On First Looking Into Nature's Chemistry

Part II The Role of Large Molecules, especially Proteins

1 Introduction

My first article, Part I, described the movements of small units, substrates, co-enzymes, ions, *etc.*, within metabolism and their control. However, it must be obvious that all these molecules function through their differential effects, usually due to binding, at protein surfaces both during chemical transformation and in the operation of control. It is then necessary to see how the presence (binding) of small molecules and ions to catalysts (see allosteric and rhesteric effects) and to other structural and control devices (all other proteins) causes the proteins to adjust their functions. The most usual mechanism is by 'conformational' change. The structural dynamics of the large molecules in the cell then become the modes through which the flows of the small molecules intercommunicate. These dynamics are also of great importance in protein–protein recognition and interaction. The nature of the structure and dynamics of proteins is the subject matter of this my second article.

Immediately it is essential to clarify the meaning of 'conformational' change. A protein is not a rigid body so that within even the most structured proteins, of one ground-state conformation, there are fluctuations amongst many vibrational and rotational energy states. At the opposite extreme the most mobile proteins are random coils. The range of proteins is such that at the one extreme a conformation change can mean a change similar to a phase change or a change from one molecular structure to another, while at the other extreme proteins change, with conditions, the occupation of states within a huge ensemble of states. The simplest model is therefore a two-state description of change, so-called relaxed-to-tense ($R \rightleftarrows T$), while the most complicated can only be represented by a detailed analysis of change of a partition function, and structures alone are inadequate in the explanation of properties (Figure 1).[1]

We must make it clear that the meaning of mobility, in the sense in which we are using it, includes all the rate processes within proteins. We will have to distinguish eventually between the extremes of very fast equilibrations over many conformers of similar energy, when a protein cannot be said to have a structure, and of slower changes in conformation between a minimum of two forms. The latter may well be quite slow and it is the type of motion which is admirably suited to the slow balanced control of the internal cytoplasmic steady state. Such conformation changes may take several tenths of a second and they would be of little use in trigger steps. We then expect that the control of isolated enzymes or of enzyme particles by co-enzymes in the cytoplasm will not be the same as the control within the trigger devices which must turn over very rapidly, *e.g.* nerve,

[1] R. J. P. Williams, *Biol. Rev.*, 1979, **54**, 389.

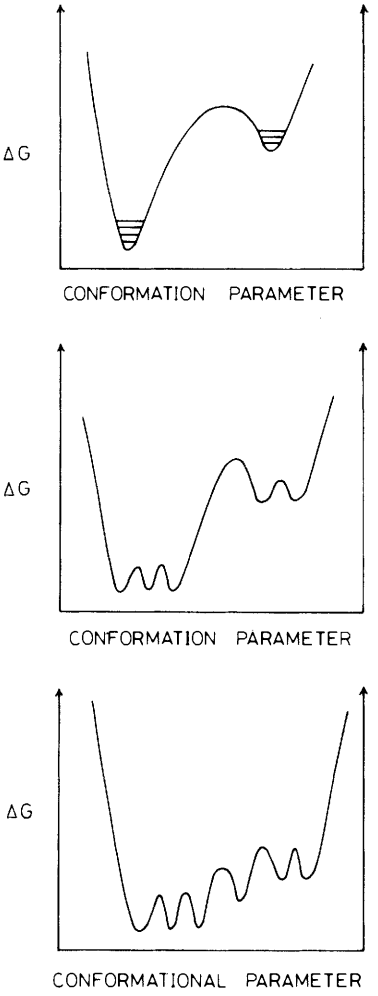


Figure 1 *Three possible free-energy diagrams for proteins, illustrating dependence upon atomic co-ordinates (conformation). At the top there are two conformers, one of which is highly energized, that have many vibrational states: in the middle a more complex set of conformers with more than one ground and one excited state is shown; at the bottom is a diagram for the case where the ground state and all excited states are populated at all reasonable temperatures*
(Reproduced with permission from *Biol. Rev.*, 1979, **54**, 389)

muscle, hormone, and transmitter activation by ions and small covalent molecules such as *c*-AMP. We are again almost forced to look for two sets of proteins with very different types of 'phase' equilibria. In the first case the changes will resemble an allotropic switch in a crystal, *e.g.* from rhombic to tetragonal sulphur,

while in the second case we shall observe a switch of balance within a dynamic equilibrium of many species as is seen, to use a loose parallel, in liquid sulphur, dynamic allotropy. It is not only the rates of such changes which are different, and we shall consider later the energetics of these controls.

Even when the change in the protein is apparently just from one conformation to one other it is highly improbable that the change could be totally co-operative, *i.e.* a single-step process. The probable nature of such a readjustment is that an initial change at one point in the structure spreads through small steps into the protein to other distant regions. This implies that the internal groups of a protein have mobility. Again, the binding of a substrate to a protein and of one protein to another may also require many steps at the active site or surface if they are not to have very high activation energies. Thus, even if it were possible to represent the initial and final states by two single states to a good approximation so that the equilibrium could be described by two structures ($R \rightleftarrows T$) this analysis would not leave any possible description of the mechanism of change. It is essential to have considerable residual motion in all states of proteins. The relationship we need to study is that of *mobility within structures to function.*

2 Structure and Mobility

A. The Mobility of the Small Molecules.—In Part I we have seen that the small molecules and ions are freely diffusing in aqueous or membrane phases. The very construction of most of these molecules shows that they also have considerable internal mobility with some barriers due to ring structure or preferred conformation. Many of the co-enzymes, formed from the units nucleotide base:sugar:carrier group, *e.g.* adenine:ribose:triphosphate (ATP), are constrained to an anti-conformation about the base–sugar bond, and this conformation is also observed in enzyme bound states, but the carrier group, here ribose triphosphate, is much more mobile (see the formulae in Table 4 of Part I), and this mobility is probably retained in enzymes. The mobile carrier group becomes the carrier arm in organized particles (see Table 4 of Part I). Of course, this very mobility of the carrier group has to have a parallel mobility in the enzyme active site, a general requirement for good catalysis. Undoubtedly, the mobility also helps the on/off binding steps of the co-enzymes and substrates with the proteins.

Apart from small organic molecules we are also concerned with the properties of the ions K^+, Na^+, Ca^{2+}, and Mg^{2+}. Of these ions the first three have sufficiently fast water-exchange reactions (involving structure changes) to be able to bind strongly, $K \simeq 10^7$ for calcium, and yet have fast binding-exchange reactions. For magnesium ions the slower water exchange means that its binding constant must not exceed 10^4 if its off-reaction is to be fast. Even so, all these ion reaction rates are fast. Moreover, we have shown[2] that there can be very fast rearrangement of ligands on the surface of the ions Ca^{2+}, Na^+, and K^+. Their complexes are fluctional, rapidly interconverting (on the surface of the ions) from one geometry

[2] B. M. Alsaadi, F. J. C. Rossotti, and R. J. P. Williams, *J. Chem. Soc., Dalton Trans.*, 1980, 597.

to another. It is not likely that this is true for the magnesium ion as the steric constraints in its octahedral complexes are very considerable. Magnesium will be able to hold a particular geometric arrangement. Once again we shall look for the corresponding properties in the proteins which will allow biological systems to take advantage of these fundamental kinetic properties of the ions. (Note in passing that internal mobility in some organic trigger molecules is smaller, *e.g.* *c*-AMP, and is certainly less than in ATP.)

B. Mobility in Proteins.—All proteins consist of linear chains of α-amino-acids. Some of them appear to have little further structure even *in vivo*, *e.g.* chromagranin,[1] but such isotropic random-coil proteins can be of little use in machinery which must transmit mechanical force anisotropically from one place to another. We must look, therefore, at structured, folded proteins and try to locate their different types of moving parts which parallel conventional parts of machines. Subsequently, organization of proteins will lead to a larger-scale machine, the cell, and organization of cells will lead to even larger machines, the organism. In general, a machine has relatively rigid parts which move very slightly but *transmit* force, *e.g.* a simple rod, and more flexible parts which move easily and are made of more elastic materials (see Table 1). The second type of machine part stores more energy during action, is deformable, and therefore can be recognized

Table 1 *Types of material and their properties*

Liquids	*Pliant solids*	*Tensile solids*
Flow	Motile	Rigid
No shape	Change shape	'Fixed shape'
Viscous	Visco-elastic	Linear elastic
Lubricate	Energy store (ΔS)	Transfer force
ΔS large	ΔS intermediate	ΔS small

readily in its different states, *e.g.* a spring. What are the corresponding parts of proteins?

While stressing the role of proteins in these machines, other polymers are not excluded but are less generally used. For example, both DNA and RNA must also have both structural and dynamic features. These polymers are used in the synthesis of polynucleotides and proteins and we shall only mention them toward the very end of the article. Two other types of polymer are very important in biological systems. First, the polysaccharides have a comparable set of roles in structural dynamics with the proteins, especially in membranes and external tissues, walls, and connective tissue, but do not handle cell controls though they assist in recognition of the environment. The second polymer class is the series of hydrocarbons $-(CH_2)_n-$ which are found in fats, lipids. Again, they have a role in structural dynamics within membranes but usually do not handle cell-control elements directly. It is the special ability of *folded* proteins to recognize specifically

small molecules and ions, either transforming them or being structurally adjusted by them, that makes the proteins so overwhelmingly important. The required combination of reactive protein side chains and controlled tertiary structure is absent in the other polymers – DNA, RNA, lipids, and polysaccharides. Many of the comments we shall make about dynamics therefore apply to all types of polymer, but the specific properties in which we are interested rest in the proteins alone.

Proteins, as we have already stated, are linear chains of α-amino-acids connected by peptide bonds. They usually exceed one hundred units in length, most shorter chains being control (hormonal) peptides. There are some twenty amino-acids, divided into four classes in Table 2, which can be incorporated into the

Table 2 *Classes of amino-acids*

Charged	*Intermediate*	*Hydrophobic*	*'Structural'*
Lysine, arginine, glutamate, aspartate	Serine, histidine, glutamine, asparagine	Leucine, isoleucine, valine, alanine, threonine, cysteine, phenylalanine, tyrosine, tryptophan, methionine	Glycine, proline

Note In many proteins modified amino-acids appear through, for example, methylation or hydroxylation

linear sequences, and the vast variety of proteins arises from the huge number of combinations which are possible. It is the exact sequence which decides the fold of the protein. Some proteins fold into tight globular structures, some are multi-lobed, some are very loose, and some form long strands. It is helpful to distinguish the *secondary* structures, helices, sheets, and perhaps β-bends, of amino-acids which arise from hydrogen bonds between the peptide links, *i.e.* the –CO–NH– units, and are independent of the nature of the amino-acid side chains, from *tertiary* structures which arise from further folding due to interactions between side chains and are specifically dependent upon the side chains. Figure 2 gives a schematic representation of this distinction.[3] The secondary structures, helices and sheets, can be quite tight strands or platforms but cannot regulate the juxtaposition in space of amino-acid side chains to any degree of subtlety.

A final feature of the amino-acids in proteins is their internal mobility. Only glycine is unsubstituted at the α-carbon and this amino-acid alone permits easy motion within the main chain of peptide bonds. Proline, by contrast, is sterically constrained as a fused ring. The other α-amino-acids are nearly all further

[3] B. W. Dijkstra, J. Drenth, K. H. Kalk, and P. J. Vandermaelen, *J. Mol. Biol.*, 1978, **124**, 53.

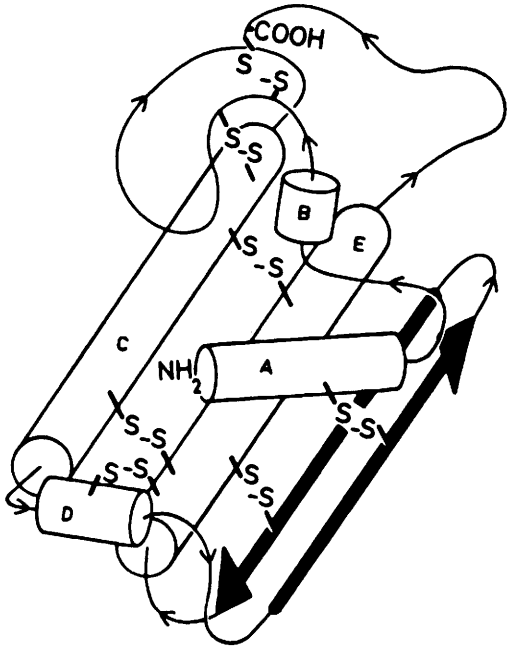


Figure 2 *An outline structure of a protein, here phospholipase A_2,[3] showing α-helical runs of amino-acids as cylinders and β-sheet runs as heavy black arrows. Disulphide crosslinks are shown, and structures of no secondary structure appear as thin lines*
(Reproduced with permission from Professor J. Drenth of Groningen, Holland)

substituted on the γ-carbon, the β-carbon being $-CH_2-$, so that much side-chain motion is possible about the β–γ bond even in folded proteins (Figure 3). The interior of the fold is not highly constrained, still less the surface. This implies, too, that the main chain can readily undergo minor movements in all proteins and major fluctuations in some.

C. Charged Groups, Fold Energies, and Fold Mobilities.—Proteins which fold in water usually denature in less polar solvents, even quite polar ones which can still form strong H-bonds, *e.g.* urea, and in both acidic and basic aqueous phases. The conclusion, that the fold is destabilized by reducing hydrophobic 'attraction' and/or weakening internal H-bonds in the less polar solvents and by increasing electrostatic repulsion between like charges at the extremes of pH, is consistent with energy calculations for folded proteins that indicate attractive components belonging to hydrophobicity and H-bonds and to electrostatic attraction of unlike charges. The hydrophobic term always acts in the sense that it leads to association of all non-polar side chains away from water. In general, it then generates co-operative globular structures. By way of contrast, the H-bond term leads largely to helical strings, β-bends, or planar sheets, *i.e.* secondary structure,

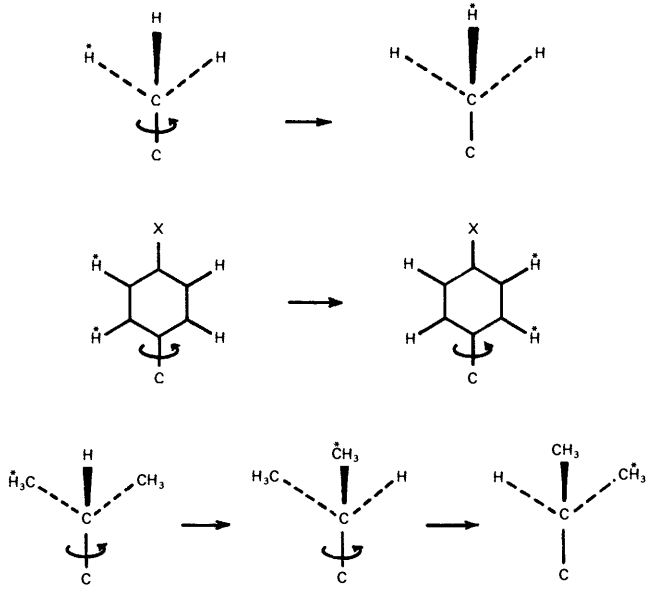


Figure 3 *Rotations of protein side chains*

but not to a globular tertiary fold since the internal H-bonds are stronger than the H-bonds to water but are only between two groups. The electrostatic term in itself appears to be too weak to generate a globular structure even if the sequence could allow a packing of opposite charges on an NaCl-like crystal lattice, and charged groups are almost invariably on the surfaces of proteins.[1,4,5] (We know that ammonium or guanidium bicarbonates and acetates are very soluble, *i.e.* such ions do not have a high affinity for one another.) Despite the affinity for water of charged residues, the H-bonded secondary structures remain stable even when they are composed of many charged amino-acids, provided that the sequence of charges is such that unlike charges are adjacent (see Section 5A). These secondary structures permit the charges to remain highly solvated. It seems likely that the best that electrostatic interactions in single proteins can do, therefore, is to combine in solvated linear alternant surface array, and only very small co-operative (Madelung) terms arise. There is, in fact, no known globular protein which has no hydrophobic amino-acids, but this does not mean that, on folding, interactions between unlike charges are unfavourable.[4,5] Gurd and his colleagues[4] have developed an extremely interesting analysis of how these interactions can contribute to fold energies. It does mean, however, that an observed tertiary fold is not likely to be that which gives the optimal free energy of the electrostatic terms. In the author's view, it is very unlikely that fragments of secondary structure will associate through electrostatic terms alone. It is for this reason that

[4] F. R. N. Gurd and T. M. Rothgeb, *Adv. Protein Chem.*, 1979, **33**, 73, and references therein.
[5] M. F. Perutz, *Science*, 1978, **201**, 1187.

segments of proteins which contain a large number of charged amino-acids will assume great importance in this article since they are likely to prove to be the least constrained within a tertiary fold. Their contribution to protein–protein interaction can be most important, as we shall see.

At the present time very many groups of experimentalists are using a wide variety of techniques to uncover the stability and the degree of mobility within the folded structures. Here n.m.r. is undoubtedly the most powerful as it can follow individual atoms. More general techniques use circular dichroism, H-exchange, specific-heat measurements, and fluorescence quenching. One of the objectives of the author's own studies using n.m.r. has been to discover the influence of charges in sequences upon the stability and mobility of a folded form (tertiary structure) of a protein and their function in assisting association between proteins. A summary of n.m.r. observations on proteins has been given elsewhere.[1]

All of the above comments apply to aqueous solutions. Many proteins are found in membranes and they can then experience two very different solvation zones. In the membrane the solvent is effectively that of a hydrocarbon, while at its surface the medium switches, though not suddenly, to an aqueous region. Taking the extreme non-aqueous phase we expect that proteins or parts of proteins in this phase will have very few charged groups exposed. The charges could now become an important cause of tertiary structure and the role of hydrogen bonds still acts to stabilize the isolated helical secondary structures in the non-polar region. If we are correct in asserting that polar interactions are likely to be found in pairs only, then the range of possible associations of helical (or β-sheet) segments in membranes is very limited. Perhaps non-aqueous media can only support simple side-by-side association of helices. A very important conclusion would then be that non-aqueous phases could contain channels but are unlikely to generate active sites of enzymes.

From the above it should now be clear that an understanding of proteins rests in the study of *composition and sequence*, *mobility within structure*, and *function*.

3 Pliable Materials

Before entering on an examination of the nature of the mobility of proteins it is as well to recognize that we shall need a change of language from the description of more or less rigid structures. The need for a new language arises from the nature of materials[6] in which there is considerable continuous motion. Such materials cannot be described by molecular structure since much of their interest lies in the rapid variation of structure. In fact, the language most suitable for description is that of statistical mechanics but here we resort to the simpler functional thermodynamics. In mobile systems we are concerned not so much with bond energy and direction, as in rigid molecules, but with entropic functions, ΔS. This is brought out in Table 1 which groups materials into three classes: liquids, pliant solids, and rigid bodies; the properties of the three classes are described in the Table. Those materials which have a very considerable rigidity

[6] S. A. Wainwright, W. D. Biggs, J. D. Currey, and J. M. Gosline, 'Mechanical Design in Organisms', Arnold, London, 1976.

are mainly used to *transmit force* and are termed *tensile*. For such materials Hooke's law holds (Figure 4), and since they deform very slightly even under a

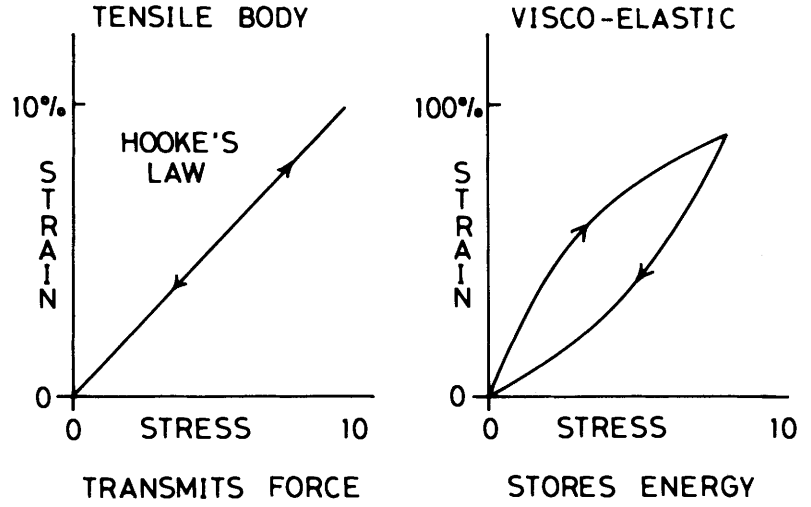


Figure 4 *The effect of applying a strain to two different types of body. The first develops little stress but transmits the force. It obeys Hooke's law. The second develops much stress and shows hysteresis on removal of the strain. It stores energy and evolves heat during the cycle*

considerable force they are not energized by this force (energy stored on compression = force × distance moved by force). A deformable body such as a rubber also transmits force but much more gradually since, when a force is applied to it, its deformation is large (its internal organization is increased, *i.e.* it loses entropy) and it stores energy (as entropy loss). Notice that both time and energy considerations are involved. *Mobile polymers are pliable solids* and store energy on impact of a force – they deform. Finally, liquids do not transmit forces if they are applied gradually but relax out of their way. Liquids must be constrained by containers to become transmission devices.

Within biology there are liquids, aqueous solutions, and, to be considered with them, random-coil polymers in solution. There are also more or less rigid bodies, bone, shell, and various fibres, and there are tightly folded proteins which are much closer to rigid materials than to liquids. Lock-and-key fitting theories of biological molecules are based upon such rigid structures. Finally, there are varying degrees of pliability in other biological materials from membranes to less tightly folded proteins. These pliable units are essential for molecular machines. My research effort in recent years has been to uncover the mobilities in large molecules, especially proteins, which might reveal the nature of the biological machinery. I expect this mobility of some proteins to be linked to the mobility of co-enzymes and the ions mentioned above so that they can act together in

control. The first major question must be 'What elements of composition of a protein confer mobility as opposed to rigidity to a structure?'

While we examine this question we must keep in our minds the important distinction between *transmission* of a message and its *transfiguration*. Transmission can be achieved by both tensile and pliant bodies, but pliant bodies transmit more slowly since they absorb the force applied at first in their own internal energies, rearrangements. To get energy, information of a message, from one place to another a tensile unit is faster, *e.g.* a tug applied to a rope. However, if the energy (of the message) is to produce a given effect (*e.g.* ring a bell) then at the end of the transmission system transfiguration of the energy to a different form is required. Pliant materials can be designed so that the required transfiguration is present in their change of average shape, which is not just a switch from one conformation to another but concerns changes in an ensemble of conformations. I shall call systems of this kind rhetropic. We can link together in our minds the functional parts of proteins as those that belong to energy transmission, more rigid, and those that belong to energy transfiguration, more mobile. In addition, since proteins are charged, mechanical changes are linked with electric-field changes.

It may be useful now to have in mind the following diagrammatic view of protein structure (Figure 5). The essence of the structure is that it contains three

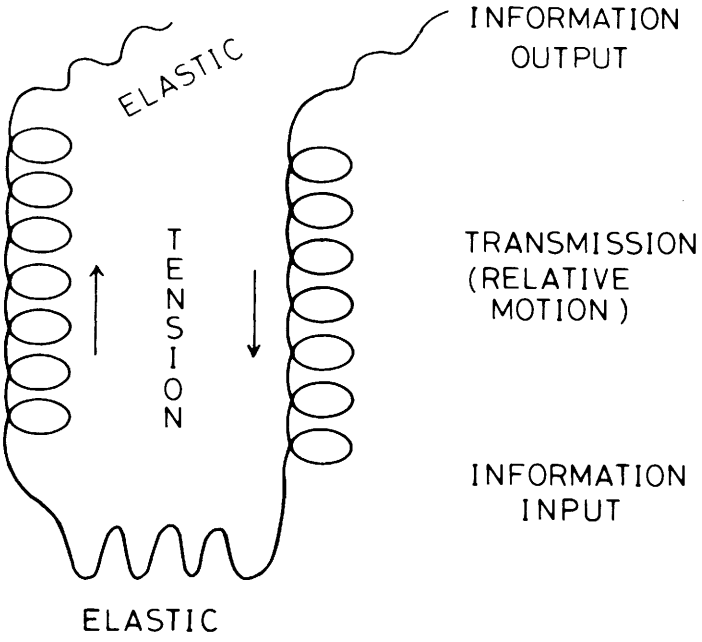


Figure 5 *A schematic diagram of a protein, showing how tension-bearing secondary structure could link more pliant sections of the protein, loops. The pliant regions retain energy and alter their conformational properties (not just a switch of conformation* $R \rightleftharpoons T$*) on applying strain*

regions. The centre portion is a more or less rigid set of frame units, helices (or sheets). There is little internal mobility of each helix, but small relative movements are possible since the helices are not crosslinked and bind together through interaction of mobile amino-acid side chains. There is at the bottom of the structure a loose loop of structure which acts as the acceptor of information not so much in the manner of the conventional lock-and-key fitting device but more akin to a mobile hand fitting a floppy glove. Koshland's[7] induced-fit model is a closer description than lock-and-key. However, it is not so much a matter of the fit of two shapes which is so important as of the binding, *i.e.* just the change of free energy, largely $\delta(\Delta S)$, altering the configurations of a receptor region. We may suppose that the adjustment of the energetics of this visco-elastic region of the protein is then transmitted by the tension and minor twisting developed in the central two helices to the upper region of the protein. Here we picture a second region of loose rhetropic structure which is the visco-elastic receptor region for a second small molecule (or protein) to which information is to be transmitted. There is a reciprocal relationship between two transfiguration sites based on minor movements of the helices which can cause gross movements, transfiguration, of the more mobile parts. The energetic effect is that energy associated with binding is to be found in the two visco-elastic regions separated by some distance by the tensile framework. The simplest model is two pieces of rubber at each end of a rod, one of which rests on a rigid support and to the other of which force is applied. The two bits of rubber are deformed and are energized to a greater degree than the rod. The deformed energized units can be recognized more easily (and differentially) than the small changes in the rod. We must see how such a model applies.

The above passage describes two extremes: rigid rods and pliant materials. An intermediate case is one in which large parts of the material are rigid but where small sections have several states open to them. Such a construction could be likened to a set of rigid elements connected by hinges, as on a door or gate. The use of this mechanical analogy will be seen later, but remember that in Part I we frequently used words like gate and pump to describe protein function.

The discussion will proceed by considering in order the simplest proteins and then increasing the complexity of organization. The simplest protein is the globular isolated enzyme. A more complex protein is one used in allosteric control where there are two binding sites ($R \rightleftarrows T$). More complicated again are proteins which have domain or subunit structures. A relatively rigid structural model will apply to these isolated simpler systems and can be extended to combinations of subunits which give rise to the structural units of cells – purely tensile proteins. The next stage of complexity is a protein exposed to more than one phase – a membrane protein. From there we can go on to protein–protein organization, and control, where control will be based on a continuum of states.

4 Enzymes

We and others have described the physical properties in solution of several

[7] D. E. Koshland, G. Nemethy, and D. Filmer, *Biochemistry*, 1966, **5**, 365.

enzymes of low molecular weight, using knowledge of their structures in crystals provided by *X*-ray diffraction studies.[1] These enzymes, which are roughly globular, are formed almost invariably from sequences of relatively hydrophobic amino-acids (see Figure 6) and are quite tightly structured, and the overall structure

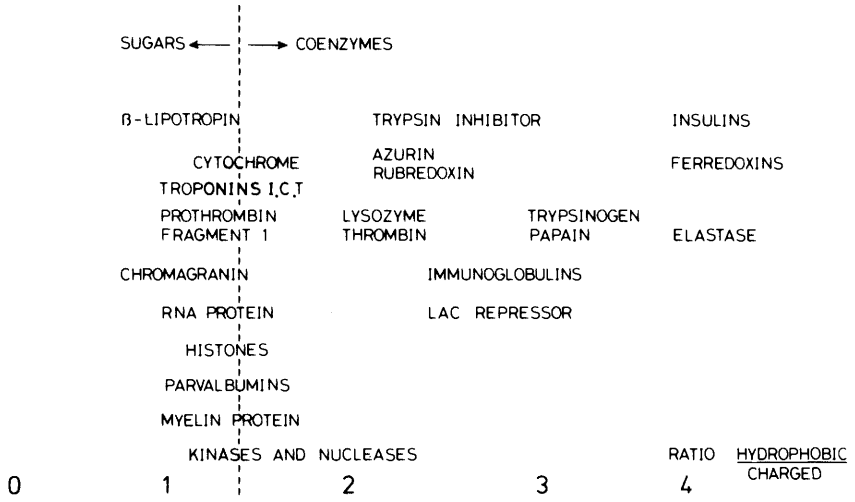


Figure 6 *A diagrammatic classification of proteins by their amino-acid composition. The charged residues are those in the first column and the hydrophobic residues those in the third column of Table* 2. *We expect the proteins on the right to be the least motile and those on the left to be more disordered. Note that nearly all enzymes are on the right but that many control proteins are on the left*

of the enzyme seen in the crystal is clearly maintained in solution. Their main-chain fold is maintained on the addition of substrate or inhibitors. However, the new feature in solution, which n.m.r. and other methods have established, is that there are several types of deviation from the crystal structures which may well have an importance quite out of proportion to their physical size. The first is local mobility of side chains both in the interior and on the surface of the proteins. Some examples are given in Table 3. The second is that on substrate or inhibitor binding there is usually, if not always, a small running change of structure which extends deeply into the protein. Finally, on the surface of these enzymes away from the active site there is usually a more motile region that is frequently part of a connecting loop between parts of the sequence which are themselves in tightly constrained secondary structures. (*Motile* is used to distinguish movement of chain segments from the mobility of side chains.) We have developed the hypothesis[1] that these mobilities and motilities can be used in four very different ways, all of which are essential parts of activity:

(i) local mobility in the active site pocket assists the substrate–product conversion;

Table 3 *Some examples of side-chain mobility and segment motility in enzymes*

Enzyme	
Lysozyme[9]	All tyrosines flip rapidly
	One tryptophan flaps slowly
	Some valines flip rapidly
	The region close to tryptophan 62 is mobile
Phospholipase A.2[10]	All tyrosines flip rapidly
	Some phenylalanines flip rapidly
	The *C*-terminus sequence is motile (Figure 7)
Cytochrome *c*[8]	Half tyrosines and phenylalanines flip rapidly and half flip slowly
	The region around isoleucine 57 is mobile
Nucleases[11]	All tyrosines flip rapidly
	Various short sequences of motility (see text)

(ii) running changes through the protein to the active site are capable of amplification into allosteric control;

(iii) mobile surface side chains of folded proteins may be of peculiar value in protein–protein recognition;

(iv) disordered sections of a protein are used to give graded responses quite different from a conformational switch (Figure 1).

This article is not particularly concerned with the mechanism of enzyme reactions. The stress here is upon the transmission of information. However, the discovery of the minor but extensive movements within enzymes on binding of their substrates immediately connects these topics since the state of an enzyme is signalled to its surrounds by these minor movements. Controls, *e.g.* co-enzymes, can act differentially on the different structures. Without detailed analysis the possibility of allosteric feed-back control can be seen to arise in these minor adjustments of the interconnected parts of proteins. If all that is required is a switch between two structures at the active site, in the presence and absence of a control chemical which binds at a second site, then in essence there need only be two conformations of the protein. Switching (R $\rightleftarrows$ T) may only require minor movements within the protein frame. This is the way in which Perutz[12] describes haemoglobin reactions with oxygen (see below). But we must note that if a more continuous series of activities is to be achieved in a single

[8] G. R. Moore and R. J. P. Williams, *Eur. J. Biochem.*, 1980, **103**, 495, and succeeding five articles.

[9] I. D. Campbell, C. M. Dobson, and R. J. P. Williams in 'Molecular Movements and Chemical Reactivity as Conditioned by Membranes, Enzymes and Other Macromolecules', ed. R. Lefever and A. Goldbeter, Wiley-Interscience, New York, 1978, p. 55.

[10] A. Aguiar, G. H. De Haas, E. H. J. M. Jansen, A. J. Slotboom, and R. J. P. Williams, *Eur. J. Biochem.*, 1979, **100**, 511.

[11] O. Jardetsky, J. L. Markley, M. N. Williams, H. Thielmann, and Y. Arata, *Cold Spring Harbor Symp. Quant. Biol.*, 1971, **36**, 257.

molecule then a range of conformational states is an essential prerequisite for an individual enzyme which then could have a continuous response to, say, pressure or temperature within each protein molecule.

The extension of our discussion from the cytoplasm to the membrane phase does not require new principles. The electron-transfer proteins undergo similar though even smaller changes on binding substrate, the electron,[8] and somewhat larger changes are seen on binding hydride.[13] Protonation also causes small changes.[14] Control over electron or hydride movement can be exerted by preventing the required movements. These are the smallest substrates, of course. Membrane-soluble co-enzymes act on these proteins in the same way as water-soluble co-enzymes act in aqueous phases (see Table 3 of Part I).

In addition to the tightly folded active site regions of an enzyme and the allosteric binding sites there are on the surface of these proteins much more mobile side chains and sometimes small stretches of motile chain. From a limited amount of data, as yet, it would appear that these regions act as initial points of recognition between enzymes and other groups such as other proteins (DNA and RNA) and between proteins and membranes. They then form an essential part of the final combined structure which carries with it some considerable motility. In the case of the simplest enzymes under examination here it is the generation of antibodies to them that reveals the principles of protein–protein binding. We have observed that it appears to be the most mobile or motile regions of enzyme surfaces which are most easily recognized, *i.e.* the most antigenic,[8] and these are often associated with charged stretches of amino-acids (see Table 4).

A. Tightly Bound Metals and Cofactors in Enzymes.—Those metal ions and cofactors which are locked in enzymes as part of the active site are usually chosen so as to be relatively rigid. We can compare the multiple fused rings of flavins, corrins, haems, chlorins, and the metal ions of relatively constrained co-ordination, Co, Ni, Cu, Fe, Mo, and Mn, all of which show little motion while exerting their catalytic functions, with the fluctional properties of the mobile co-enzymes NAD, Co–A, and ATP and the metal ions Ca^{2+}, Na^{+}, and K^{+} (with Mg^{2+} as an ion of somewhat intermediate character), all of which have control functions. We must stress that even the more rigid cofactors and ions have limited mobilities, much as organic active site residues do, which are of extreme value in reaction mechanisms, but they do not show the more gross conformational adaptability of the second group of molecules and ions. Immediately we see that it is the first group which is involved in the most specific and selective activities of enzymes, catalysis. The second group is more largely concerned with simple material transfer and control, where mobility has an obvious significance. There

[12] M. F. Perutz, *Scientific American*, 1978, **239**, 68.
[13] W. W. Smith, R. M. Burnett, G. D. Darling, and M. L. Ludwig, *J. Mol. Biol.*, 1977, **117**, 195.
[14] G. R. Moore, G. W. Pettigrew, R. C. Pitt, and R. J. P. Williams, *Biochim. Biophys. Acta*, 1980, **590**, 261.

Table 4 *Antigenic regions of some proteins*[1]

Lysozyme* (hen egg)	I	Arg(125)	Arg(5)	Glu(7)	Arg(14)	Lys(13)	
	II	Trp(62)	Lys(97)	Lys(96)	Asn(93)	Thr(89)	Asp(87)
	III	Lys (116)	Asn(113)	Arg(114)	Phe(34)	Lys(33)	
Cytochrome *c* (horse)		66—73 Asp-Tyr-Val-Asp-Glu-Pro-Lys-Lys					
		(This is the region near Ileu 57)					
Myoglobin (sperm whale)		16—21 Lys-Val-Glu-Ala-Val					
		56—62 Lys-Ala-Ser-Glu-Asp-Leu-Lys					
		94—99 Ala-Thr-Lys-His-Lys-Ileu					
		113—119 His-Val-Leu-His-Ser-Arg-His					
		146—157 Tyr-Lys-Glu-Leu-Gly-Tyr					
TMV coat protein		108—113 Leu-Asp-Ala-Thr-Arg-Arg					

*The region of Trp(62) is a mobile region of lysozyme

is an indication here that the binding region of the protein and its metal ion may be matched in mobility so that each assists the function of the other whether it be in catalysis or in control. We describe the special catalytic enhancement which arises from the combination of more rigid proteins and metals in the next paragraph and return to the control functions later.

B. The Entatic State.[15]—When a protein folds tightly, as is true of many enzymes, the overall free energy of fold lies in a deep minimum [Figure 1 (top)]. There are wide variations in the energies of local protein regions for, although most individual side chains will be stabilized in the folded protein, a few will find themselves constrained by electrostatic or repulsive forces. If exposed they are expected to be of heightened reactivity since strain can be relieved by chemical modification. Again, when a cofactor, haem or metal ion, binds to a tightly folded protein it can happen that the protein fold is somewhat distorted, energized, and the metal ion or cofactor binding also cannot be of that geometry which it would form with freely mobile binding partners. The total free energy of the combination is a minimum, but local strains in protein and cofactor are induced. Vallee and Williams[15] referred to this locally strained condition as the entatic state of a group and indicated ways in which it could be detected. They pointed out that such metal ions or other groups could be highly effective catalysts. If the idea is correct, and there is now much evidence in favour of such a description of active sites in some enzymes, then a major contribution to enzyme activity is the overall stability of the fold which generates heightened catalytic power locally in the strained environment. There then must be some compromise between the strongest binding of the metal, which is possible, and activation, and clearly a further compromise is required in that enzyme active sites must be of required, though limited, mobility. The very different binding of calcium is described in Section 6, 'Control Proteins'. Note that in Figure 6 many calcium proteins lie to the left whereas metallo-enzymes lie to the right.

C. Enzymes with Larger Complements of Charged Amino-acids.—All the 'rigid' enzymes we have discussed so far have in common a high complement of hydrophobic amino-acids in their sequences (see Figure 6), which we begin to correlate with the relative rigidity of their fold and the lack of internal mobility. As further examples of enzymes we consider now phospholipase A.2,[16,17] staphylococcal nuclease,[11,18] and two cytochromes b[19,20] since all of these enzymes have considerable runs of charged amino-acids in their sequences.

Phospholipase A.2 is an enzyme of known structure[3,16] in crystals (Figure 7),

[15] B. L. Vallee and R. J. P. Williams, *Proc. Natl. Acad. Sci. USA*, 1968, **59**, 498.
[16] B. Dijkstra, 'Structure and Mechanism of Phospholipase A_2', Doctorate Thesis, University of Groningen, Holland, 1980.
[17] A. Aguiar, G. H. De Haas, and R. J. P. Williams, unpublished results.
[18] F. A. Cotton, E. E. Hazen, and M. J. Legg, *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 2551.
[19] F. S. Matthews and E. W. Czerwinski in 'The Enzymes of Biological Membranes', ed. A. Martonosi, Plenum Press, New York, 1976, Vol. 4, p. 143.
[20] F. S. Matthews, P. H. Bethge, and E. W. Czerwinski, *J. Biol. Chem.*, 1979, **254**, 1699.

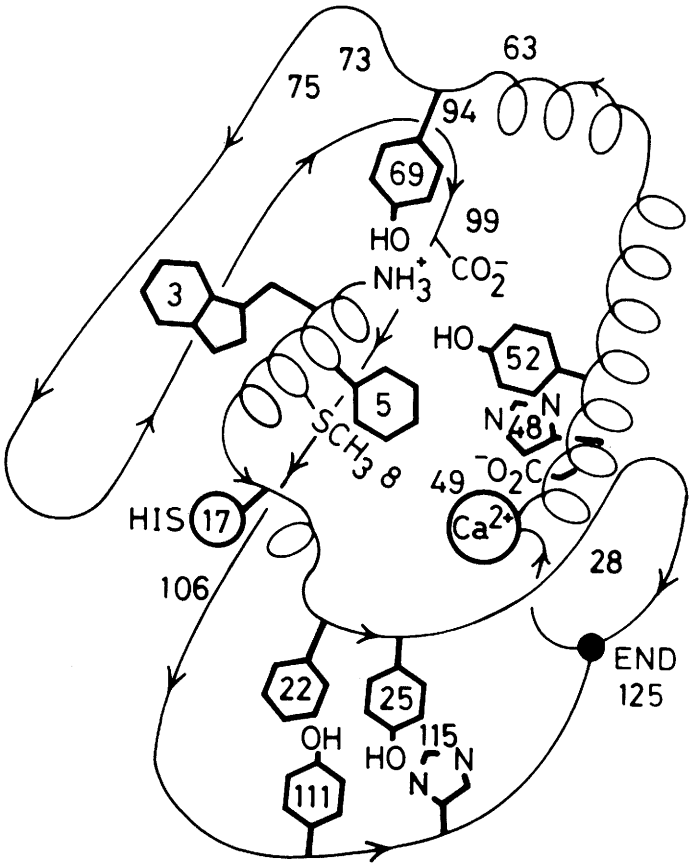


Figure 7 *A schematic diagram of phospholipase A*.2 (*compare Figure* 2). *The motile chain is toward the* C-*terminal* (125) *and its motility is following by the changes in its n.m.r. spectrum due to the interactions of* Tyr-111 *and* His-115 *with the aromatic rings of residues* 22 *and* 25

and its solution properties have been studied in great detail by De Haas in Holland and ourselves.[17] Though much of the work remains to be published it has been shown that in solution and crystal states this enzyme has a tightly constructed, hydrophobic, stable active site held together by seven –S–S– cross-links (Figure 2). However, two regions of the protein are not tightly bound. The *C*-terminal sequence occupies a variety of states, and its properties are modified continuously with changing temperature.[17] The sequence here includes many charged residues: (104)-Cys-Phe-Ser-*Lys*-Val-Pro-Tyr-Asn-*Lys-Glu-His-Lys*-Asn-Leu-*Asp-Lys-Lys*-Asn-*Cys*-(125). A second somewhat mobile region is the charged *N*-terminus which carries a tryptophan. This region of the protein is

concerned with both activity and binding to membranes, and it is then intriguing that the two properties are strongly interconnected by structural changes.

Cytochrome b_5 is a highly charged haem protein which splits into two peptides 1—93 and 94—140 when the protein is exposed to proteases.[19] The peptide 1—93 has a well defined crystal structure which is largely a set of helices surrounding a haem pocket. Its *C*-terminus is disordered. The *N*-terminus of the peptide joins the second peptide in the following run of charged amino-acids: (82)-*Asp-Asp-Arg*-Ser-*Lys*-Ile-Thr-*Lys*-Pro-Ser-*Glu*-Ser-(93). A part of this sequence is disordered in the crystal structure. The *N*-terminus, 94—140, amino-acids are hydrophobic and are thought to insert into the membrane. A good description of the protein is that it floats on the surface of the membrane with the hydrophobic tail in the membrane itself (see Section 5B). It has a highly motile charged connecting link between the water-soluble haempeptide and the membrane-soluble peptide. Now, the haem-containing peptide 1—93 is a globular unit not very different in percentage of charged amino-acids from cytochrome *c*. Cytochrome *c* is only held together by the hydrophobic haem group without which it is highly motile. It is probable that both these charged loose sequences are designed to allow easy insertion of the haem into the apoproteins. Similar comments apply to the apoprotein of cytochrome b_{562}, which as an apoprotein is not a well ordered set of short helices. On binding haem the helices become roughly parallel, two running in each direction.[20] Thus, the loose apoprotein helices are able to form around the haem, enclosing it (see Figure 8).

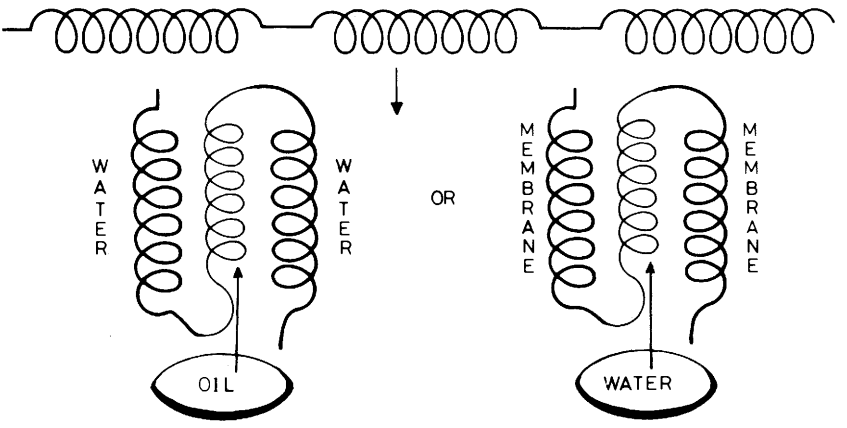


Figure 8 *An illustration of the way in which helices can come together to give a binding hole or a channel for a hydrophobic molecule, a lipid or a haem, or for a string of water molecules*

Cotton and his co-workers have studied the crystal structure of staphylococcal nuclease (Figure 9), and much of the solution work on this protein has been

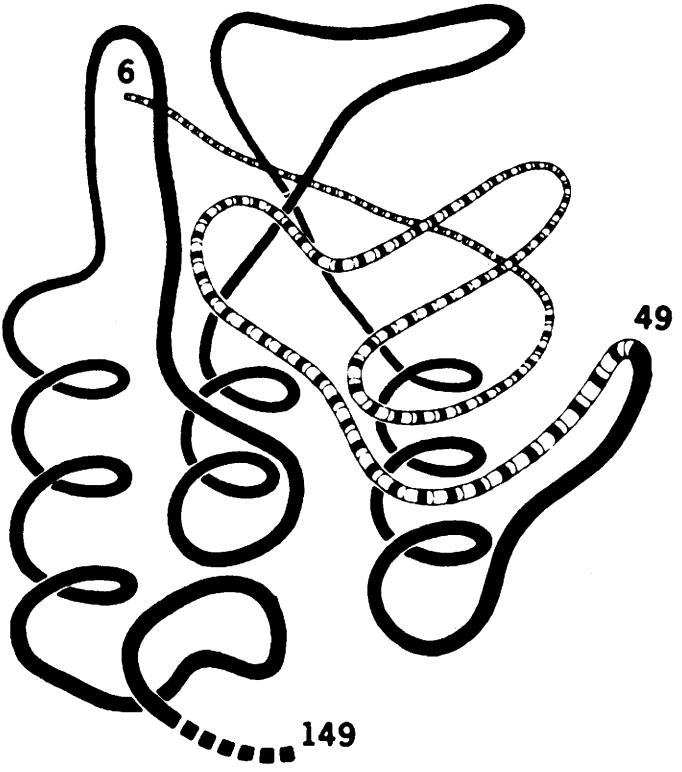


Figure 9 *The structure of* Staph. *nuclease (after Cotton and his co-workers*[18]*). The two parts of the chain are easily split apart near motile sequence around* 49

carried out by Anfinson and his group and by Jardetsky and his co-workers.[11] The following observations were made:

(*a*) The sequences at the two termini cannot be defined in the crystal. The two sequences are *Glu*-Asn-*Asp*-Ala-*Asp*-Ser-Gly-*Gln*$^-$ and $^+$*Ala*-Thr-Ser-Thr-*Lys*-*Lys*. Clearly these sequences are highly disordered and they contain many charges and few hydrophobic residues. In the nuclease-B versions of the enzyme a further acidic tail of some nineteen amino-acids are added, which are also disordered and contain up to seven charged amino-acids, two glycines, and only six hydrophobic residues.

(*b*) There is a middle section from residue 45 to 54 which is not defined in the lattice. It has the sequence *Lys-His*-Pro-*Lys-Lys*-Gly-Val-*Glu-Lys*. This sequence is highly charged and again is highly disordered. If these three regions of the

enzyme are removed it is a much more hydrophobic and normal enzyme in its composition.

(*c*) There are considerable regions of the protein which change conformation on addition of calcium and inhibitor. Very much of the protein undergoes some change in tertiary structure, *i.e.* change of side-chain positions, although the fold pattern is constant.

(*d*) Separated peptides made from sequences 1—44 and 48—134 are disordered.

We see that this enzyme is undoubtedly the most motile of any studied by crystallographers but that really it is two more or less hydrophobic domains with three flexible attachments. We also note that it is a very non-specific nuclease and in this respect differs from the many enzymes discussed so far and from DNA-nucleases and RNA-nucleases generally, which have much more hydrophobic and rigid conformations. I guess that this nuclease is related more closely to the transcriptase proteins than it is to conventional enzymes. In passing, I stress that there must be in the machinery for the synthesis of DNA and RNA (and proteins) proteins which handle these polymers within grooves like that in the nuclease but which transfer the information on picking up one polymer to a second groove where another type of polymer is made. Specificity is not the essence of this activity but it is required to be connected to a third site of energy input. It is in the connection between these activities that mobility in the machinery protein is essential. I wonder if in a way the nature of staphylococcal nuclease has been mistakenly described. For the hydrolysis of nucleotides (or proteins) a much more rigid-framed protein is the general rule and specificity is usually found. Is this nuclease related to a loose part of the central machinery of polynucleotide synthesis and is it being studied while it runs backwards? It even seems to bind DNA and RNA in different ways.

Before turning to other proteins there is a final feature of enzymes: they usually contain tryptophan. This large aromatic residue can act as a stabilizing platform in the structure undergoing little motion. In phospholipase A.2 and the staphylococcal nuclease the tryptophan is on the surface and more mobile. This very residue is absent from many control proteins – see below.

The feature that emerges from this analysis of enzyme structure and mobility is that it is the charged amino-acids, both their side chains, and, when they are grouped together, their regions of main chain, together with glycine, which can generate mobility and motility in structures. Curiously, from their antigenic properties it may be these very regions which are sites of protein–protein recognition (Table 4).

Now we have to resolve an apparent contradiction. We have observed that electrostatic forces do not on their own stabilize a tertiary protein fold, *i.e.* two pieces of the same protein do not bind together through ionic interactions alone. Yet we are finding that these interactions are major components in protein–protein binding. The contradiction is resolved by remembering that once a tertiary

structure has folded by virtue of hydrophobic forces the entropy of the segments of the protein which carry the charged amino-acids and which lie on the surface of the protein has been considerably reduced. Indeed, on the surface some oppositely charged amino-acids are known to form weak ion pairs and in so doing help to stabilize the folded structure. It is the interaction of two sets of such partially ordered charges on two different proteins which is now favourable through a type of multi-chelate effect. It remains true that the charged regions on the surface of either protein are mobile and may well be somewhat motile, and even after the binding together of two such regions considerable mobility and motility may well remain so that the surfaces are adjustable. Of course, some hydrophobic interaction can also assist in the binding of the two surfaces. (Similar principles can be applied to the binding together of proteins to the charges of membrane surfaces or to DNA and RNA, or to polysaccharides, since all these molecular systems have regularly spaced charges.) In essence, then, we are saying that two random polymers of opposite charge type interact feebly, but if one or both of the polymers is folded the other may well bind to it strongly, whether or not it is folded, if the proper matching of charges can occur. We shall develop this theme later.

The two proteins, cytochrome b_{562} and phospholipase A.2, provide further insight into organization. Both the apoprotein of cytochrome b_{562} and phospholipase A.2 illustrate how motile sections of proteins interact with hydrophobic molecules, *e.g.* lipids singly or in membranes, and can then be organized around them. Cytochrome b_{562} shows this in two ways. Firstly, it does not finalize its fold until a hydrophobic molecule, the haem, is incorporated. We can imagine a general principle of design which is that a protein of highly charged helices or other secondary structured elements will only form a stable tertiary fold if it is provided with additional hydrophobic molecules, *e.g.* of lipid-soluble molecules. One face of each helix needs to be hydrophobic. Alternatively, there can be parts of a tertiary fold that are motile in the form of the protein seen in water but that will readjust to a somewhat different tertiary fold on binding to a hydrophobic molecule. This is the property of the *N*-terminal membrane-binding part of phospholipase A.2. The further possibility illustrated by cytochrome b_5 is that the protein is built in segments, one of which is as already described but remains soluble in water on binding the lipid-soluble molecule, here the haem, while the second segment is like a section of phospholipase A.2 and enters the membrane. Finally, aqueous segments of the proteins can combine with other proteins by using the specific electrostatic interactions described here. We stress the essential part of mobility and motility in all these surface sections of the proteins and in the connecting links between segments. We shall now extend the discussion to larger enzymes and then to different types of protein.

D. Bi-lobed and Multi-lobed Enzymes.—Many of the enzymes we have described, *e.g.* lysozyme,[21] are not really simple globular proteins but are composed of two

[21] C. C. F. Blake, D. E. P. Grace, L. N. Johnson, S. J. Perkins, D. C. Phillips, R. Cassels, C. M. Dobson, F. M. Poulsen, and R. J. P. Williams, *Ciba Foundation Symposium*, 1978, **60**, 137.

or more connected domains. In some larger enzymes the construction of domains becomes more obvious. In the domains themselves the organization of the protein is much as that found in a truly globular protein, and there is usually little motility. However, the domains can move relative to one another. This movement allows the protein to entrap a small volume between domains when the connecting strands are called a hinge (see Figure 10). The amino-acid

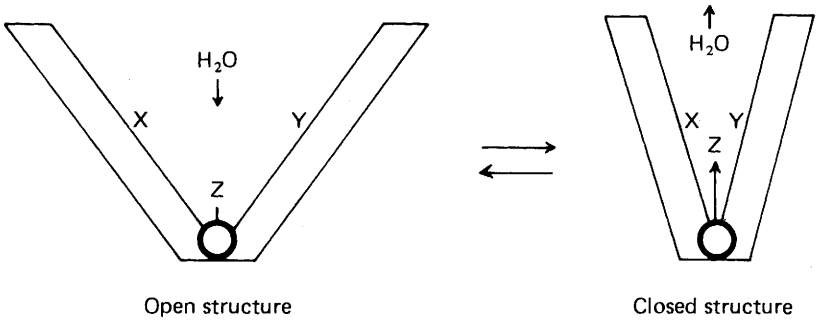


Figure 10 *A schematic diagram of a hinged protein which has two relatively rigid regions and a more mobile loop region. Water or other molecules can be excluded from an active site by the hinge action. The discussion of real examples is to be found in reference* 1

composition of the hinges, which are connecting loops of little secondary structure, are given in Table 5. We see that these loops usually contain a high content of the more polar amino-acids and/or a high content of glycine. This confirms the general idea of motility associated with such sequences.

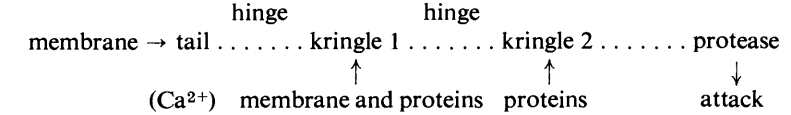
A typical example of an enzyme which is built in even more well defined domains is prothrombin. This is a very large protein with four quite distinct domains. At one end there is a domain which is a tightly folded enzyme contain-

Table 5 *The amino-acid composition of hinges in proteins*

Protein	*Hinge sequence*
Cytochrome b_5*	*Asp.Asp.Arg*.Ser.*Lys*.Leu.Ser.*Lys*.Pro.Met.*Glu*
Phosphoglycerate kinase†	*Lys*.Ala.Gly.Phe.Leu.Met.*Lys.Lys.Glu*.Leu
Pyruvate dehydrogenase‡	Gly.*Asp. Lys*.(co-enzyme).Ala.Ser.Met-*Glu*
Lactate dehydrogenase§	Gly.Ala.*Arg*.Gln.Gln.*Glu*.Gly.*Glu*-Ser-*Arg*-Leu
Tropomyosin¶	Ala.Gln.*Lys.Asp.Glu.Glu.Lys*
Prothrombin‖	(i) *Arg*.Ser.Gly.Gly.Ser.Thr.Thr.Ser
	(ii) *Glu*.Gly.*Arg*.Thr.Ser.*Glu.Asp*.His

*S. Tajima, K. Enomoto, and R. Sato, *J. Biochem.* (*Tokyo*), 1978, **84**, 1573.
†R. D. Banks, C. C. F. Blake, P. R. Evans, R. Haser, D. W. Rice, G. W. Hardy, M. Merrett, and A. W. Phillips, *Nature*, 1979, **279**, 773.
‡G. Hale and R. N. Hale, *FEBS Lett.*, 1979, **105**, 263. (This is a swinging arm hinge)
§M. Adams, personal communication.
¶see text.
‖see reference 22.

ing a high percentage of hydrophobic amino-acids and one or two hyper-reactive (entatic) amino-acid side chains. The middle two units are called kringles and have a looser structure and no very reactive amino-acid side chains, and, finally, there is a long tail of disordered, highly charged amino-acids.[22] The four domains are linked by polar, disordered, sequences which are readily chopped by proteases. The overall structure is, then, one of one tightly knit enzyme and two more loosely structured signalling devices, all joined by two hinge regions, and followed by a very loose, short, charged binding region for calcium ions. This dynamic structure is matched to the activity of the protein as a whole. Prothrombin belongs in the cascade of blood-clotting enzymes. Its most rigid end acts as a protease so as to amplify the initiating signal to clot blood by generating enzymes from proenzymes further down a cascade. It needs to recognize these proenzymes and does so in part through the first kringle. The recognition site is typically a charged region of the surface (compare antigenic sites and see Section 6). The second kringle and the calcium binding tail act as the trigger function in recognition and binding of a cell membrane. The organization is then



Organizations of this kind, which greatly limit free diffusion of units but which are required to act together, are found elsewhere, *e.g.* in the complement system plasminogen, and in the lipoxygenase 1 of seeds.[23] There are also the beginnings of such combination of activities within one protein in the phospholipases A.2 (see Section 4C). We begin to see that ways of incorporating enzymes into organizations are very like those for the incorporation of co-enzymes (see Part I).

E. Swinging Groups.—The parallel between the constraints on diffusion of enzymes and the description of the restrictions on diffusion of small molecules must not be missed. At first in Part I we described freely mobile co-enzymes and then co-enzymes bound on swinging arms in larger freely diffusing assemblies. In Part II we have described freely diffusing (usually extracellular) enzymes, *e.g.* lysozyme, and then enzymes bound in larger units in which the enzyme moves on a swinging arm or floats in a membrane. The next step is to control diffusion further by building the proteins into larger structures, especially in particles and membranes. We need to see how large structures can be built and we can do this by looking at the two major structural systems, fibrous structures and the different membranes of vesicles and of the cytoplasm of the cell.

(The swinging-arm feature of some bound co-enzymes referred to above must restrict them to the surface of a protein. At present we have no general knowledge of the sequences to which the arms are attached, but see Table 5. It is not

[22] N. Pluck, P. Esnouf, and R. J. P. Williams, *Thrombosis and Haemostasis*, 1979, **42**, 95.
[23] M. R. Egmond and R. J. P. Williams, *Biochim. Biophys. Acta.*, 1978, **535**, 418.

then possible to make general statements about the protein participation in this mobility. It is probably a feature of it and of other modifications of protein side chains, *e.g.* phosphorylation, that the modifications are on loops of enzyme structure and that these loops are more motile than the other parts of the structure. We expect them to be charged. An example of a swinging-arm or side-chain motion which has been very well described is the case of the tyrosine or carboxypeptidase[24] which can pop in and out of the active site. This is an illustration of the obvious fact that the moving segment of a protein can be of any size from one amino-acid to most of the protein.)

5 'Rigid' Structural Proteins

A. In Water.—There are several types of protein which form relatively rigid structures in the aqueous and organic media of cells but which do not have a tertiary fold nor a catalytic function. First, we describe those which are found in water. The most obvious are the strands of connecting tissue such as the long multiple helices of collagen. These proteins are composed mainly of simple amino-acids, glycine, proline, and alanine, and only in rare stretches are their sequences broken by isolated runs of rather polar amino-acids or by large hydrophobic residues. The H-bonding of the helices, not charge interactions or hydrophobic forces, are the main source of stability. Packing such chains together provides very strong connective tissue which can be further strengthened into hard materials, bones, by incorporating needle crystals of ionic solids such as hydroxyapatite between the collagen strands.[25] The strands of protein which form flagellae are probably similar.

A somewhat different secondary structure can be built by interrupting α-helical stretches regularly so that the protein is really made in short lengths of helix, domains, which are connected one to another by floppy strands. Such sets of helices can be bent around on the edge of a disc. When all the hydrophobic residues are placed in the sequence so that they point inwards an ideal cavity is formed for the transport of lipids (Figure 8). Such a disc of helices will remain water-soluble. This device is similar to the fold used in cytochromes b_{562} to bind haem. A feature of these proteins is that large hydrophobic amino-acids are placed regularly with respect to the hydrophilic or charged residues. It is the periodic run of groups which gives the secondary structure, and these proteins have little tertiary structure without crosslinking agents such as metals, co-factors, or lipids.[26] Another example can be taken from amongst the muscle proteins where strong fibres are required, but which must also have the ability to flex. These proteins extend through cells and connect to other cells or external structures. In tropomyosin, for example, there is a regular helical repeat unit of amino-acids [(hydro)-phobic-phylic-phylic-phobic-phylic-phylic-phylic]$_n$, *e.g.* (-Ala-Glu-Lys-Tyr-Arg-Asp-Glu-)$_n$, which gives an helical tension-bearing

[24] J. T. Johansen and B. L. Vallee, *Biochemistry*, 1975, **14**, 649.
[25] C. Berthet-Colominas, A. Miller, and S. W. White, *J. Mol. Biol.*, 1979, **134**, 431.
[26] J. R. Segrest, R. L. Jackson, J. D. Morrisett, and A. M. Gatto, *FEBS Lett.*, 1974, **38**, 247.

strand (Figure 9). This repeat breaks some four times in tropomyosin, *e.g.* -Ala-Gln-*Lys-Asp-Glu-Glu-Lys*, which has five charged amino-acids in sequence. It is observed that such a sequence 'break' forms a natural cleavage point for enzyme attack and is thus very exposed.[27] It appears very likely that such sequences give the motile loops, protein–protein binding sites, and potential hinges in the structures (Figure 10). In tropomyosin the other sequences of irregular nature are 12—17, 26—31, and 215—220. Note that the break sequences are all highly charged and out of register. These fibrous proteins have no tertiary fold but form secondary and quaternary structures. Similar sequences are found in the above lipoproteins and in myosin and keratin.

The structural proteins which are described above belong to muscle and connective tissue often on the outside of cells. In the membrane regions there are great differences in structural units between bacteria and plant cells and animal cells in that the former have strongly crosslinked retaining walls. In animal cells the shape of the membrane is maintained loosely by the proteins on its outer and inner faces. Examples are spectrin[28] and myelin.[29] These proteins have a high content of charged amino-acids and in the absence of the membrane phospholipids have reduced tertiary structure. (Myelin is somewhat like parts of histones; see below.)

Inside the cell the connection between membranes and between membrane and nucleus is maintained by a protein called tubulin and actomyosin fibres (see Section 8A). It is thought that the protein tubulin folds into an elongated structure and then polymerizes. Of great interest here is that at the *C*-terminal of tubulin there is an extremely acidic sequence with some fifteen glutamic acid residues which we take to be disordered in the monomeric protein.[30] The state of polymerization may well depend upon this terminal sequence since it is known that the polymerization also depends on Mg^{2+} and Ca^{2+} concentrations. Tubulin may well be somewhat similar to certain virus coat proteins.

Throughout the above description it was important to notice the role played by the out-of-register sequences of charged amino-acids in more motile segments. It is now appropriate to use as examples the structures of two non-enzyme proteins which may be used in biological structures in very different ways – the tobacco mosaic virus (TMV) protein[31] and pre-albumin.[32] The features of both of them which we wish to stress are that they have a central core from which there extends either one long loop of sequence which is not seen in the *X*-ray electron-density map, TMV protein, or two loops which are less distinct than expected in the electron-density map and are more mobile than the core. The notable

[27] L. B. Smillie, *Trends in Biochem. Sciences*, 1979, **4**, 151.
[28] G. B. Ralston, *Trends in Biochem. Sciences*, 1978, **3**, 195.
[29] L. A. T. Littlemore, *Aust. J. Chem.*, 1978, **31**, 2387.
[30] R. C. Lu and M. Elzinga, *Biochem. Biophys. Acta.*, 1978, **537**, 320.
[31] (*a*) A. C. Bloomer, J. N. Champress, G. Bricogne, R. Staden and A. Klug, *Nature*, 1978, **276**, 362; (*b*) A. D. McLachlen, A. C. Bloomer, and P. J. G. Butler, *J. Mol. Biol.*, 1980, **136**, 203.
[32] C. C. F. Blake, M. J. Geisow, S. J. Oatley, B. Révat, and C. Révat, *J. Mol. Biol.*, 1978, **121**, 339, personal communications.

feature of these loops is that they can be likened to anchoring chains for making contact either with neighbouring subunits or with polynucleic acids. The sequences in the loops are highly charged. It is considered that these motile loops of sequence can search out the corresponding binding site on another polymer by making a stepwise series of connections, using electrostatic attraction, both for binding and to guide their coming together. Motility at the junction may be maintained almost as required, and it will be possible to separate the binding faces again by a stepwise process so that large activation energies of assembly and disassembly are avoided. We suppose that this type of protein is not too dissimilar from control proteins, histones, and some ribosomal proteins, for example, many of which are thought to have long terminal mobile segments when isolated. They all play a structural role but obviously they can also exert control over structure changes if they are open to modification by diffusing messenger, *e.g.* the calcium ion (see later).

Structural proteins have now been described for the production of threads, for binding RNA, and for binding DNA. These proteins also help to stabilize the membrane and the particles in Figure 3 in Part I. It is obvious enough that enzymes must be bound into these particles and thread structures in order that material and energy can be handled in association with them. It is also obvious that energy must modify them. It does so in a continuous or rhetropic manner (see Section 3, 'Pliable Materials') through control proteins, described below.

B. In Membranes.—Helices which are more hydrophobic than the above, due to the inclusion of greater numbers of such residues as leucine, isoleucine, and valine in the repeating sequences, are soluble in the lipid matrix of the membrane. These helices often associate one with another. The best known examples of these helices are alamethicin,[33] gramicidin A,[34] the purple protein,[35] the gap junction protein,[36] and glycophorin.[37] Each helical stretch is about twenty amino-acids long so that each just spans the 40—50 Å membrane (see Figure 5 of Part I). By itself, a single such helix is a rigid-tension rod connection between the two aqueous phases inside and outside the cell membrane (see glycophorin below). When the helices associate they can form membrane channels (as in the disc of the lipoprotein and the haem site of cytochrome b_{562}), and the sizes of these channels, now in the membrane, depend upon the precise nature of the helix (see Figure 8). A narrow water channel can bind protons only while a wider channel may also permit binding of large cations and anions. (The gated channel described by Edmonds[38] is of this kind.) However, it is possible to imagine almost any type of channel binding specificity. The more mobile parts

[33] D. R. Martin and R. J. P. Williams, *Biochem. J.*, 1976, **153**, 181.
[34] E. Bamberg, H.-J. Apell, H. Alpes, E. Gross, J. L. Morrell, J. F. Harbaugh, K. Janko, and P. Läuger, *Fed. Proc.*, 1978, **37**, 2633.
[35] R. Henderson and P. N. J. Unwin, *Nature*, 1975, **257**, 28.
[36] G. Zampigh and P. N. T. Unwin, *J. Mol. Biol.*, 1980, **135**, 451.
[37] M. R. Egmond, D. Rees, J. Welsh, and R. J. P. Williams, *Eur. J. Biochem.*, 1979, **97**, 73.
[38] D. T. Edmonds, *Chem. Phys. Lett.*, 1979, **65**, 429.

of the proteins are likely to be in loops connecting the helices or in terminal regions of the protein exposed to solvent.

The channels through the membrane are the pathways for food, waste, and ions to the surrounds of the cell. Once again, they will not be simple channels but must have enzymes connected to them so that the channels become pumps (Figure 11) or are gated by hinged energized action (Figure 10). Once the chan-

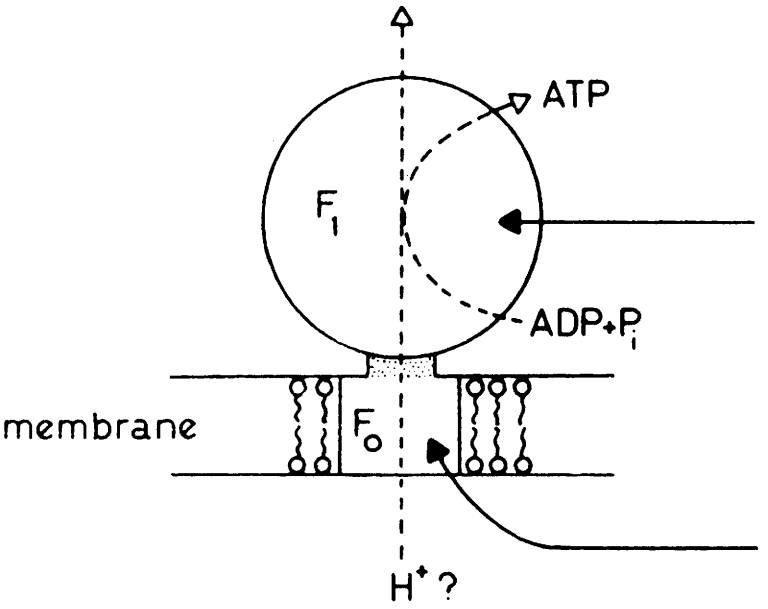


Figure 11 *An illustrative diagram of an ion channel, here for protons, connected to an ion pump, here operated by the energy of ATP hydrolysis. The reaction can work in reverse to synthesize ATP. The two arrows point to* F_0, *the channel in the membrane, and to* F_1, *the ATP-utilizing pump, which may work somewhat like a hinge protein (Figure* 10)

nels are controlled in this way then the electrical apparatus of the cell can be constructed. Again, a simple helix segment of a protein in a membrane between a sensor section on the outside of the cell and a connecting link to energy on the inside of the cell can provide the cell with knowledge of the outside world. Returning to Figure 1 of Part I, we are beginning to see how the whole organization of proteins, structural and enzymic, are connected together to permit the small molecules, ions and co-enzymes, to function in control.

6 Control Proteins

A final class of proteins which functions only in activating enzymes or structures is grouped together in Table 6. Each group of such proteins is not of very well defined structure except in the presence of the chemicals to which it binds, both the small molecules of Part I and the large proteins with which it forms a unit.

Table 6 *Some control proteins*

Protein	*Composition characteristics*	*Function*
Troponin C	Negatively charged	Binds calcium/muscle control
Troponin I	Positively charged	Muscle regulation
Histones	Positively charged	DNA regulation
Some RNA proteins	Positively charged	Ribosomal control
Myelin (membrane) protein	Positively charged	Membrane control
Chromagranin	Negatively charged	Internal vesicle structure

The histones clearly fold readily in conjunction with DNA, some of the ribosomal proteins fold only in part in the absence of ribosomal RNA, and the calcium-binding proteins fold but loosely in the absence of calcium, *e.g.* the proteins of muscle, troponin C, I, and T, are not well structured. In these polypeptide chains there will be sections of secondary and of loose tertiary structure, but much of the chain is disordered. These proteins have in common a large number of charged residues which do not run in simple repeating sequences regularly broken by hydrophobic residues (see Figure 6 and Table 7).

Table 7 *The composition of calcium trigger proteins and metallo-enzymes*

Protein	*% Amino-acid composition*		
	Hydrophobic	*Charged*	*Ratio*
Troponin-C (rabbit) (Ca)	43	38	1.1
Calmodulin (bovine) (Ca)	46	34	1.2
Parvalbumin (frog) (Ca)	46	31	1.5
Azurin (*P. fluorescens*) (Cu)	50	21	2.4
Ferredoxin (taro) (Fe)	48	26	1.8
Rubredoxin (*P. elsdenii*) (Fe)	57	31	1.9
Carboxypeptidase (bovine) (Zn)	55	17	3.2

In Figure 6 we see that while very few enzymes have other than a high content of hydrophobic amino-acids, and while those proteins which are used in structures can be either hydrophobic or hydrophilic sequences of regular repeat, control proteins have highly charged sequences of no pattern. The control protein sequences are then in large part similar to the short disordered sequences found in some enzymes (Section 4C) and in regions of hinges in both enzymes and structural proteins. While stress is placed on the charged amino-acids it should be noted that glycine is also particularly important in these motile sequences since it provides no constraint on chain motions. Another residue which can become very valuable is serine since it can be phosphorylated, whereupon the

loop containing it will become highly charged and of different conformational motility. Charges are the basis of local disorder in individual proteins but the basis of the assembly of proteins into organization.

A. The Structures and Dynamics of Fast Control Proteins.—Not many control proteins have been examined in detail as yet. They must be examined in the free state, bound to the molecules and ions which affect them and to those which are affected by them. Their motility in the absence of binding is seen in several ways: (i) the rapid rotation of all aromatic residues, (ii) they do not contain tryptophan, (iii) the low denaturation temperatures and the wide range of temperature over which sections of them melt, (iv) very rapid hydrogen–deuterium $\rangle$NH exchange, (v) if they have saccharides bound to them then they, too, are extremely motile.[1] Such properties are common to the calcium-control proteins, calmodulins, and some proteins associated with them, troponin I and T, to the histones, to some of the ribosomal proteins, and to some membrane proteins, glycophorin. We consider that this motility is an essential part of their function. The basic suggestion is that their rapid fluctuations allow them to respond quickly and to transmit information rapidly. It also allows them to respond continuously to their environment since they do not occupy one of two major states ($R \rightleftarrows T$) but adjust their ensemble of states (ΔS) continuously as increasing stress is applied to them. They are rhetropic (Table 1). This ability is shared by other polymers of high configurational entropy and in particular it is a property of many polysaccharides. The glycoproteins may well be particularly fine sensors of stress. Table 8 gives a listing of polymers in different classes to be compared with Table 1 (and see Figure 5).

Table 8 *Types of protein* (*polysaccharide*)

Random	*Flexible*	'*Fixed shape*'
Highly charged	Highly charged or glycine-rich	Hydrophobic or regularly spaced charges
No crosslinks	Crosslinked	Many enzymes with or without crosslinks
Chromagranin A	Control proteins rubbers	Collagen
(RNA, DNA?)	Glycoso-amino-glycans	Cellulose

B. Functions of Loops of Charged Amino-acid Sequences.—We must now stress the particular functions which rest in these irregular charged sequences. We describe our view of them, using Figure 12 which illustrates the construction of the calcium-binding proteins (see Figure 5). We consider that in the absence of calcium the proteins have several motile regions, loops, often as many as eight,

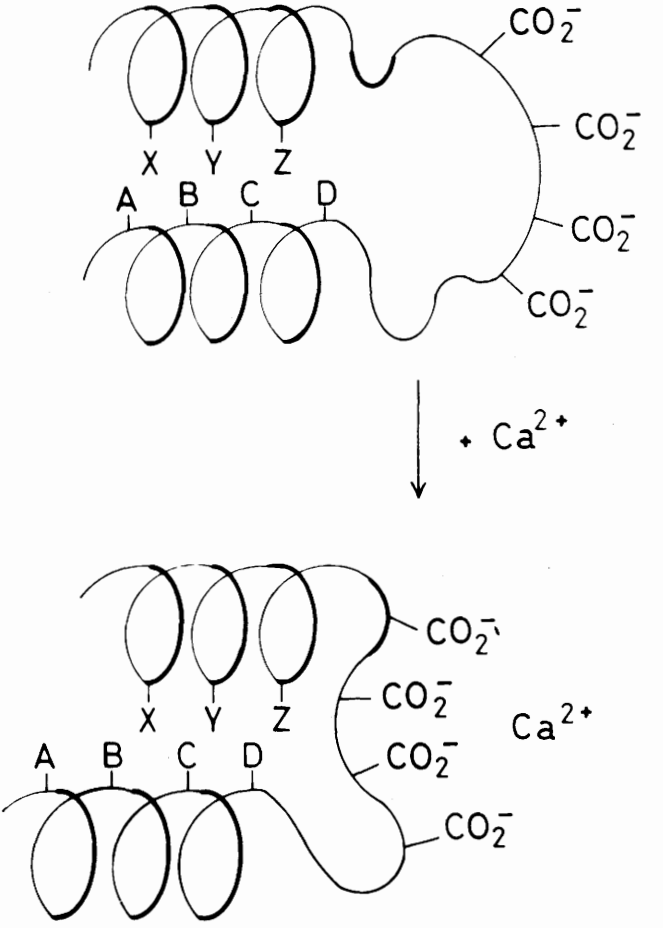


Figure 12 *Many calcium-binding proteins are control proteins which operate by the binding of calcium to groups of carboxylate residues in a limited sequence on a loop connecting two helices. The binding adjusts the interaction of the protein to enzymes and mechanical structures* (*see Figure* 5)

connected by more rigid helices which can move relative to one another.[39] Calcium binding causes four of the mobile regions to become structured, and this binding causes restrictions on motion at the other four loops.[40] There is energy storage in the eight loops and tension transmission by the helices. The loops that do not bind calcium interact through several mobile attachments with the proteins which they activate, *e.g.* calcium binding to troponin C activates troponin

[39] R. Kretzinger, *C.R.C. Crit. Rev.*, 1980, **8**, 179.
[40] B. A. Levine, J. M. Thornton, R. Fernandes, C. M. Kelly, and D. Mercola, *Biochim. Biophys. Acta.*, 1978, **535**, 11.

I and T so that muscle action is initiated (Figure 13).[41,42] Such proteins as I and T placed between the calcium receptor, protein C, and the muscle fibre are motile, as is the head group connection between them and the muscle fibre protein itself. The motility of the binding site of tropomyosin has been described above. In such an arrangement there is a relay of ordering within the proteins C to I + T to muscle proteins which can be continuous in its strength and not just an R ⇄ T switch. We can see this in the following way.

The parts of the proteins which interact are the most mobile parts. It follows that they are not to be described by structures but only by ensembles of structure.

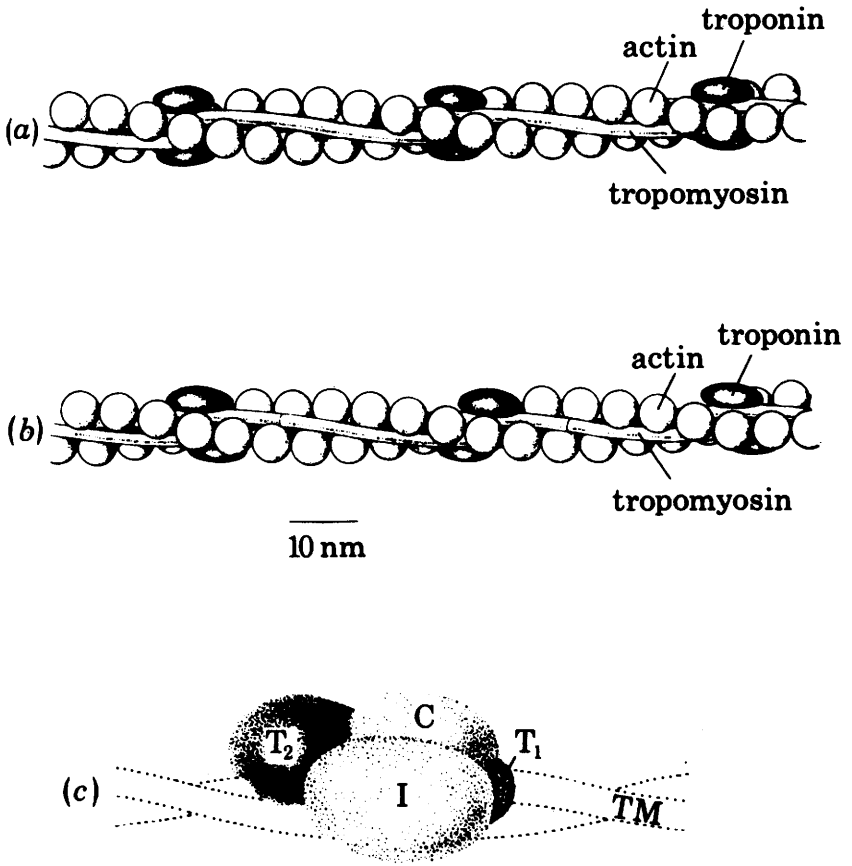


Figure 13 *A description of muscle proteins developed by Ebashi*[41] *and Perry.*[42] *Troponin C (calcium-binding), I, and T connect through tropomyosin TM to activate an enzyme (energy-generating) which causes tension to increase in the filaments.*[41]
(Adapted by permission from S. Ebashi, *Proc. R. Soc. London, Ser. B*, 1980, **207**, 259)

[41] S. Ebashi, *Proc. R. Soc. London, Ser. B*, 1980, **207**, 259.
[42] S. V. Perry, *Biochem. Soc. Trans.*, 1979, **7**, 593.

If we look just at these regions we would classify them, using Table 1, as pliant materials, rhetropic, and as such the variation of entropy in them with force applied can be continuous. The connections are then like rubber joints. Clearly, this must be very important if there is to be a continuous variation in tension in the muscle fibres *via* the variable interaction strengths between the proteins, troponin C, I, T,–tropomyosin–actin. Finally, we ask about the role of the chosen ion, calcium in this organization.

The calcium-ion co-ordination sphere is fluctional. Thus, of all ions its interactions with a protein are most like the mobile interactions of the charged residue sequences discussed above as the binding sites between proteins. In fact, calcium, too, is an excellent flexible crosslinking agent and it, too, has fast on/off reactions. Table 7 makes the distinction between the motility of the calcium proteins and the rigidity of iron, copper, or zinc proteins. Calcium and the control proteins are highly selected partners.

The binding constant of calcium is now an extremely interesting *variable*. It is not to be represented by a constant $\log K_{Ca}$ independent of the energy storage in the troponin C except when calcium binding occurs to the isolated protein. Even to this protein calcium binding proceeds with a continuous series of movements since the fluctional properties of this ion permit the *continuous* adjustment of ligands on its surface. These ligands gradually lose the disorder of the unbound state to give well structured sites. There is a continuity of binding constants. If the protein is interacting through similarly motile regions to other proteins and the two regions, calcium-binding and protein-binding, are linked by tension rods as in Figure 12, then the strengths of calcium binding and protein–protein binding are continuously and mutually adjusting. This is not an allosteric $R \rightleftarrows T$ switch but an infinitely variable rheostat. The binding constants for calcium could even be related to the tension in the muscle fibres *via* the variable interaction strengths between proteins C, I, and T – muscle proteins.

Now, the above description is applicable to any series of binding reactions in which there are many conformations of sets of binding groups. Energization of one alters them all in a continuum of states. The value of the control device rests not only in its final thermodynamic states, unlike a two-state switch, but also in its kinetics. The very fact that there is a continuum of states with no barriers between means that the machinery responds smoothly and very rapidly. We believe that protein–protein, protein DNA–RNA, protein–membrane, as well as protein–fast messenger interactions, are based on these principles. Electrostatic forces are extremely valuable in this context. The mechanism is quite different from that of a slow ($R \rightleftarrows T$) allosteric switch by a co-enzyme.

Before we analyse these structures further we should look at the nature of man's machines and the demand on the materials of which they are made. In this way the essential requirement for mobility within structure and its relation to function can be brought home.

7 Biological Structure

Before going forward to a description of machines as it applies to the proteins

in a living organism, *e.g.* a single cell, we must again inspect the language we are using, especially the concept of a structure. There is a close relationship between the structure of sodium chloride in a crystal and the structures we have just described of various types of *extracted* protein, despite the internal dynamics of the proteins. This close relationship is not maintained in biological 'structure'. In Part I the point was made that a biological structure is not stable, it is metastable through a constant energy flux – it is a dissipative structure.[43] We used the analogy of a river. Without a constant flow of material and energy biological structures collapse, *e.g.* plants, muscles, *etc.* This is not true of inorganic or protein structures. Of necessity this means that through many of the proteins which we have described there is a flow of energy and material and that the steady-state shape of a particle or a cell is not the sum of the structures of the isolated component structures. Each protein in a cell has a structure dependent upon the energy flow. If we look at an operating man-made machine we may see an apparently static object although it has not the same structure as the machine at rest. While this is also true of a biological machine, when the biological machine is stopped the whole assembly begins to crumble. If left in this state for a relatively short while it is dead. The structure seen is a reflection of the energy content of the cell or vesicle or particle and the 'structure', we must understand, is the sum of a number of states which contribute to the constant fluctuations, only some of which are easily observed by any study of isolated objects. Moreover, the 'structure' of each part will vary with the energy input over a continuum of 'structures'. A plant wilts gradually and its component parts, proteins, must also wilt gradually. We can now turn to a description of machines in whole cells.

8 Machines

A machine must have the following elements: a connected source of power, a frame, control units, and a connection to an output.[6] In biology all these pieces must be made from proteins. The source of power is chemical transformation (of ATP) and therefore enzymes are required as parts of the biological machines. Transmission of power is conformational (mechanical) and must require considerable movement, but our description of enzymes makes it likely that they will be rather rigid (close to their active sites) and so indeed are parts of the structural (frame) units. There is a requirement for connections, sequences which are much more mobile. We have pointed to the conventional on/off description of allosteric or switch (R $\rightleftarrows$ T) of co-enzyme control of enzymes. However, it is obviously advantageous to devise a continuously tunable control in machines since then (*a*) any one unit can be varied continuously, (*b*) the machine is likely to work with greater speed of response, and (*c*) it works with greater sensitivity in all its individual parts. Although (R $\rightleftarrows$ T) switches if present in large numbers can give, statistically, any activity between those of R and T states, individually they cannot respond in this way, and collectively the

[43] I. Prigogine and R. Lefever in 'Molecular Movements and Chemical Reactivity as Conditioned by Membranes, Enzymes and other Macromolecules', ed. R. Lefever and A. Goldbeter, Wiley, New York, 1978, p. 1.

activation energy for change will be large and the response slow. The control proteins that we have described through their lower level of structure (high motility and rapid equilibrium between many conformations) can change continuously and rapidly. Returning to Section 2, the entropy of these proteins (pliable materials) is large. It is ΔS which is continuously adjusted by constraints, and the forces which act from the control are entropy-driven forces. The steric constraints are now of lower value and these proteins connect to other proteins, *etc.*, through a different type of recognition from that seen in enzymes. Of course, such devices are sensitive to temperature and pressure in a continuous manner. Biology is thus exquisitely sensitive. We shall refer to proteins which behave in this way as rhestatic.

It should be very clear that the essence of a machine is not structure but the ability to transfer energy and that, although structure is required, movement is equally essential. Theories of machines or biopolymers which are just structural must be replaced by dynamic descriptions, and we could well look at concepts from materials and engineering sciences.

Here, a useful parallel is any simple machine such as a bicycle. It is easy to see one function of the frame – to carry the weight of the man. The energy input is also easily understood – pressure on the pedals. It is not so easy to describe the transference of power to the chain and then to the drive on the wheels. When different forces are applied the tension in all the elements of the structure changes. Notice, too, that some parts of the bicycle have to be flexible, and various choices of materials have been made which include metal frame, linked-segment metal chains, rubber tyres, and oil in a contained volume (see Table 1). Every machine is composed of the same categories of materials and functions and the rules must apply to biological machines. There is a relationship between the energy input to the bicycle, the energy storage in its various parts, and the tensions and distortions developed.

A. The Nature of Biological Machines.—*Mechanical Machines.* We are now in a position to discuss some biological machines. We shall describe mechanical devices first and then turn to electrical devices. Chemo-mechanical devices can be linked directly to electrical devices in biology as the current carriers cause conformational changes in *charged* proteins as they hop from site to site. In man's devices there is usually a separation of mechanical from electrical devices since the electrical device is electronic and the electron is the least potent of current-carrying particles in bringing about chemo-mechanical changes; contrast the electron with Ca^{2+}, H^+, and Mg^{2+} especially, but Na^+ and K^+ are more like the electron.

We shall use glycophorin as a model of the probable way in which the simplest mechanical devices work. The nature of glycophorin is shown in Figure 14. It is a membrane protein with a short helical stretch through the membrane and two ill-organized peptides both highly *charged* on either side of the membrane.[37] The helical stretch we shall assume to be a tensile connecting rod between the two pliable units on either side of the membrane. Any charged body or any

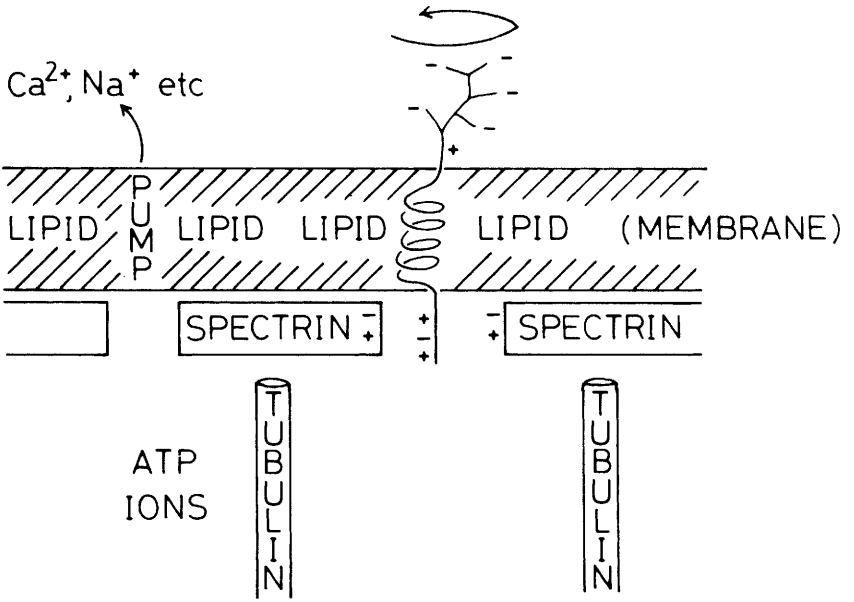


Figure 14 *The probable structure of the membrane protein glycophorin.*[37] *Two random sections outside (glycoprotein) and inside the cell are joined by a helix running through the membrane*

object which comes close to the external side of the glycophorin must cause the disordered polymer to alter the number of conformational states open to it. Under the applied force this region initially stores energy, $\partial(\Delta S)$, and then transmits force through the helix to the second region of low order inside the membrane. Here, a second redistribution of configurations, $\partial(\Delta S)$, occurs so that energy is now stored inside the membrane from effects outside. This energy contributes directly to a message relayed to the protein underlying the membrane and to any structural fibres of the cell. In turn the structured fibres connect (see below) through a mechanical muscle-like device which switches on enzymes, *e.g.* ATP-hydrolyses, causing amplified contraction of the fibre and adjustment of the membrane in the region where external force has been felt. Such a device depends on a combination of rigid rods and pliant elements, and it is not open to a simple structural description since the energy of transmission is locked in entropic as well as enthalpic functions. Recovery on removal of the external force is signalled by relaxation of the external disordered section of glycophorin, which was storing energy after the initial force was applied as in a rubber, and there follows a reversal of the signalling system and a return to the steady state at rest. This device is not unlike the muscle or other contractile units in principle, though the details of action are different. It is a continuously variable sensor with a continuously variable response.

As a second example of the combination of tensile and pliant materials in biological machines we consider the proteins of the muscle. The major units are myosin, actin, tropomyosin, and troponin. Table 9 gives their amino-acid compositions, and Figure 13 gives an artistic impression of their relationship to their functions, for which the sliding filament hypothesis is used to describe the muscle action. The model requires that binding of calcium (the muscle trigger), to troponin or light-chain myosin, should allow enzymic hydrolysis of ATP (energy source) in the myosin headgroup region to reorganize the interactions of the myosin headgroups with tropomyosin so that the filaments of one polymer, myosin, slide over and intermesh with the filaments of the other, actin. Given the amino-acid composition (Table 9) and knowledge of some structural features,

Table 9 *Amino-acid composition of muscle proteins*

Protein	*% Amino-acid composition*		
	Hydrophobic	*Charged*	*Ratio*
Troponin T	37	50	0.75
Troponin I	39	42	0.93
Tropomyosin	39	47	0.82
Troponin C	41	37	1.1
Myosin (light chain)	54	30	1.8
(heavy chain)	43	30—40	~1.4
Actin	54	23	2.3

Note In troponin C there is a highly charged site for binding to troponin I. The sequence is Gly.Lys.Ser.Glu.Leu. See J. S. Evans and B. A. Levine, *J. Inorg. Biochem.*, 1980 (accepted)

e.g. that tropomyosin and myosin are multiple helices, we may suppose that the tensile parts of this machine are formed from the actin, tropomyosin, and myosin, but not its headgroup, and that the moving pliant parts rest in the myosin headgroup, in the troponin, and in parts of tropomyosin. Our n.m.r. data show already that indeed the units of troponin, all three proteins, are flexible and that their interactions cause considerable geometric changes. Moreover, they are all dependent on the presence or absence of calcium.

We see that indeed these biological machines are composed of the same categories of units as man's machines. We believe that the description can be extended to all biological apparatus.

The Electric Devices. Current flow in semi-conductors (hopping transmission) and electrolytic conductors requires both structured and mobile devices. In place of batteries, wires, and resistors the major biological currents require pumps, aqueous solutions, and gates. In place of condensers there can be buffered regions of space in vesicles but also due to binding proteins. A comparison diagram between man's and biology's electric circuits is shown in

Figure 15. (In biological membranes electron carriers operate in the same way as in man's circuits, but in the aqueous phases the carriers are quite different.) In

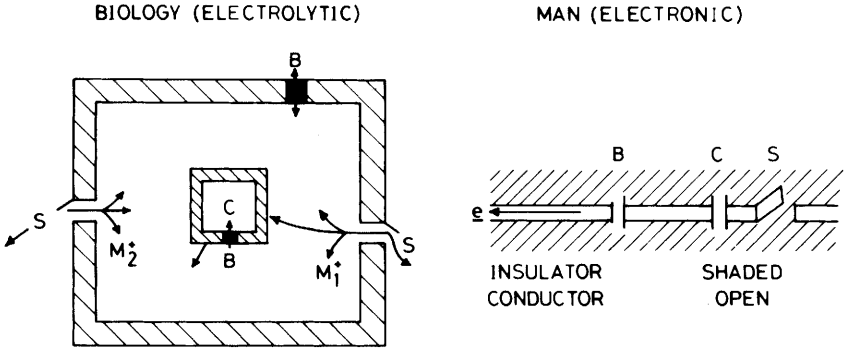


Figure 15 *A comparison between man's electronic circuits which are similar to biological electron-transfer circuits in membranes (right) and particles and the electrolytic circuits of biology (left). Note the general conducting space on the left is used by at least four current carriers, and gating devices allow all to act independently*

this article we are interested in the way in which proteins can or could act, not just as ion buffers but in the construction of the major devices which either create the required ionic concentration gradients or regulate the downhill flow of ions in these circuits. We note that the constraints on ion movement are not all mechanical. Just as in electronic circuits, there can be voltage control of downhill flow. It is usual to consider that Na^+/K^+ flows in nerve membranes have a voltage control, and Edmonds[38] has shown how this need not demand much protein conformation change. Flow is then in rigid-channel membrane proteins (see above). However, pumping of ions or the conversion of the electrical energy to chemical bond energy (ATP) does require a mechanical transduction. What are these mechanical devices? We can base a working hypothesis (see Figure 11) on our knowledge of the enzymes, kinases, that transfer phosphate since the activities of pumping ions, making ATP from ion gradients (especially H^+), and moving phosphate from ATP to another chemical in kinases are all related by the following equation

$$ATP + X \rightleftharpoons ADP + X\text{–}P$$

where, in a kinase, X is a substrate and, in an ion pump or the making or hydrolysing of ATP, X is water. It does not matter whether there is a phosphorylated protein as an intermediate in any or all of these reactions – they have in common phosphate transfers. The general suggestion is, therefore, that as a consequence of the change in structure of the pyrophosphate group of ATP to ADP plus P there is engendered a hinged movement in the protein (Figure 10).

The hinged movement can be driven by the ion gradient to make ATP or it can be a way of driving an ion up a gradient of chemical energy. The exact description of this type of mechanical activity must await detailed studies.

B. Polymers other than Proteins.—We stressed in the introduction that the polymers of sugars, nucleotide bases, and lipids ($\rangle CH_2$) must also have mobility. The major role of sugars on the surfaces of animal cells may well be in the rhetropic recognition, *i.e.* they do not have much structure (see glycophorin above). Plant saccharides are often structural. Table 8 brings out the distinction in charge type which is very like the distinction between proteins. Of course, chemical crosslinking blurs the distinction between mobile and rigid as it does in rubbers. The mobility of lipids is obvious, but a lipid membrane is a highly constrained liquid. It transmits force and is a visco-elastic material. Like motile proteins and sugars it responds continuously, it is rhetropic. Now, putting together the flexible proteins, sugars, and lipids in the complete membrane there is generated the locally, highly, and continuously responsive sensing of the external environment of the cell which is connected to the highly responsive control proteins (mechanical and electrolytic) of the interior. These mechanical and electrolytic devices connect to the enzymic machinery and not only through allosteric switches but by infinitesimally graded response locally. The whole structure and activity become a network of responses. In higher organisms this has allowed structure to develop as a reflection of required function.

9 Larger-scale Motions

In Section 8A we showed how the inner tubular system of a cell could be connected to the outside environment, using glycophorin as a model connector protein. The back-action of tubulin could then cause changes in the membrane of the cell, *i.e.* change cell shape. In addition to changes brought about by mechanical action the unequal distribution of charge, ions, across the membrane interacts with charges in the membrane. Again, the membrane is asymmetric and binds charges differently on its two surfaces. There is then a deforming energy acting on the membranes, and an effect will develop which is dependent upon the local construction of the membrane. Thus, membranes, and therefore cells, have shapes which reflect the state of activity of the cell or of an organelle. The best known local effects are in organelles such as mitochondria, chloroplasts, and such membranes as the sarcoplasmic reticulum of muscle. In these cases the construction is of deep invaginations such that very thin aqueous spaces are created between two membranes. The exact shapes depend upon energization of the organelle. All these three membranes use or make ATP in energy transduction between chemical energy and ion gradients and all then must be expected to act locally from both our knowledge of their gross physical construction[44] and the local disposition of the proteins. I have developed these points elsewhere.[45,46]

[44] G. Adam and M. Delbrück 'Structural Chemistry of Molecular Biology', ed. A. Rich and N. Davidson, Freeman, San Francisco, 1968, p. 198.
[45] R. J. P. Williams, *Biochim. Biophys. Acta*, 1978, **505**, 1.
[46] R. J. P. Williams, *Biochim. Biophys. Acta*, 1975, **416**, 237.

Here I wish to discuss the machinery of the membrane.

The membrane is partly lipid and partly protein but is not a solid construction. We must not expect to describe it by structure alone for as in the case of proteins it undergoes fluctuations. In the membrane some of these fluctuations are gross since the lipid is far closer to a liquid that is the interior of many proteins. Although the protein constructions, such as the helices of channels, and the proteins which float on the membrane, such as cytochrome b_5, may be moderately rigid all such units move relative to one another and only a statistical description is correct. It is then the case that the shape of the membrane reflects its state at a given instance in time. Just as there are intercommunicating network particles of metabolism in the aqueous cytoplasm, and yet these particles are freely moving, so there are intercommunicating network particles in the membrane. In these membrane-bound particles we shall find swinging arms and so on. However, the restriction of diffusion to a two-dimensional organization[44] and the fact that we are dealing with a lipid phase means that tumbling motion is also significant. The exact motions which are important in the transport of material across membranes, small or large molecules, is quite unknown but already we know that special protein sequences are involved in the export of proteins through membranes. During the actual movement these sequences may very well undergo considerable rearrangement.

Biological systems have often been described at the molecular level as a set of on/off equilibria. My view is quite different in that many proteins, membranes, or assemblies are seen to have a continuum of states. The individual particles then respond to pressure (*i.e.* all forms of pressure including mechanical forces), radiation falling upon them, temperature, and chemical or electrical potential continuously. This means that states are characterized not by structures but by fluctuations between a large number of states, and small changes at one site lead to flow. Each cell has many circuits of flow of different potential carriers (electrons, ions, co-enzymes), and the structure of the cell gives these flows selected specified directions based on local pumping, gating, and chemical transformation by catalysts. In such a system there are no fixed constants such as equilibrium constants or solubility products as each such constant, defined relative to a conventional standard state, is a continuous function of variables:

$$RT\log K = \Delta G^0 + F_1(T, P) + F_2(\psi, [\mathrm{X}], etc.)$$

where ψ is any applied field, electrical or mechanical, and F_1 and F_2 are functions. Finally, since the cell accumulates its products and these alter the chemical potentials of all the components, when they alter many of the above functions, the values of $RT\log K$ are functions of time (see Part I). Of necessity the shapes of organisms, cells, and even individual macromolecules become a function of time. It is the dynamic nature of polymers which allows this possibility.

10 Summary

In this article I hope that I have indicated how the proteins control and are controlled by the small molecules and ions of a cell. The whole network of inter-

actions is *continuously* monitored and adjusted by internal and external stresses. Since, in addition, development of the cell continuously alters the protein and other polymer complements of the cell, the internal stresses are open to programming. This programming is seen in extreme cases as a switch to a new activity, differentiation, or to cell division. A major task confronts us. It is to learn from the study of fragments the potential functions of the units, but also to study whole organelles and cells to uncover the flow of energy and material, currents, to which the cell's activities are tuned. It is the ability to organize the units locally which will make the circuits so effective in keeping the cell in communication with its environment (including other cells). Organisms higher than unicellular do not use new principles but just larger numbers of divisions of space by membranes.