The role of mobility in the substrate binding and catalytic machinery of enzymes

TOM ALBER*, WILLIAM A. GILBERT†, DAGMAR RINGE PONZI† and GREGORY A. PETSKO†

*Departments of *Biology and †Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

Abstract Recent theoretical and experimental studies have demonstrated that proteins are fluctuating systems capable of large, seemingly random, excursions from the equilibrium conformation. Attention is now focusing on the functional consequences of these motions. X-ray diffraction is a powerful tool for mapping the spatial distribution of protein dynamics; studies on the temperature dependence of the apparent Debye-Waller factors of crystalline myoglobin demonstrate that proteins are flexible in the solid state. Crystallographic studies of a Michaelis complex of ribonuclease A show that a mobile lysine adapts its conformation to the changes in stereochemistry and charge distribution in the substrate during catalysis. The structure of the triose phosphate isomerase–substrate complex shows that a mobile region of 10 amino acids becomes ordered when ligand binds. These studies suggest several roles for protein mobility in enzymic catalysis: providing access to internal sites, allowing changes in substrate structure during the reaction, and reducing the observed binding constant of substrate and product to the enzyme by decreasing entropy. A flexible enzyme also does not need a communication system to signal binding or transformation, since a pre-existing equilibrium can be used. More speculative ideas, such as the guiding of thermal vibrations along the reaction coordinate, can only be tested when more detailed data are available.

Enzymes as machines

The earliest concepts of enzyme structure viewed these biological catalysts as colloids without a defined conformation. Since that time, our view of these molecules has oscillated from one of great flexibility to rigidity and back again. In the last five years, virtually the entire arsenal of biophysical methods has been trained on the problem of protein dynamics. We now understand

1982 Mobility and function in proteins and nucleic acids. Pitman, London (Ciba Foundation symposium 93) p 4-24

that, in aqueous solution at ordinary temperatures, atoms in globular proteins can undergo a wide variety of motions, ranging in amplitudes from a few hundredths of an Ångstrom to several Ångstroms and in frequencies from less than one picosecond to seconds (Gurd & Rothgeb 1979). Attention is now beginning to focus on the functional consequences of the random fluctuations in protein structure, on the not unreasonable assumption that nature usually turns the inevitable to her advantage.

We are interested in the structural basis for enzymic catalysis. Some enzymes speed up the rate of their chemical reactions by factors of 10^{10} or more. Simple organic and inorganic catalysts cannot approach this efficiency. Since one characteristic these small molecule catalysts lack that enzymes possess is conformational flexibility, we have tried to identify what chemical advantages a dynamic structure might confer on a protein.

The concept of a changing structure for an enzyme is hardly revolutionary. Indeed, our conventional way of viewing biological catalysis is by analogy to machinery. Like the elaborate cause-and-effect contraptions created by the American cartoonist Rube Goldberg and his British counterpart Heath Robinson in the 1920s, enzymes are thought of as moving logically from one well-defined conformation to another in response to a set of signals. These signals originate with the binding of substrate to the enzyme or its conversion to an intermediate, and are propagated to distant parts of the protein by a linked series of small changes: a hydrogen bond breaks, leading to a shift in a helix–helix contact, which in turn, etc. (the pre-eminent example of this view is the transformation from the R to T state of haemoglobin).

The validity of this picture is long established. But this picture is not quite what we want to discuss here. The Rube Goldberg/Heath Robinson picture of conformational changes in a protein is a description of transitions from one equilibrium state to another. We are concerned, in this paper, with the fluctuations of the enzyme about any given equilibrium state, and the role of these thermal energy-driven motions in the catalytic process. Such movements are random and difficult to observe, but ubiquitous. And we have found, by examination of their traces in the crystal structures of two enzyme–substrate complexes, that they may be extremely important for efficient catalysis.

Proteins are flexible in the crystalline state

Since we shall rely on the results of X-ray diffraction to map the spatial distribution of protein fluctuations, it is essential to establish that these small random motions really occur in the crystalline state and can be measured. There is considerable indirect evidence that proteins maintain their flexibility

even on crystallization: protein crystals are typically 50% solvent by volume (Matthews 1968), so there is room in the lattice for substantial mobility; buried amide hydrogens in crystalline proteins exchange with solvent deuterons at rates comparable with their exchange rates in solution, implying that whatever motions admit water to the protein interior or expose the inner structure to solvent occur in the solid state (Tüchsen & Ottesen 1979), and there is considerable fall-off in the electron density distribution around each atom in a protein crystal structure, presumably from thermal motion of those atoms (Artymiuk et al 1979, Frauenfelder et al 1979). This last point is crucial to our use of diffraction as a tool to study protein motion, so we want to discuss it in some detail.

The fall-off of electron density around each atom can be fitted to a Gaussian curve whose width at half-height, the apparent mean-square displacement or $\langle x^2 \rangle$ of that atom, is related to the conventional Debye-Waller factor, B, of that atom by the expression $B = 8\pi^2 \langle x^2 \rangle$ (Willis & Pryor 1975). If high resolution X-ray data (i.e. 0.2 nm or better) have been collected for a structure, the apparent B for each atom can be refined in a restrained least-squares method, as described by J. H. Konnert and W. A. Hendrickson (Konnert 1976). Then the apparent $\langle x^2 \rangle$ for each atom can be calculated. This method requires two incorrect assumptions, the seriousness of which is unknown: first, that the motion is isotropic, and second, that it is harmonic. Interpretation of the $\langle x^2 \rangle$ values in terms of actual motion requires the additional assumption that the observed fall-off in electron density from the average position of each atom is really due largely to atomic motion. Unfortunately, it could be dominated by a build-up of errors in the structure determination and/or by static disorder in the crystal lattice. We do not believe this is the case, since the observed $\langle x^2 \rangle$ values correlate sensibly with structural features in the protein (Frauenfelder & Petsko 1980), but direct evidence is desirable.

Atomic motion should be temperature-dependent; lattice disorder should not be. Further, the magnitude of any temperature dependence should enable pure harmonic vibrations (small and linear dependence on T) to be differentiated from larger, more complex motions that require a potential energy barrier to be overcome (large and complex T dependence). In collaboration with Professor Fritz Parak, Drs W. Steigeman and H. Hartmann in Munich, and Professor H. Frauenfelder in Illinois, we have refined the crystal structure of sperm whale Met-myoglobin at 0.2 nm resolution at a temperature of 80 K (Hartmann et al 1982). Fig. 1 shows a comparison of the average $\langle x^2 \rangle$ of the backbone atoms of myoglobin as a function of residue number at 80 K with the values obtained earlier at room temperature (Frauenfelder et al 1979). The overall $\langle x^2 \rangle$ is reduced by more than a factor of two, and individual residues show reductions even larger than this. The apparent $\langle x^2 \rangle$

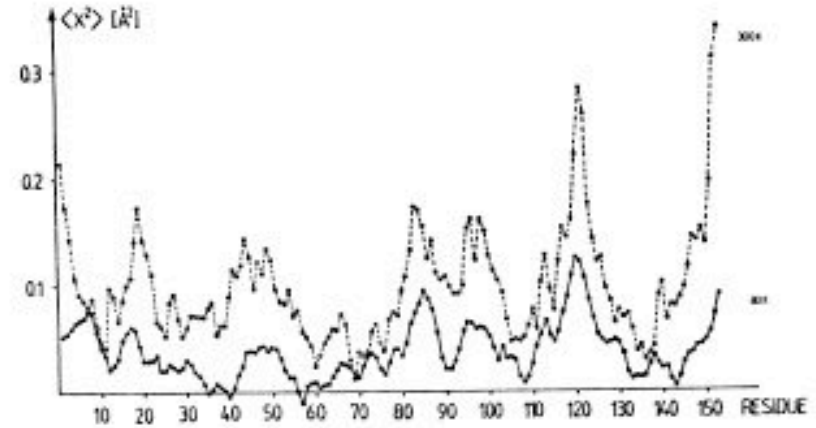

FIG. 1. Plot of the average mean-square displacement of the N,C^{α}, and carbonyl carbon atoms of myoglobin against the residue number at 300 K (open circles; broken lines) and 80 K (solid lines and filled circles). Both curves have been corrected for the same estimated lattice disorder (Frauenfelder et al 1979).

values for myoglobin are predominantly the result of atomic motion. The temperature dependence also shows that many residues in myoglobin have motions that are more complex than simple Debye-Waller vibrations. Even in the crystalline state, proteins are dynamic structures.

The problem of ligand access

Fig. 1 suggests one important functional role for the fluctuations in the atoms of myoglobin. If a Gaussian model is assumed for isotropic atomic motion, the observed $\langle x^2 \rangle$ values imply that some regions of the protein have a reasonable probability of undergoing displacements of 0.1 to 0.2 nm in amplitude. A series of such motions could open a transient channel from the surface of myoglobin to its interior. Myoglobin is an oxygen storage protein; the oxygen binds to an iron atom buried inside the protein in a hydrophobic pocket. Yet there is no pathway in the static crystal structure of myoglobin from outside the protein to the iron that is large enough to admit oxygen. The random fluctuations in the myoglobin structure provide access to the shielded ligand-binding site.

Many enzymes may face a similar problem. Chemical considerations could require a buried active site, with restricted accessibility to bulk solvent. Yet substrate must be able to enter. For small substrates, entry can be achieved by

small movements of the protein which allow penetration from outside the molecular envelope.

Interpretation of hydrogen exchange data is controversial, but it seems clear that small-scale fluctuations in the protein structure are also essential to permit ions derived from solvent water to make contact with buried protein hydrogens. We often forget that water is an essential reactant in many enzyme-catalysed reactions. Fluctuations in the positions of atoms in the enzyme could allow water to reach even relatively hydrophobic pockets where the reaction takes place. Thus, because enzymes are flexible, they can provide a highly apolar environment for a particular group on the substrate as well as allowing water to reach and react with an adjacent group. No rigid catalyst could satisfy these conflicting conditions. One role for protein mobility in catalysis is clear, then: to permit access to seemingly buried regions of an enzyme.

The problem of response to changing substrate structure

Enzyme–substrate complexes, even in the crystalline state, normally have lifetimes of one second or less. Unfortunately, it takes several days or longer to collect a set of X-ray diffraction data on an enzyme crystal. This limitation has prevented direct observation of productive enzyme–substrate complexes by crystallography. We have overcome this limitation by combining the techniques of cryoenzymology with protein crystal structure determination (Fink & Petsko 1981). We cool enzyme crystals to very low temperatures in a fluid cryosolvent and then let substrate flow into the crystal. Binding still occurs at subzero temperatures, but turnover is extremely slow. The 'frozen' Michaelis complex is often stable for weeks, and its structure may be determined at high resolution by standard crystallographic methods.

Recently, we have applied X-ray cryoenzymology to the reaction catalysed by ribonuclease A and have determined the structures, at atomic resolution, of every kinetically significant step in the reaction pathway (W. A. Gilbert & G. A. Petsko, 1982, unpublished). Ribonuclease A is interesting because its substrate undergoes a number of large structural rearrangements during catalysis. The enzyme hydrolyses single-stranded RNA and is specific for pyrimidine bases on the 3′ side of the phosphodiester linkage to be split. The phosphate group is initially in a 3′, 5′ bridge, is converted to a pentacoordinate intermediate (or transition state—its lifetime is uncertain) and then to a 2′, 3′-cyclic monophosphate. Finally, via another trigonal bipyramidal intermediate, the cyclic phosphate intermediate is hydrolysed to a 3′-monophosphate ester. During these transformations the phosphate atom moves by 0.2 nm from its initial position in the enzyme–RNA complex. The

oxygen atoms around it change position, bond angles, and charge. How does the enzyme respond to the alterations in substrate geometry?

Some of the active site residues simply move along with the substrate atoms to which they are bonded, but there is one interesting exception. Lysine 41 has long been implicated in the ribonuclease reaction: chemical derivatization of this residue destroys catalytic activity. Yet Lys 41 is not directly bonded to the substrate in the enzyme–substrate, enzyme–phosphate or enzyme–cyclic phosphate complexes. However, in the crystal structure of the complex of ribonuclease A with uridine vanadate, a transition-state analogue which is trigonal bipyramidal in geometry, Lys 41 has moved so as to interact with one of the equatorial trigonal oxygens (Fig. 2).

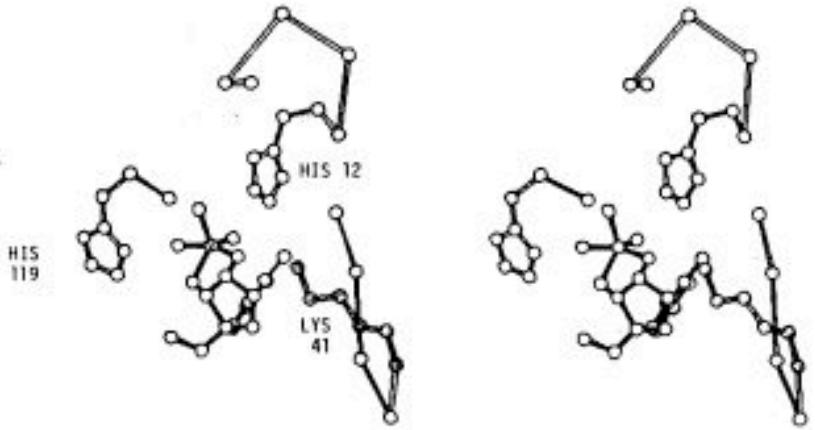

FIG. 2. Stereo drawing of some of the residues in the active site of ribonuclease A when the transition-state analogue uridine vanadate (shown in black) is bound. The vanadium (V) is at the centre of a distorted trigonal bipyramid of oxygens. Lys 41 (shaded grey) is rigid in this complex, presumably due to interaction with the rightmost equatorial oxygen.

Presumably this interaction also occurs in the transition state of the reaction and stabilizes the negative charge that must develop on that oxygen. But why does Lys 41 not interact with the oxygen all the time?

The answer lies in the large changes in substrate structure referred to earlier. If Lys 41 were always in the position it assumes in the transition state complex, that complex could not form. The movements of the phosphorus and oxygen atoms would be blocked sterically by the terminal amino group of the lysine. Ribonuclease solves the difficulty of keeping Lys 41 out of the way but bringing it into play when needed by making it a flexible residue. In all the structures we have examined except the uridine vanadate complex structure,

Lys 41 has very high $<x^2>$ values for its side-chain atoms, among the highest in the entire enzyme. Further, the temperature dependence of these $<x^2>$ values indicates that they are due to real thermal motion, not static disorder. But in the structure of the transition-state analogue complex, ribonuclease–uridine vanadate, the $<x^2>$ values for Lys 41 decrease threefold. The highly mobile side-chain is now anchored (temporarily, for in the actual transition state the pentacoordinate species is unstable and rapidly hydrolyses) by interactions with the substrate.

The fluctuations undergone by Lys 41 are important for catalysis. They allow the enzyme to adapt to the changes in stereochemistry and charge distribution that take place in the substrate. A rigid catalyst would not be able to bind both a tetrahedral substrate (and product) and a trigonal bipyramidal intermediate as efficiently, or would prevent the rearrangement from one to the other sterically. What is important to note here is that the enzyme does not move from one static conformation to another as the substrate is converted to product. It is constantly fluctuating among a number of conformations, and the required one is stabilized when needed. Thus, there is no energy barrier (or at least, no large one) to be overcome when the enzyme goes from the native (or substrate-bound) active site structure to the transition-state structure.

When tight binding is not so tight

Recently, a number of enzymes have been observed to contain regions of poorly defined three-dimensional structure (Huber 1979). These disordered segments may range in size from a few amino acids to an entire structural domain, and may occur either at the termini of the polypeptide chain or in the middle. We have been fortunate in having one of these to study, since they represent extreme examples of conformational mobility. Our specimen is in the yeast glycolytic enzyme, triose phosphate isomerase (TIM).

TIM is a dimeric enzyme with a subunit relative molecular mass of 27 000. It catalyses an extremely simple reaction: the movement of one proton to convert dihydroxyacetone phosphate (DHAP) to D-glyceraldehyde 3-phosphate. The phosphate group is essential for substrate binding. The crystal structure of the chicken muscle enzyme has been solved by Phillips and his co-workers at Oxford (Banner et al 1975). We have concentrated on the yeast enzyme, but *all* our results reported here were observed initially at lower resolution with chicken TIM (Phillips et al 1977). The results are thus independent of crystal form.

Native TIM is a symmetrical dimer. Each subunit is composed of a central core of eight strands of parallel β-sheet, twisted around the surface of a barrel. Each strand is connected to the next by an α-helix (Fig. 3). We have solved the crystal structure of yeast TIM at 0.3 nm resolution (Alber et al 1981) and also sequenced the gene for this protein (Alber & Kawasaki 1982). In the native structure there is no contiguous electron density for a 10-amino acid loop from residues 168 to 177. This loop is the only stretch of more than five amino acids that is not involved in either a helix or strand of sheet (except for the intersubunit loop that holds the dimer together).

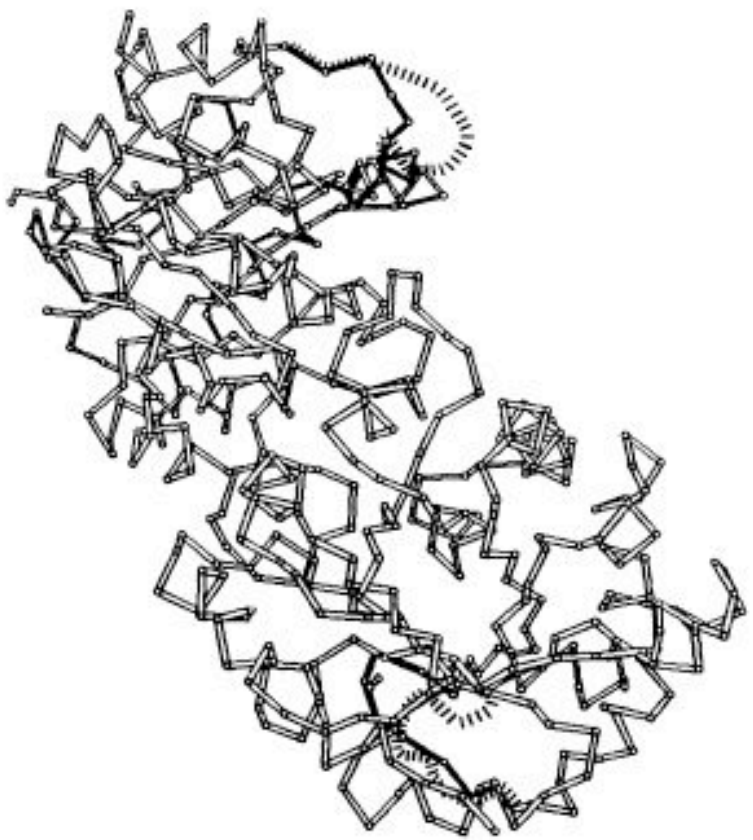

FIG. 3. Alpha carbon drawing of the yeast triose phosphate isomerase dimer. The lower subunit is oriented so that the viewer is looking down the central β-barrel. Helices surround this core of pleated sheet. The two average conformations of the loop from residues 168 to 177 are highlighted. The 'open' form in the native enzyme is disordered, so its position is indicated by broken lines. The rigid, 'closed' conformation is shown with solid black bonds.

When the structure of the enzyme–substrate complex TIM–DHAP is examined at 0.35 nm resolution, a large change is seen in the average position of the residues of this loop (Alber et al 1981). In the native structure it is 'open', extended away from the molecular surface (Fig. 3). On DHAP binding it moves by over 1 nm to fold down into a 'closed' conformation onto the active site pocket (Figs. 3, 4). The disordered loop becomes highly ordered, with strong electron density for all residues. At least two of the main chain amides in this loop make hydrogen bonds to two of the oxygens of the phosphate moiety of DHAP.

Since this conformational change is seen in both orthorhombic chicken TIM crystals and monoclinic yeast TIM crystals, it is not an artifact of crystal packing adjustments. Since these studies have been performed with the physiological substrate DHAP, the loop movement is not an artifact of incorrect binding. Since the subunit where the loop is held in an 'open' conformation by intermolecular contacts in the chicken TIM crystals does not bind substrate, even though the active site pocket is accessible, the loop movement is essential for both substrate binding and catalysis (Phillips et al 1977). All these conclusions are reinforced by the observation that the amino acid sequence of the loop is highly conserved (Alber & Kawasaki 1982).

The folding of the 168–177 loop over the phosphate end of the substrate (Fig. 4) has several obvious consequences for catalysis. First, it provides an important source of specificity through an 'induced-fit'. Only substrates capable of anchoring the loop in the 'closed' conformation will be bound

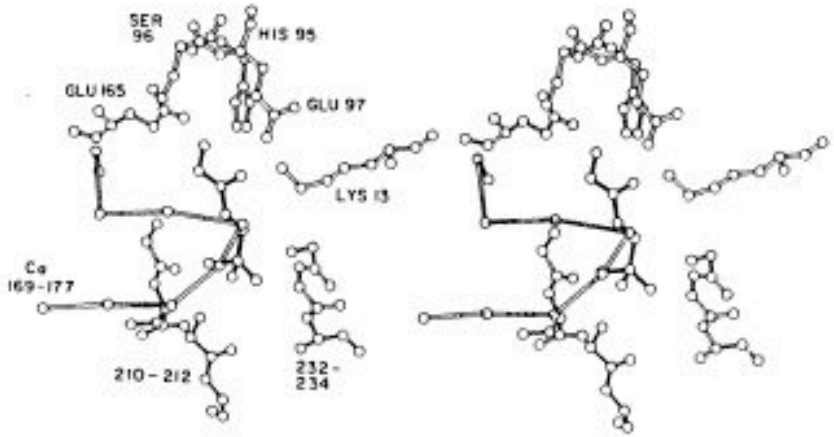

FIG. 4. Stereo drawing of the active site of the yeast triose phosphate isomerase–dihydroxyacetone phosphate Michaelis complex. The substrate is shown with solid black bonds, and a portion of the loop 168–177 is highlighted by grey shading. Only the C^{α} atoms of the loop are shown, as the substrate would be obscured if all the atoms were included.

tightly. Second, the folding of the loop sterically excludes solvent from part of the active site. Of course, a readily apparent but essential function of the flexible loop is to coordinate the phosphate group. At least one of the terminal oxygens of the phosphate, and possibly the bridging oxygen as well, is hydrogen-bonded by peptide nitrogens from the loop in our current model. These various roles demand a large conformational change in the loop. The conflicting requirements for ready substrate access to the active site and envelopment of the substrate when it is actually bound can easily be met by a 'flap' which goes from an open to a closed conformation on complexation. Additional support for this loop movement occurring in TIM-catalysed isomerization in solution comes from the observation (Albery & Knowles 1977) that a physical step connected with product (glyceraldehyde 3-phosphate; GAP) release—presumably, the reopening of the loop to allow GAP to escape—is rate-limiting.

Thus, the conformational change in the 168–177 segment in the TIM structure is mechanistically reasonable, but there is an additional aspect that deserves discussion. The loop does not merely go from a unique open structure to a unique closed one; it changes from a flexible region with many conformations to a much less flexible one. We believe this disorder–order transition has at least two consequences for the energetics, and consequently the rate, of the TIM reaction. Because of the structural and thermodynamic problems of large conformational changes, a flexible to less flexible transformation has a number of advantages so great that it has been predicted that this pattern will appear often in protein–ligand interactions (Alber 1981).

First, the disorder in the loop eliminates the problem of how information concerning substrate in the active site is transmitted to the loop. Initially, the average position of the phosphate-binding atoms in the loop is at least 0.8–1.0 nm from where they will need to be to coordinate the phosphate oxygens. If the loop were rigid, a signal would have to be sent from the active site to the loop indicating that substrate had entered the pocket, and the signal would have to trigger a large rearrangement in the loop. Since the loop makes no contacts with the rest of the protein, it is difficult to imagine how this could be done. It is even more difficult to understand why the loop would then reopen to allow release of product. The TIM reaction is readily reversible, and *in vivo* the enzyme may encounter either substrate. Further, the two substrates do not differ very much in stereochemistry, and not at all in charge. (The binding question can be partly answered by postulating that the substrate drags the loops with it as it approaches the active site. However, this does not resolve the reopening question.)

All these difficulties vanish if the loop is initially flexible. No signal need be sent to the loop to cause it to move: it already is moving. There is a pre-existing conformational equilibrium, whose average position is many

Ångstroms from the active site but one of the extreme positions of which is at or near the desired 'closed' conformation. Driven by the kinetic energy available to it at ordinary temperatures, the loop samples a number of conformations. At least one of these is at or near the active closed state. If a substrate molecule is present in the active site when the loop finds itself at this extreme point, the loop can interact with the phosphate, binding and desolvating the substrate and shifting the equilibrium from one favouring the open state to one favouring the closed state. Product release is easy to visualize in this model: although the conformational equilibrium has been shifted to favour the substrate-stabilized structure, it is still an equilibrium. There is still a finite probability that a thermal fluctuation will break the loop–phosphate bond and the loop will open. When it does, product can escape. But the probability that this rare event will happen is greater than it would be if the open form were a unique, rigid conformation. The transition from closed to open state is an order-to-disorder transition, which is favourable entropically. One way of viewing this is that a random thermally driven fluctuation is not as likely to carry the loop from one discrete state to another of greatly different structure as it is to move the loop along a continuum of structures, each differing only slightly from the next in atomic positions and energy.

The suggestion of entropy as a driving force for the opening of the loop, permitting release of product, leads to a consideration of the thermodynamic role of a flexible region of a protein in catalysis. Alber (1981) has discussed this in detail; his arguments apply to the binding of any molecules to proteins, but here we shall consider only TIM catalysis.

First, the activation energy for binding is minimized, because the interactions between substrate and enzyme do not require the disruption of a stable structure in the protein. Second, if the reduction in the number of conformational states of the loop on substrate binding is sufficiently large, some of the binding energy will be dissipated by the resulting decrease in the conformational entropy of the polypeptide chain. The quantitative limits of this effect have been considered by Alber (1981); they are sufficient to increase the apparent binding energy for substrates by several kilocalories per mole (Fig. 5). Thus, the disorder-to-order transition raises the free energy of the enzyme–substrate complex relative to the free energy of the transition state for its formation. Analogously, the enzyme–phosphate complex has a more positive binding energy than it would otherwise have, allowing more rapid release of product. Fig. 5 shows the effect of a disorder-to-order transition on the free energy profile for the TIM reaction (Albery & Knowles 1977). The entropic loss contributed by the loop makes both substrate and product less tightly bound than the sum of their interactions with the enzyme would suggest. Since efficient catalysis requires tight binding of the transition state

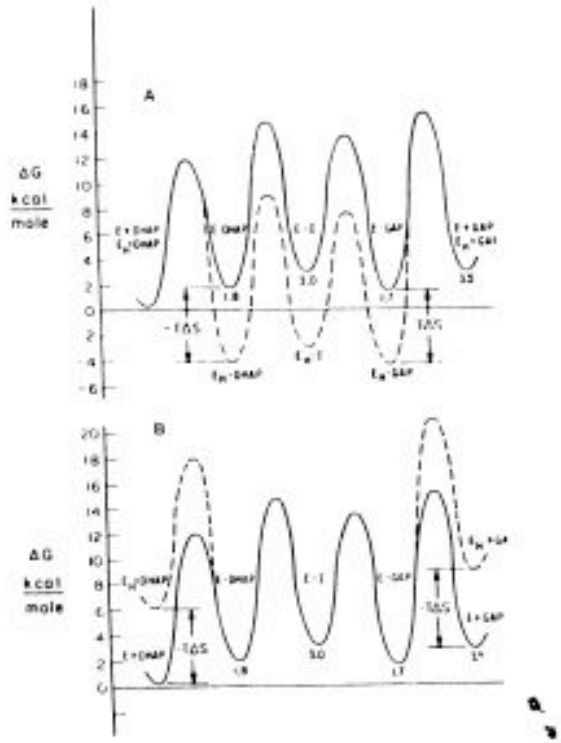

FIG. 5. Free-energy profile for the reaction catalysed by triose phosphate isomerase (TIM). The solid line shows the free-energy profile derived from the kinetic and thermodynamic results of Albery & Knowles (1977). The broken line shows the profile for a hypothetical enzyme (E_H) which is identical to TIM except that the 168–177 loop is ordered before its movement to interact with the substrate. The value of $-T\Delta S$ indicates the extent to which the enzyme–substrate (E–S) complex is destabilized by the ordering of the loop. Assuming, for example, that 6 kcal/mol (−25 kJ/mol) of binding energy is dissipated to order the loop, at 37°C the natural enzyme (E) is over 10^4 times faster than the hypothetical one. (A) Free-energy profiles assuming no change in the standard state. In this case, the amino acid sequence in the loop of E_H is assumed to be different from that of its real counterpart. (B) Free-energy profiles assuming that the amino acid sequences of E and E_H are the same and that the energy of E_H is raised by the loop being in only one conformation before substrate binding. DHAP, dihydroxyacetone phosphate; GAP, glyceraldehyde 3-phosphate; E-I, intermediate enzyme. (From Alber 1981.)

but not too tight binding of substrate, product or intermediates, this is a simple mechanism for combining specificity with optimal rate acceleration.

Directed fluctuations: fact or fantasy?

For some time we have been considering an additional role that fluctuations in an enzyme might play in assisting catalysis. We have been unable to devise an experiment to test this supposition, but feel it deserves some mention. We call it the directed fluctuations hypothesis.

Consider the free-energy profile for an enzyme-catalysed reaction, as in Fig. 5. Enzymes lower the free energy of the transition state, but they do not reduce it to zero. Energy is still required to overcome the remaining barrier and proceed along the reaction coordinate to product. Presumably this energy is just the kinetic energy available to the enzyme from collisions with the solvent at the temperature of interest. It is normally assumed that this energy is absorbed and distributed randomly in the enzyme–substrate complex. What if it is not random?

In order for a reaction to go from substrate to product, certain bonds must break and others must form. The transition state consists of partially formed and broken bonds. At any given point in the reaction there are thermally driven fluctuations in bond lengths and angles. If these vibrations coincide with the reaction coordinate, that is if they move atoms in the directions they must move during the reaction, catalysis will proceed. Our hypothesis is that these fluctuations may not be random, and that they could tend to favour directions along the reaction coordinate. They are 'directed'—by the protein structure—to be useful.

Conceptually, we view this effect in terms of rate enhancement. At a given temperature, there is some finite probability that enough energy will be absorbed in vibrational modes to allow the substrate to cross the transition-state energy barrier. If the three-dimensional structure of the protein favours (energetically) vibrations in directions which lie along the reaction coordinate, the observed rate of crossing this barrier will be faster than that expected when completely random bond vibrations and librations are assumed.

Do directed fluctuations occur? We do not know. If anisotropic $<x^2>$ values could be calculated for the atoms in, say, the ribonuclease A–substrate complex, it might be possible to observe a non-random distribution of the largest moments along a productive direction. This calculation requires more X-ray data than we have measured at present, but we are trying to obtain them. Alternatively, we may find directional trends in the amplitudes of picosecond motions calculated in a molecular dynamics simulation of the reaction. (Calculations of this kind are just beginning.) But at present there is no evidence for directed fluctuations, or against them.

Conclusion: enzyme mobility is important for enzyme function

We began this paper with a discussion of enzymes as machines. It is clear that this view is too simplistic: the machines of Rube Goldberg and Heath Robinson moved from one rigid state to another in response to carefully regulated stimuli, but enzymes fluctuate continuously among families of closely related states. They may still change their average conformations, but these conformations are indeed just averages. If enzymes are machines, they are stochastic machines.

Analysis of the structural data accumulated in our laboratory suggests that the mobility of enzymes is of functional significance. The fluctuations of atoms in the protein can provide access to buried sites, can allow the catalyst to respond to changes in the structure of the substrate, and can, by a disorder-to-order transition, raise the apparent free energy of the enzyme–substrate complex. We do not know whether they also are spatially distributed in such a way as to favour the formation and breakdown of the transition state.

Finally, we conclude that flexible enzymes need no signalling system or linked series of bond rearrangements to produce a desired effect (e.g. movement of the 168–177 loop in TIM or the movement of Lys 41 in ribonuclease A) from a given cause (e.g. the binding of substrate to the active site). This is not to imply that signalling does not exist; there are many proteins where it probably does. We merely wish to point out that it is not essential for all proteins. Indeed, there may be advantages in not using it. A continuum of structures of similar free energy, from which the desired conformation is stabilized when needed, requires no overcoming of a large potential energy barrier for the generation of the desired conformation. In this view, the change from one functional average structure to another is not really separate from the fluctuations about the observed averages. A conformational sub-state at one extreme of the distribution of excursions can become the new average structure if it is stabilized by, say, interactions with the substrate (or with DNA, or with another protein). Enzyme mobility and enzyme function are two horns on the same bull.

REFERENCES

Alber T 1981 Structural origins of the catalytic power of triose phosphate isomerase. PhD thesis, Massachusetts Institute of Technology

Alber T, Kawasaki G 1982 Nucleotide sequence of the triose phosphate isomerase gene of *Saccharomyces cerevisiae*. J Mol Appl Genet, in press

Alber T, Banner DW, Bloomer AC et al 1981 Structure and catalytic mechanism of triose phosphate isomerase. Phil Trans Roy Soc Lond B Biol Sci 293:159-171

Albery WJ, Knowles JR 1977 Perfection in enzyme catalysis: the energetics of triose phosphate isomerase. Acc Chem Res 10:105–111

Artymiuk PJ, Blake CCF, Grace DEP, Oatley SJ, Phillips DC, Sternberg MJE 1979 Crystallographic studies of the dynamic properties of lysozyme. Nature (Lond) 280:563–568

Banner DW, Bloomer AC, Petsko GA, Phillips DC, Pogson CI, Wilson IA 1975 Structure of chicken muscle triose phosphate isomerase determined cystallographically at 2.5 Å resolution. Nature (Lond) 255:609–614

Fink AL, Petsko GA 1981 X-ray cryoenzymology. Adv Enzymol Relat Areas Mol Biol 52:177–246

Frauenfelder H, Petsko GA 1980 Structural dynamics of liganded myoglobin. Biophys J 19:465–483

Frauenfelder H, Petsko GA, Tsernoglou D 1979 Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. Nature (Lond) 280:558–563

Gurd FRN, Rothgeb M 1979 Motions in proteins. Adv Protein Chem 33:73–165

Hartmann H, Parak F, Steigemann W, Petsko GA, Ringe Ponzi D, Frauenfelder H 1982 Conformational substates in a protein: structure and dynamics of metmyoglobin at 80 K. Proc Natl Acad Sci USA 79: 4967–4971

Huber R 1979 Conformational flexibility and its functional significance in some protein molecules. Trends Biochem Sci 4:271–276

Konnert JH 1976 A restrained-parameter structure-factor least-squares refinement procedure for large asymmetric units. Acta Crystallogr Sect A Cryst Phys Diffr Theor Gen Crystallogr 32:614–617

Matthews BW 1968 Solvent content of protein crystals. J Mol Biol 33:491–497

Phillips DC, Rivers PS, Sternberg MJE, Thornton JM, Wilson IA 1977 An analysis of the three-dimensional structure of chicken triose phosphate isomerase. Biochem Soc Trans 5:642–647

Tuchsen E, Ottesen M 1979 A simple hydrogen exchange method for crosslinked protein crystals. Carlsberg Res Commun 44:1–10

Willis BTM, Pryor AW 1975 Thermal vibrations in crystallography. Cambridge University Press, Cambridge

DISCUSSION

Steitz: Why do you think myoglobin wasn't 'designed' in the first place *not* to have the access to haem iron blocked? This gives us a function for mobility, certainly, but what is the function of the blocking?

Petsko: One cannot have too much access to the oxygen-binding site, because free haem has a considerable tendency to form μ-oxo dimers. Thus the necessity for burying the haem within the protein is a chemical necessity. If dimers were formed, that would block the function of the haem molecule. But why it is so deeply buried is unclear.

Steitz: Is the haem iron buried in haemoglobin too?

Petsko: Yes; it is buried to about the same extent as in myoglobin.

Berendsen: I have difficulty in understanding your point about the facilitation of product release. You say that the ordering of a loop of the triose phosphate isomerase chain weakens the substrate binding, because of the

entropic effect, but there must be an equilibrium between a non-ordered and an ordered configuration. If the ordered one is preferred, it still means that the free energy of that configuration is lower, even if the entropy is lower. How would that weaken the binding? It would have to be compensated for by a more negative enthalpy.

Petsko: Yes. I didn't discuss compensation, which is clearly a difficult subject. Nor did I discuss the role of water in all these changes. These are complex aspects of the problem. One aspect is how the protein compensates for some of these changes in flexibility. For example, if lysine 41 moves, what happens to the residues with which it was previously in van der Waals contact? Do they become more disordered? From our structural data, it seems that they might. The triose phosphate isomerase reaction is relatively reversible and the enzyme 'sees' both substrates all the time, physiologically, so that what we think of as 'product release' is in fact, from the point of view of the enzyme, the binding of substrate. So I think my argument is still pertinent to the enzyme, in that for this or any other enzyme one wants specificity and multiple points of contact, without making the binding of substrate too tight.

Berendsen: A measure of the tightness of binding is given by the free energy; having a higher entropy will not change the fact that free energy has to be lower in order to get binding. I don't see how it helps, to make the binding *weaker*.

Frauenfelder: I can answer that question in a different context. Consider O_2 binding to myoglobin or haemoglobin, where the dissociation rate is given by $k_{\mathrm{dis}} \approx \nu \exp(S^*/R)\exp(-H^*/RT)$. Nature can do little to change H^*; H^* is mainly determined by the strength of the covalent Fe–O_2 bond and it is of the order of 80 kJ/mol. The term $\exp(-H^*/RT)$ then is about 10^{-14}. The constant ν lies between 10^{11} and $10^{12}\ s^{-1}$ so that $\nu \exp(-H^*/RT) \approx 10^{-3}$ to $10^{-2}\ s^{-1}$. If $S^*=0$, oxygen remains bound for a long time and is not available where needed! Indeed, we know that the entropy decreases considerably on binding of O_2 to Hb and Mb; it must therefore increase in dissociation and the factor $\exp(S^*/R)$ makes life possible. A major contribution to S^* may come from protein flexibility.

Petsko: We are talking in a sense about the difference between an *intrinsic* binding energy and the observed binding parameters. This is where the confusion arises. It is true that the free energy of the bound species must be lower than that of the free entities, but what I am saying is that the free energy of the bound species is not as low as it should intrinsically be in view of the type and number of bonds made. It is not as low because some of that intrinsic binding energy is offset by a loss in entropy when substrate binds and the flexible loop becomes ordered. So the observed binding is weaker than the sum of the interactions would lead you to expect. Therefore, product release is easier than the intrinsic binding energy would suggest, because when product is released the loop can become disordered again, with a gain in entropy.

Deisenhofer: How do your examples compare with the case of trypsinogen and trypsin? When trypsinogen is activated, a large domain of the enzyme undergoes a transition from a disordered to an ordered state and so forms the specificity pocket (Fehlhammer et al 1977). This seems to be the opposite of what you described.

Petsko: I don't mean to imply that this particular mechanism will apply to all enzymes. Here (i.e. with the trypsinogen-to-trypsin conversion) one has a specialized problem, namely turning the enzyme on and off at the physiologically appropriate time. It has been argued convincingly that there are good reasons for using a transition between order and disorder to do that. This is slightly different from the case of triose phosphate isomerase, say, where you may want to exclude solvent from contact with the substrate but also provide a relatively facile mechanism for getting substrate on and off the active site. There will be many enzymes where one wants a steric solvent exclusion. If you don't want it to be difficult to have binding and release, some type of flexible region is likely to be found in those enzymes.

Steitz: Could the substrate come off triose phosphate isomerase in the closed conformation?

Petsko: If you look at the solvent accessibility of the atoms by Richards-type calculations, it doesn't look as if that is an easy thing to happen, but we don't know that experimentally. We are now trying to trap the enzyme in the closed conformation, and also to cut the loop with proteases, so that the two new ends of the loop are free to move around. We want to see what happens if we make the loop totally flexible and do not even constrain it in the central portion. Those experiments are not complete yet.

Steitz: Do the crystallization conditions make a difference to the flexibility observed and thus to its expected function?

Petsko: This flexible loop has been seen in three different crystal forms of triose phosphate isomerase, including two crystal forms of the yeast enzyme which have been grown at different pH values, both from polyethylene glycol with no salt present, and a form of chicken muscle triose phosphate isomerase where the mother liquor was 75% saturated ammonium sulphate. It is interesting, however, that in the chicken enzyme the packing of molecules in the crystal lattice is such that the loop is blocked in one subunit and does not move readily. Therefore, if you diffuse substrates or inhibitors into crystals of the chicken isomerase, binding is to one active site only—the active site that is free to move. This is alarming for people who do crystallographic investigations of enzyme activity, because it implies that in some cases the lattice forces are strong enough, relative to other forces, to inhibit the necessary conformational change.

Williams: It is difficult to understand how you get from a crystal structure to a discussion of entropy. All you seem to be seeing are certain crystal states, and

you have to make assumptions to describe entropy. I can see only a kind of 'fuzz' in your maps. From these to entropy is a long step, because the fuzz could come about in different ways. Can you explain how a crystallographer goes from a fuzzy map to entropy?

Petsko: By fuzzy thinking, of course! There are two reasons for believing that, for this particular loop in that enzyme, we are looking at something that is conformationally flexible rather than something that exists in a number of discrete *states*. The first reason comes from nuclear magnetic resonance spectroscopy, which shows a sharp resonance for an amino acid that is found in that loop sequence in triose phosphate isomerase. The second reason is that if we look at the structure of the isomerase in two totally different crystal forms, the yeast and chicken enzyme crystals, the same loop is disordered in both. From those two pieces of evidence we think it likely that the residues in the loop are mobile, and not just a static distribution of different conformations. If we now assume their mobility, we have the curious fact that on binding the physiological substrate to the crystalline enzyme, the loop folds over and becomes much more rigid, as demonstrated by both the quality and clarity of the electron density, and the displacement parameters. Given that this is so, we can ask why it might happen. We realized at this point that the transition of the loop from the disordered to the ordered condition was an entropy contribution that would serve to diminish the observed binding energy of the substrate over what it might be if the loop were always in a single conformation. We considered whether this might have catalytic advantages, given what we knew about the observed energy profile for this enzyme. This is the chain of reasoning that took us from the crystallographic observation to a discussion of entropy.

Williams: Can you distinguish whether you have three states, five states, 10 states, and so on? If you can't, then three states could contribute nothing, or very little, to entropy, five states a little more, and 10 states a bit more. Thus we do not really know what the entropy contribution is unless we know the number of states.

Petsko: I don't agree that three states contribute nothing. I can at least say what the minimum number of states is. Given the noise level of an electron density map, one can say that there have to be at least four distinguishable states for the electron density for a group of atoms to disappear.

Blow: At lower temperatures, radiation damage, as well as movement of the side-chains, becomes less. It is therefore possible that you are just solving your structure better at low temperature. This presumably will have a 'cooperative' effect: because you have solved the structure better, the accuracy with which you determine phase angles is not so bad as you extend the resolution. This will reduce the blurring in your map. Is it possible that these changes are artifactual and connected with the technical problem of working out the structure at different temperatures?

Petsko: We can't rule this out absolutely but we don't think this is a major factor. With the exception of chicken triose phosphate isomerase, we have been working deliberately with molecules which diffract very well at room temperatures and where radiation damage at ordinary temperatures is small. We chose to study myoglobin, ribonuclease and yeast triose phosphate isomerase in part for this reason. Also, in the structure of myoglobin at 80 K, for instance, some striking things are observed in the electron density itself. (The same, I believe, has been seen for trypsin, too.) For example, the side-chain whose electron density is extremely fuzzy at room temperature not only becomes sharp but in several cases resolves itself into several alternative positions, at 80 K. It seems difficult to imagine a specific mechanism, based on radiation damage or phase quality, that would produce such a chemically sensible result. It is also true that these regions are correlated with structurally 'sensible' parts of the molecule—that is, regions where you may have a hindered motion of this residue which becomes apparent as temperature is lowered. So in part we believe these results because they satisfy our prejudices about how these molecules ought to work, but we are at least aware that that is what we are doing! We are trying, nevertheless, to use experimental conditions that will minimize artifacts.

Deisenhofer: Colleagues of mine (Walter et al 1982) compared the crystal structures of trypsinogen at room temperature and at 103 K, paying special attention to the four disordered regions of trypsinogen (one at the N-terminus and three inside the polypeptide chain). These regions behave differently at low temperatures. The N-terminus becomes more ordered but the three loops remain disordered. This result cannot be explained by the artifacts mentioned by Dr Blow.

Karplus: R.M. Levy, D. Perahia and I (1982) did calculations for a single isolated α-helix between 5 K and 300 K, to see how the motions change and when anharmonic effects become important. When we looked at the change in the mean square fluctuations as a function of temperature over the range for which measurements were available for myoglobin, the results for the helix were very similar to the myoglobin results. At 80 K one is in the range where the motion is fairly harmonic, and the room temperature measurements which give larger fluctuations have significant anharmonic contributions. The good agreement suggests that both the approximate analysis of the X-ray experiments and the approximate calculations are telling us the truth.

You suggest that the change from one equilibrium state to another happens purely randomly, without any 'pulling' due to ligand binding or other interactions. I think that is possible, but there is one case where the data are sufficient to show that this is not true, namely for haemoglobin. It was originally suggested that there are two haemoglobin structures and that oxygen binds only to the relaxed structure; i.e. that oxygen does not help to 'pull' the haemoglobin

molecule from one quaternary structure to the other. From the kinetics of the ligand binding reaction one can show that this cannot be true; oxygen must bind to the 'tense' (T) structure and shift the equilibrium, rather than just selecting the molecules that are in the 'relaxed' (R) structure. In your cases one doesn't know yet whether the rates would be fast enough for it to be done just by selecting from the equilibrium rather than by 'pulling'. My guess is that each case will be different. It will be interesting to analyse ribonuclease kinetically to see whether the probability of a given state is large enough to explain the rate on a random basis.

Petsko: That is a good point. We examined various sequences of triose phosphate isomerase for residues in the loop that might bind substrate when the loop was open, to see if that might be an outer capture point for substrate. But there are no positively charged residues in the loop, in most of the enzyme sequences, and there is no obvious special interaction between those residues and the phosphate moiety of the substrate.

Karplus: It could be that the general electric field of the molecule would change to produce the required effect.

Petsko: It becomes complicated. The structure has not been analysed enough yet, but we were looking to see whether, with the loop closed, it would form something that might have a turn of helix, say, giving a dipole effect.

Steitz: Your difference Fourier map struck me as showing that the negative electron density was as strong as the positive electron density. That would imply as much ordered structure before adding the ligand as after. Why, therefore, do you feel that the structure in the absence of the ligand is the more flexible one?

Petsko: The negative density is weaker and more diffuse than the positive density, when seen in three dimensions.

Jardetzky: Could you control for potential artifacts by looking at a series of mutants or modified proteins where there is other evidence—for example, spectroscopic evidence—that the modification changes the mobility? Artifacts presumably would be common to all the structures.

Petsko: That raises the problem of cytochrome *c*, where there is much evidence from a variety of spectroscopic observations (including your own) showing a substantial change in mobility between the oxidized and the reduced forms of the protein, and the crystallographic data do *not* show this up.

Jardetzky: But if they did, in a given case, you would be on safer ground.

Petsko: This is the only case where it has been looked at and the changes did not appear! But you are correct.

Frauenfelder: The entire discussion may be on weak grounds because the way the data are usually evaluated is incorrect. Considering substates and using a standard program for the X-ray data evaluation is inconsistent; the standard program is based on the harmonic approximation and many substates are far

from harmonic. There is one saving grace, because what one really measures are form factors. The first term in the expansion of form factors that is different from 1 is proportional to the mean-square radius. Thus the first approximation isn't quite as bad as one would expect initially. Comparison of the x^2-values measured by Mössbauer effect with those obtained from X-ray diffraction shows that they have the same temperature dependence where expected and very similar values. Radiation damage is not important in Mössbauer effect. The agreement therefore implies that we can't be completely wrong.

REFERENCES

Fehlhammer H, Bode W, Huber R 1977 Crystal structure of bovine trypsinogen at 1.8 A resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. J Mol Biol 111:415-438

Levy RM, Perahia D, Karplus M 1982 Molecular dynamics of an α-helical polypeptide: temperature dependence and deviation from harmonic behavior. Proc Natl Acad Sci USA 79:1346-1350

Walter J, Steigemann W, Singh TP, Bartunik H, Bode W, Huber R 1982 On the disordered activation domain in trypsinogen: chemical labelling and low temperature crystallography. Acta Crystallogr Sect B Struct Crystallogr Cryst Chem 38: 1462-1472