand RP, Parsegian VA (1976) Measurement of forces between lecithin bilayers. Nature (Lond) 259:601–603

Li CH, Gordon D, Knorr J (1973) The primary structure of sheep pituitary growth hormone. Arch Biochem Biophys 156:493–508

Liljas A, Kannan KK, Bergsten PC, Waara I, Fridborg K, Stanberg B, Carlbom U, Jarup L, Lovegren S, Petel M (1972) Crystal structure of huamn carbonic anhydrase. Nature New Biol 235:131–137

Lipscomb WN, Reeke GN, Hartsuck JA, Ouiocho FA, Bethge PH (1970) The structure of carboxypeptidase A. Phil Trans Roy Soc Lond B 257:177–214

Luby-Phelps K, Lenni F, Taylor DL (1988) The submicroscopic properties of cytoplasm as a determinant of cellular function. Ann Rev Biophys Biophys Chem 17:369–396

MacLennan DH, Brandl CJ, Korczak B, Green NM (1985) Amino acid sequence of Ca + Mg-dependent ATPase from rabbit muscle sarcoplasmic reticulum deduced from its complementary DNA sequence. Nature (Lond) 316:696–700

Masters CJ, Reid S, Don M (1987) Glycolysis. A new concept in an old pathway. Mol Cell Biochem 76:3–14

Matthews BW (1977) X-ray structure of proteins. In: Neurath H, Hill R, Boeder C (eds) The proteins, vol 3. Academic Press, New York, pp 403–590

Matthews BW, Jansonius JN, Cloman PM, Schuenborn BP, Dupourque D (1972) Three-dimensional structure of thermolysin. Nature New Biol 238:37–41

Matthews BW, Alden RA, Bolin JT, Filman DJ, Freer ST, Hamlin R, Hol WGJ, Kisliuk RL, Pastore EJ, Plante LT, Xuong N, Kraut J (1978) Dihydrofolate reductase from *Lactobacillus casei*. J Biol Chem 253:6946–6954

Mavridis IM, Hatada MH, Tulinsky A, Lebioda L (1982) Structure of 2-keto-3-deoxy-6-phosphogluconate aldolase at 2.8 A° resolution. J Mol Biol 162:419–444

McLachlan AD (1979) Gene duplication in the structural evolution of chymotrypsin. J Mol Biol 128:49–79

McLachlan AD, Walker JE (1977) Evolution of serum albumin. J Mol Biol 112:543–558

Moore WJ (1956) Physical chemistry, 2nd edn. Longmans, London, pp 48–51

Newcomer MÈ, Jones TA, Aquist J, Sundelin J, Eriksson U, Rask L, Peterson PA (1984) The three-dimensional structure of retinol binding protein. EMBO J 3:1451–1454

Ninham BW (1980) Long-range versus short-range forces. J Phys Chem 84:1423–1430

Norrish K (1954) The swelling of montmorillonite. Disc Faraday Soc 18:120–134

Novotny J, Haber E (1985) Structural invariants of antigen binding. Proc Natl Acad Sci USA 82:4592–4596

Perutz MF (1979) Regulation of oxygen affinity of hemoglobin. Ann Rev Biochem 48:327–386

Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North ACT (1960) Structure of hemoglobin. Nature (Lond) 185:416–422

Pierce JG, Parsons T (1981) Glycoproptein hormones, structure and function. Ann Rev Biochem 50:465–495

Plumb RC, Bridgeman WB (1972) Ascent of sap in trees. Science 176:1129–1131

Porter KR, Berkerle M, McNiven A (1983) The cytoplasmic matrix. Mod Cell Biol 2:259–273

Priestle JD, Grutter MG, White JL, Vincent MG, Kania M, Wilson E, Jardetzky TS, Kirschner K, Jansonuis JN (1987) Three-dimensional structure of the bifunctional enzyme N-5-phosphoribosyl anthranilate isomerase — indole-3-glycerol phosphate synthase from *E. coli*. Proc Natl Acad Sci USA 84:5690–5694

Privalov PL (1979) Stability of proteins. Adv Prot Chem 33:167–241

Ptitsyn OB, Finkelstein AV (1980) Similarities of protein topologies. Ouart Rev Biophys 13:339–386

Reeke GN, Becker JW, Edelman GM (1975) The covalent and three-dimensional structure of concanavalin A. J Biol Chem 250:1525–1547

Richards FM, Vithayathil BJ (1959) The preparation of subtilisin-modified ribonuclease and the separation of protein and peptide components. J Biol Chem 234:1459–1465

Richardson JS (1981) The anatomy and taxonomy of protein structure. Adv Prot Chem 34:167–339
Rossmann MG, Argos P (1981) Protein folding. Ann Rev Biochem 50:497–532
Scholander PF, Hammel HT, Bardstreet ED, Hemmingsen EA (1965) Sap pressure in vascular plants. Science 148:339–346
Schulz GE, Elizinga M, Marx F, Schirmer RH (1974) Three-dimensional structure of adenylate kinase. Nature (Lond) 250:120–124
Semenza G (1986) Anchoring and biosynthesis of stalked brush border membrane proteins. Ann Rev Cell Biol 2:255–313
Shotton DM, Watson HC (1970) The three-dimensional structure of crystalline porcine pancreatic elastase. Phil Trans Roy Soc Lond B 257:111–118
Srere PA (1987) Complexes of sequential metabolic enzymes. Ann Rev Biochem 56:89–124
Srivastrava DK, Bernhard SA (1987) Biophysical chemistry of reaction sequences. Ann Rev Biophys Biophys Chem 16:175–204
Stroud RM, Kay LM, Dickerson RE (1974) The structure of bovine trypsin. J Mol Biol 83:185–208
Sweet RM, Wright HT, Janim J, Chothia CH, Blow DM (1974) Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor at 2.6 A° resolution. Biochemistry 13:4212–4224
Tainer JA, Getzoff ED, Beem KM, Richardson JS, Richardson DC (1982) Determination and analysis of the 2 A° structure of copper, zinc superoxide dismutase. J Mol Biol 160:181–217
Tanford C (1968) Protein denaturation. Adv Prot Chem 23:121–282
Tanford C (1973) The hydrophobic interaction. Wiley, New York
Tardieu A, Laporte D, Licinio P, Krop B, Delaye M (1986) Calf lens α-crystallin quaternary structure. J Mol Biol 192:711–724
Taylor DL, Fechheimer M (1982) Cytoplasmic structure and contractility, the solation-contraction coupling hypothesis. Phil Trans Roy Soc Lond B 299:185–197
Thoenen H, Barde YA (1980) Physiology of nerve growth factor. Physiol Rev 60:1284–1335
Van Olphen H (1954) Interlayer forces in bentonite. Clays Clay Miner 2:418–438
Watterson JG (1982) Model for a co-operative structure wave. In: Franks F, Mathias SF (eds) Biophysics of water. Wiley, New York, pp 144–147
Watterson JG (1987a) Does solvent structure underlie osmotic mechanisms? Phys Chem Liq 16:313–316
Watterson JG (1987b) Solvent cluster size and colligative properties. Phys Chem Liq 16:317–320
Watterson JG (1987c) A role for water in cell structure. Biochem J 248:615–617
Watterson JG (1988a) The role of water in cell architecture. Mol Cell Biochem 79:101–105
Watterson JG (1988b) A model linking water and protein structures. Bio Systems 22:51–54
Watterson JG (1989) Wave model of liquid structure in clay hydration. Clays Clay Miner 37:285–286
Weber G (1975) Energetics of ligand binding to proteins. Adv Prot Chem 29:1–83
Welch GR (1977) On the role of organized multienzyme systems in cellular metabolism. Prog Biophys Mol Biol 32:103–191
Welch GR, Somogui B, Damjanovich S (1982) The role of protein fluctuations in enzyme action. Prog Biophys Mol Biol 39:109–146
Wetlaufer DB (1973) Nucleation, rapid folding, and globular intrachain regions in proteins. Proc Natl Acad Sci USA 70:697–700
Wierenga RK, de Jong RJ, Kalk KH, Hol WGJ, Drenth J (1979) Crystal structure of p-hydroxybenzoate hydroxylase. J Mol Biol 131:55–73
Wiggins PM, MacClement BAE (1987) Two states of water found in hydrophobic clefts. Int Rev Cytol 108:249–303
Wyckoff HW, Hardman KD, Allewell NM, Inagami T, Johnson LN, Richards FM (1967) The structure of ribonuclease-S at 3.5 A° resolution. J Biol Chem 242:3984–3988