A. Plückthun

Towards New Enzymes: Protein Engineering and Catalytic Antibodies

Towards new enzymes: Protein engineering and catalytic antibodies

Abstract: The article first discusses the principal factors contributing to rate accelerations in enzymes. Then, the chances and problems associated with four strategies to new enzymatic activities are scrutinized: the screening of microorganisms, random mutagenesis of a cloned enzyme, protein engineering and the generation of catalytic antibodies. Each of these topics is illustrated by several examples from the literature (80 refs.).

1. History and introduction

One of the most tantalizing dreams of organic chemists, ever since the discovery of enzymes and their intriguing potential, has been to one day have at their disposal enzymes specifically catalyzing any difficult synthesis. The second winner of the Nobel prize for chemistry, *Emil Fischer,* showed exceptional foresight in his award lecture, nota bene given in 1902[1], when he said: ". . . if we wish to catch up with Nature, we shall need to use the same methods as she does, and I can foresee a time in which physiological chemistry will not only make greater use of natural enzymes but will actually resort to creating synthetic ones."

This statement is remarkable for a number of reasons. We do not know if *E. Fischer* had any idea that it would take virtually a century of world-wide research for this goal to be brought closer. Indeed, it may still require several decades before this goal can duly be considered as achieved. *Fischer* also made his remarks without knowing the ways and means which researchers would have one day at their disposal. He could not possibly have known how important immunology and gene technology would become. His appreciation of the possibilities, in view of the general lack of knowledge about the structures and properties of proteins and enzymes in 1902, is truly astonishing. There is another reason, however, for quoting *Emil Fischer* at the beginning of this article. For him, to work at the interface between biology and chemistry was perfectly natural. His successors, however, especially those in Germany, often took a more purist line on chemical research and were consequently less farsighted. One would do well to remember *Fischer's* scientific work when discussing the significance of biochemistry within chemical science and education.

In considering how artificial enzymes might be synthesized, one should first appreciate how enzymes function[2]. Here,

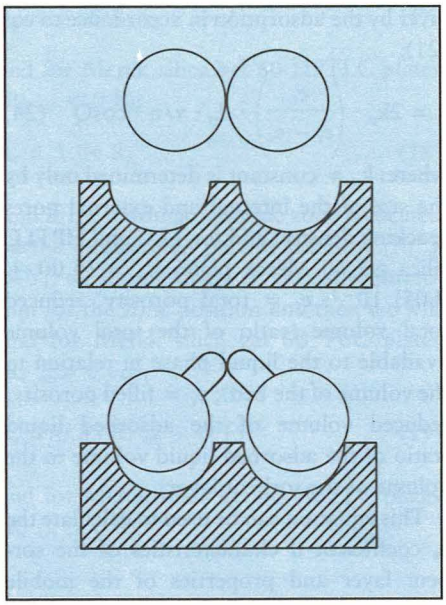


Figure 1: Haldane's concept of enzyme catalysis. The enzyme neither fits the substrate nor the product perfectly but is complementary to an intermediate state between the two. It thereby exercises a certain strain both on the substrate (top) and on the product (bottom).

too, we should perhaps delve back into history and start again with *Emil Fischer.* He recognized from his studies on sugar-converting enzymes that many of them are extremely specific. From this work comes the famous analogy of the substrate fitting the enzyme like a key fitting a lock[3]. This was a tremendously important realization, and today the analogy still serves to illustrate the concept of substrate specificity. It has since been directly borne out by crystallographic analysis of the structures of innumerable enzyme-substrate complexes.

One thing this analogy does not do, however, is explain why an enzyme should promote a chemical reaction at all. Indeed, an enzyme binding a substrate perfectly would simply leave it at that; the substrate would then be prevented from taking part in any reaction. A slightly amended theory leading us out of this dilemma (Fig. 1) was proposed by *J. B. S. Haldane* in 1930[4]. His theory allows that "the key does not fit the lock quite perfectly but exercises a certain strain on it". We can nowadays explain enzyme function better in terms of transition state theory[5, 6], which is based on chemical reaction kinetics and dates from roughly the same period. The first thing to be noted about the transition state is that it is only a conceptual model for a transient structure which exists between the product and the reactant. The stucture is that of the highest energy on the reaction pathway and the reaction pathway is energetically the most favorable path from substrate to product, not unlike a mountain pass. It can then be argued that a lowering of the energy of this transition state is the same as a rate acceleration[7]. In transition state theory, the transition state is treated as if it were a stable entity for which equations can be formulated and calculations performed. Although the transition state is only a model, it is a very useful and productive concept[2].

Haldane argued, therefore, if not quite in these words, that an enzyme would do better by being structurally complementary to the **transition state** (not the substrate), in order to stabilize it. It was in 1946 that *Linus Pauling,* another of the great names in chemistry (who also, incidentally, cared little for textbook definitions of chemistry), examined the theory more closely. He went one step further postulating that, if enzymes really function in this fashion, they should bind the transition state much more effectively than the ground state and should thus also bind any substances more tightly, which structurally resemble the transition state more than the substrate. Such substances are now known as "transition state analogs"[9].

There has been much discussion as to whether this is an appropriate name for these substances or whether they might not better be termed "intermediate analogs". In physical theory, the difference is fundamental but, in practice, merely semantic.

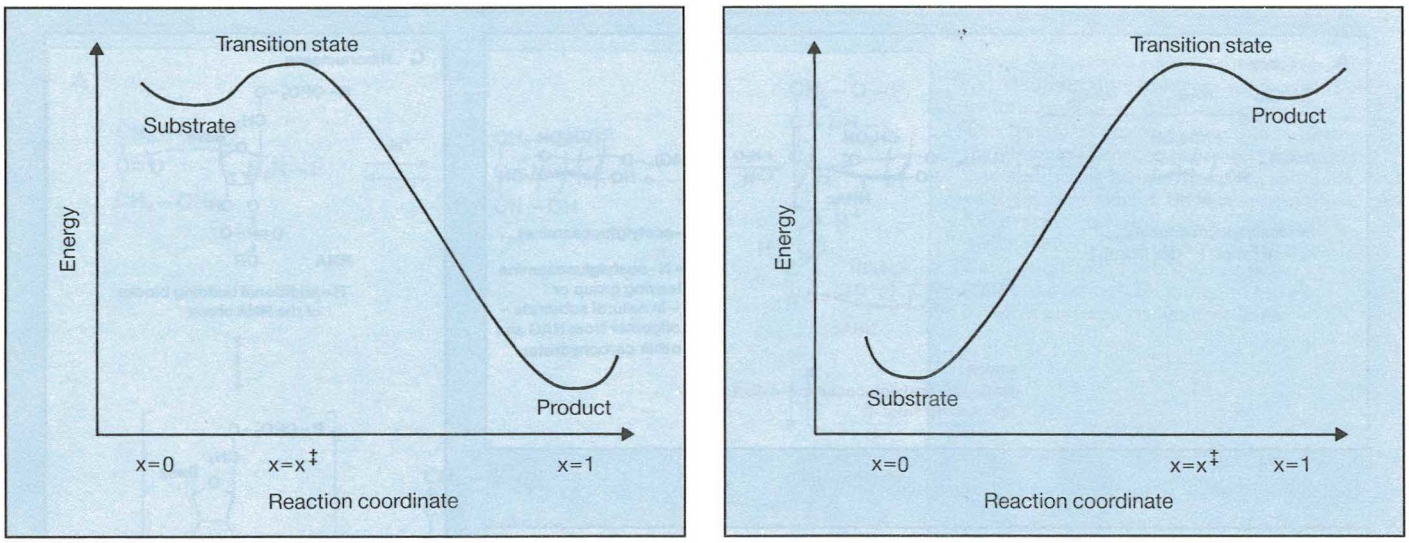


Figure 2A: Hammond's postulate: In an exergonic reaction the transition state occurs "early" along the reaction coordinate, i.e. the transition state resembles the substrate in structure and energy, while in an endergonic reaction the situation is reversed: the transition state occurs "late" and is similar to the product in energy and structure.

Hammond[10] postulated that a transition state is similar, in terms of energy and structure, to an unstable intermediate immediately preceding or succeeding it along the reaction coordinate (Fig. 2). The structural differences are probably so slight that the active center of the enzyme would be incapable of distinguishing an intermediate from the nearby transition state. Also, the term "transition state analog" has now come into general usage. The main consequence of this enzyme model is, therefore, that such a substance would be bound **more strongly** to the enzyme than the substrate. In fact, this has been borne out for a number of reactions and substrates[9] (Fig. 3), underpinning the model.

2. How do enzymes work?

We should now take a closer look at the question of why the reaction proceeds more swiftly in the active center of an enzyme than in the solvent. One of the most important points to have been recognized over time is that there is no single mechanistic reason, but rather that the enormous rate acceleration achieved is due to a number of mechanistic factors which have different weight in individual enzymes and can combine to elicit large effects[2]. The various factors shall now be examined in more detail.

2.1 Covalent catalysis

In covalent catalysis, the reaction in the active site of the enzyme may not be the same as in solution. Covalent intermediates may occur which are more reactive for chemical reasons or reasons of entropy (see below). Fig. 4 provides two examples. Actually, this type of catalysis is a rather "unfair" comparison of the reaction in the enzyme and in solution, therefore not warranting a more detailed discussion here. The organic chem-

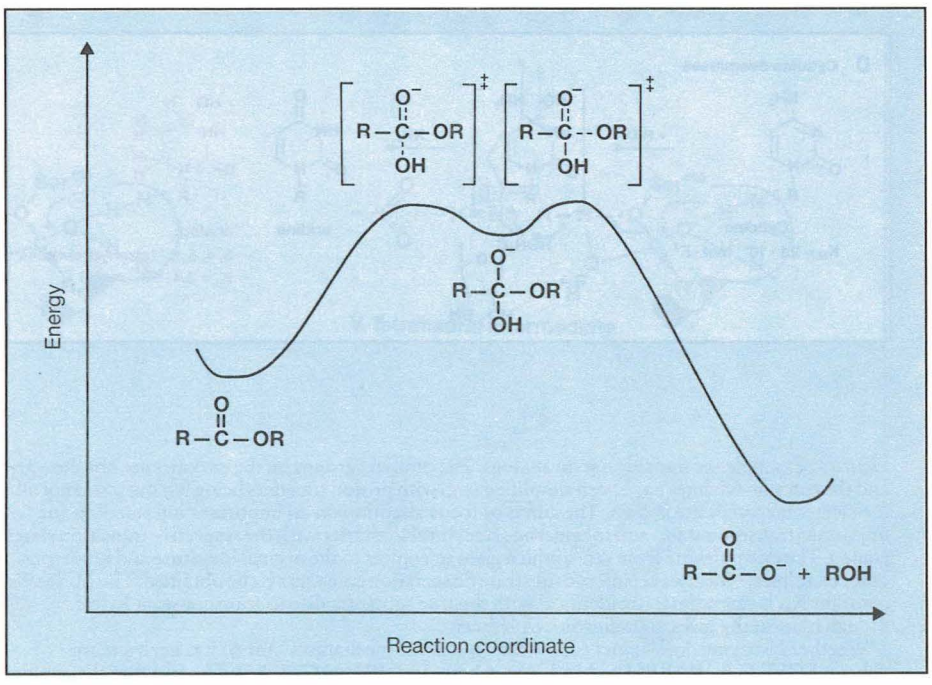


Figure 2B: When a reaction proceeds via an unstable intermediate, the transition states are more similar to this intermediate than to the substrate or product. The diagram is a simplified representation of an ester hydrolysis since proton transfers are ignored.

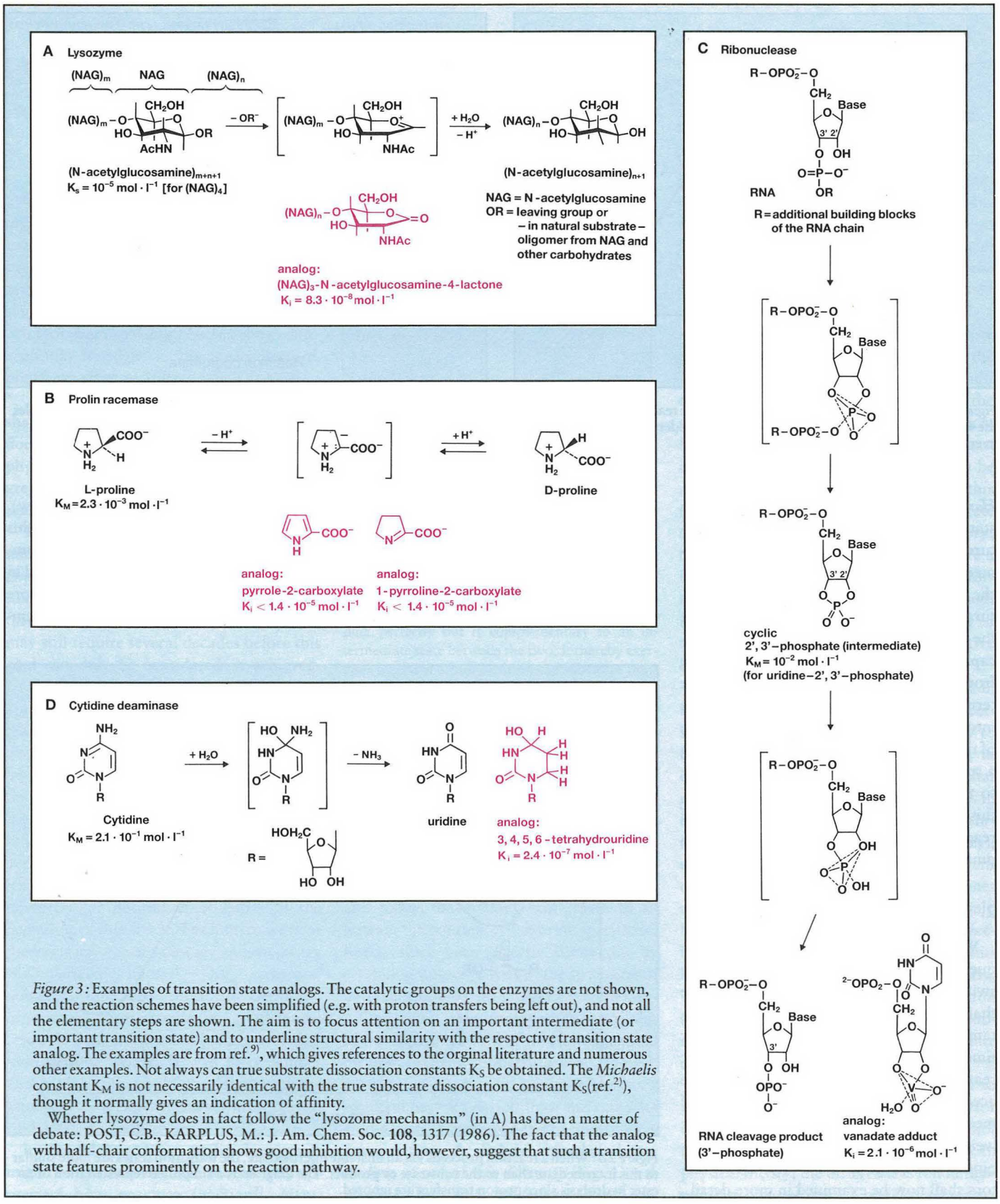


Figure 3: Examples of transition state analogs. The catalytic groups on the enzymes are not shown, and the reaction schemes have been simplified (e.g. with proton transfers being left out), and not all the elementary steps are shown. The aim is to focus attention on an important intermediate (or important transition state) and to underline structural similarity with the respective transition state analog. The examples are from ref.[9], which gives references to the orginal literature and numerous other examples. Not always can true substrate dissociation constants K_S be obtained. The *Michaelis* constant K_M is not necessarily identical with the true substrate dissociation constant K_S(ref.[2]), though it normally gives an indication of affinity.

Whether lysozyme does in fact follow the "lysozome mechanism" (in A) has been a matter of debate: POST, C.B., KARPLUS, M.: J. Am. Chem. Soc. **108**, 1317 (1986). The fact that the analog with half-chair conformation shows good inhibition would, however, suggest that such a transition state features prominently on the reaction pathway.

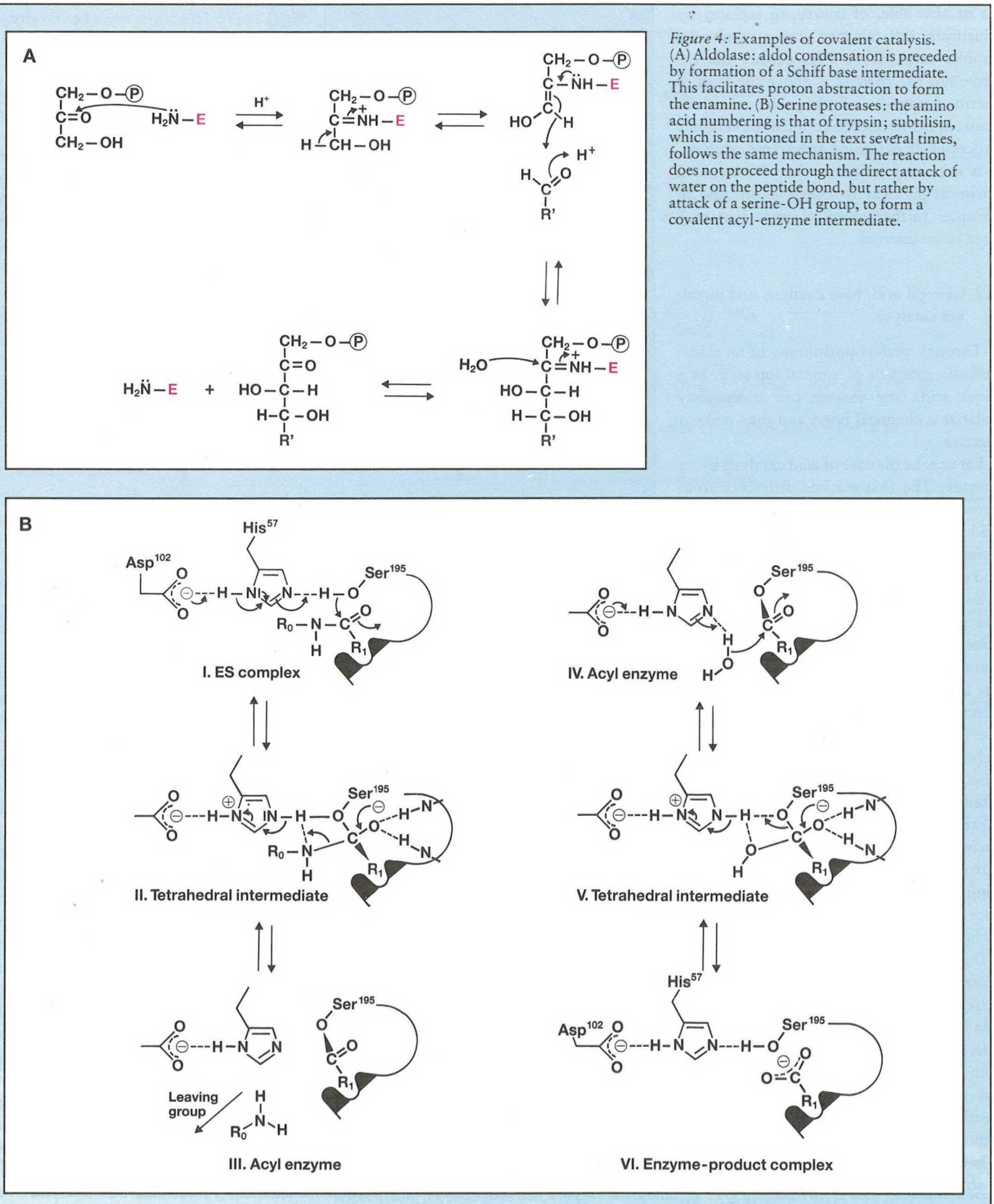


Figure 4: Examples of covalent catalysis. (A) Aldolase: aldol condensation is preceded by formation of a Schiff base intermediate. This facilitates proton abstraction to form the enamine. (B) Serine proteases: the amino acid numbering is that of trypsin; subtilisin, which is mentioned in the text several times, follows the same mechanism. The reaction does not proceed through the direct attack of water on the peptide bond, but rather by attack of a serine-OH group, to form a covalent acyl-enzyme intermediate.

ist ist also able, of course, to replace one kinetically difficult step with a number of simpler ones. Obviously, neither the enzyme nor the chemist is above the laws of thermodynamics. Endergonic reactions must, in either case, be coupled to exergonic processes. In designing an enzyme catalysis, it is necessary to assure that the intended chemical pathway is a feasible one and to consider further chemical steps that may need to be inserted.

2.2 General acid/base catalysis and metal-ion catalysis

Through perfect positioning of an acidic or basic group or of a metal ion (e. g. as a *Lewis* acid), the enzyme can enormously polarize a chemical bond and thus make it reactive.

Let us take the case of acid catalysis by an enzyme. The first marked difference from chemistry in solution is that, in the enzyme, acid catalysis can occur selectively at one point in the active center (i. e. regioselective and enantioselective catalysis), while in 1 M hydrochloric acid all of the sufficiently reactive groups are attacked. The second difference is the high local concentration of amino acid side-chains in the enzyme active site functioning as proton donor. This high effective concentration occurs by virtue of the fact that the substrate ist positioned tightly in the active site (see also point 2.4). Also, geometries (distances and angles) are often optimal for proton transfers. In order for the substrate to have an equal chance of encountering a proton in solution as in the active center of the enzyme, the acid concentration would have to be so high as to produce unwanted secondary reactions or even higher than physically possible.

Due to local electrostatic effects in the protein, individual amino acids may have extreme pK_a values. Consequently (and also because, when substrate is present, the active center will not always make contact with the surrounding solvent), certain proton transfers may take place which the chemist might find surprising at first sight[2, 11] (Fig. 5).

In the active site of the enzyme as well as in solution, rates of proton transfers must obey general physical laws, and *M. Eigen*[12]

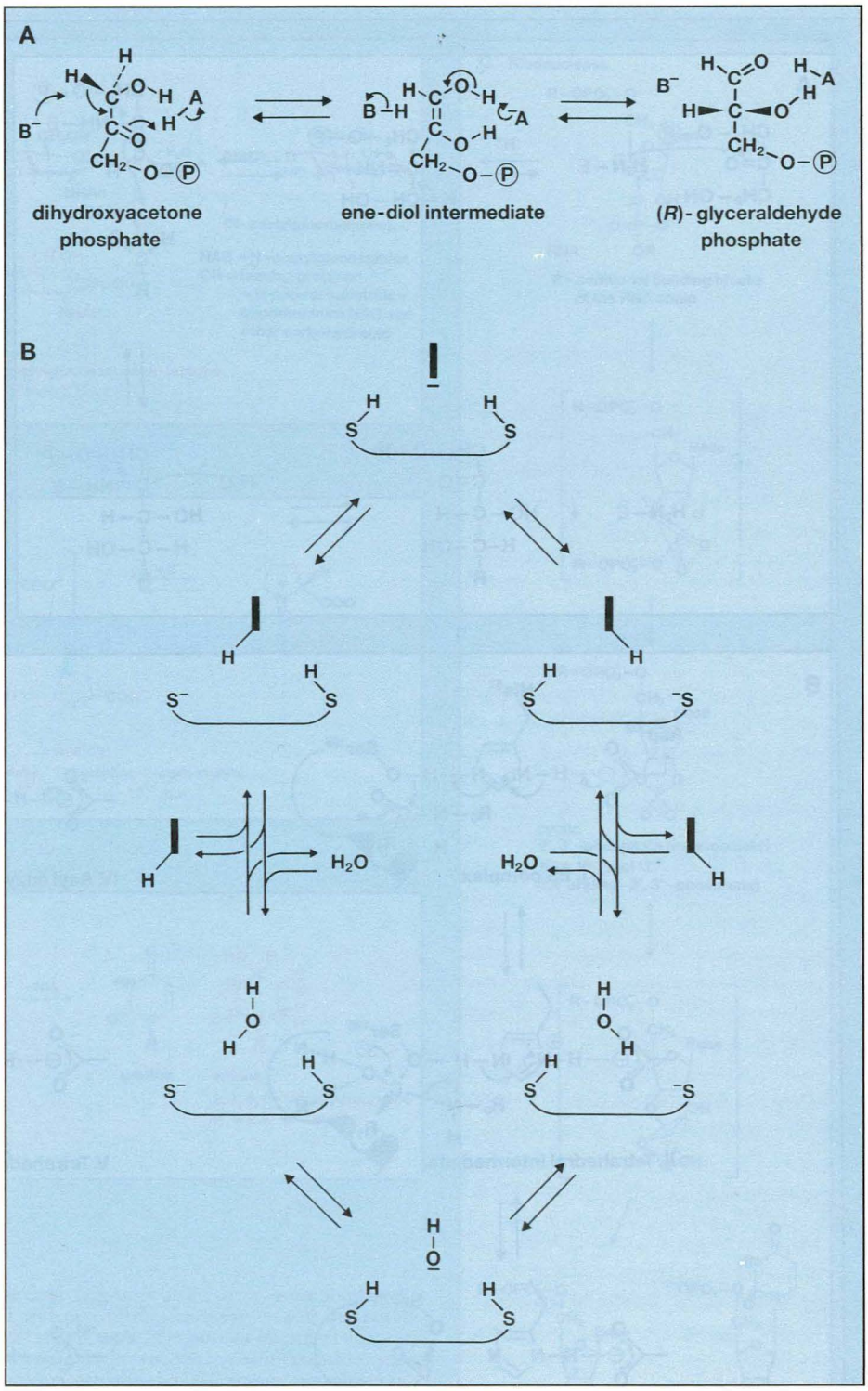


Figure 5: Examples of proton transfer in enzymes. (A) Reaction mechanism of triosephosphate isomerase whose catalytic base B is the carboxyl group of a glutamate residue. The pK_a for this carboxyl group is approximately 7 in the enzyme, while for a free glutamate it is about 4.6. (B) Schematic representation of the reaction mechanism of proline racemase. The proline ring is shown edge-on as a black bar. The pK_a for both cysteines in proline racemase is approximately 8, and for a free cysteine it is around 9.1-9.5. The C-H acid proton of proline has a pK_a of approximately 23 in water, while in the enzyme it is close to 17.5 (*J.R. Knowles*, personal communication). The racemase thus lowers the substrate pK_a by approximately 6 pH-units.

described the kinetic principles which apply to **any** proton transfer between two groups of known pK_a.

2.3 The enzyme as "super-solvent"

Slow-reacting substrates may not necessarily be intrinsically inert. It may be, for instance, that the strong hydration shell, which forms around a charged particle, markedly reduces its nucleophilicity or electrophilicity. In this case the active center of an enzyme may resort to "solvation substitution"[13], i.e. some of the water molecules are replaced by groups from the protein. Removing the hydration shell from between two reactants can enable substrates to manifest greater activity in the active site of the enzyme[2]. An active center also alters the distribution of electrons in the substrate at the desired position.

2.4. Entropy effects and geometric effects

Considerable rate acceleration for bimolecular reactions catalyzed by an enzyme is achieved simply by the fact that the reactants do not need to find each other in dilute solution but are already bound at the active site at the right distance and at the right angle. We know, for example, from studies on the formation of addition complexes of carbonyl compounds that nucleophilic attack of the carbonyl carbon can only occur from within a certain cone[14]. In innumerable kinetic investigations of organic model reactions (Fig. 6), attempts have been made to quantify this effect[2, 15]. Quantification of entropy loss by binding and approximation and "freezing" of rotational degrees of freedom remains controversial, but need not concern us in the following discussion.

Not only nucleophiles but also electrophiles (normally reactive groups in coenzymes) or acidic and basic groups, as well as metal ions, must have optimal geometric arrangements for very high reaction rates to be achieved. For example, simply by exchanging the catalytic glutamate residue for the slightly shorter aspartate in the enzyme triosephosphate isomerase, *J. R. Knowles* observed that the k_{cat} value of the enzyme was reduced by three orders of magnitude[16].

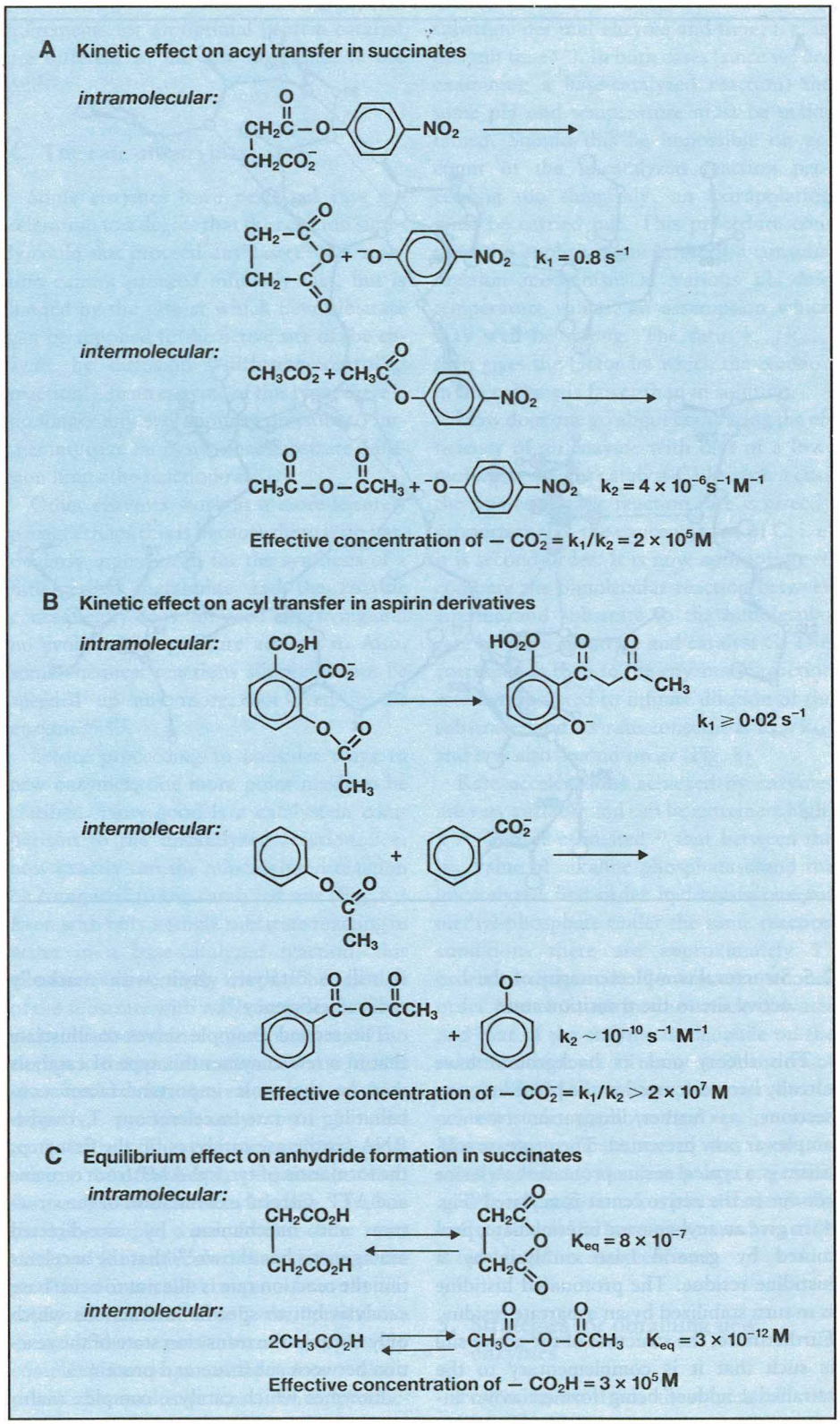
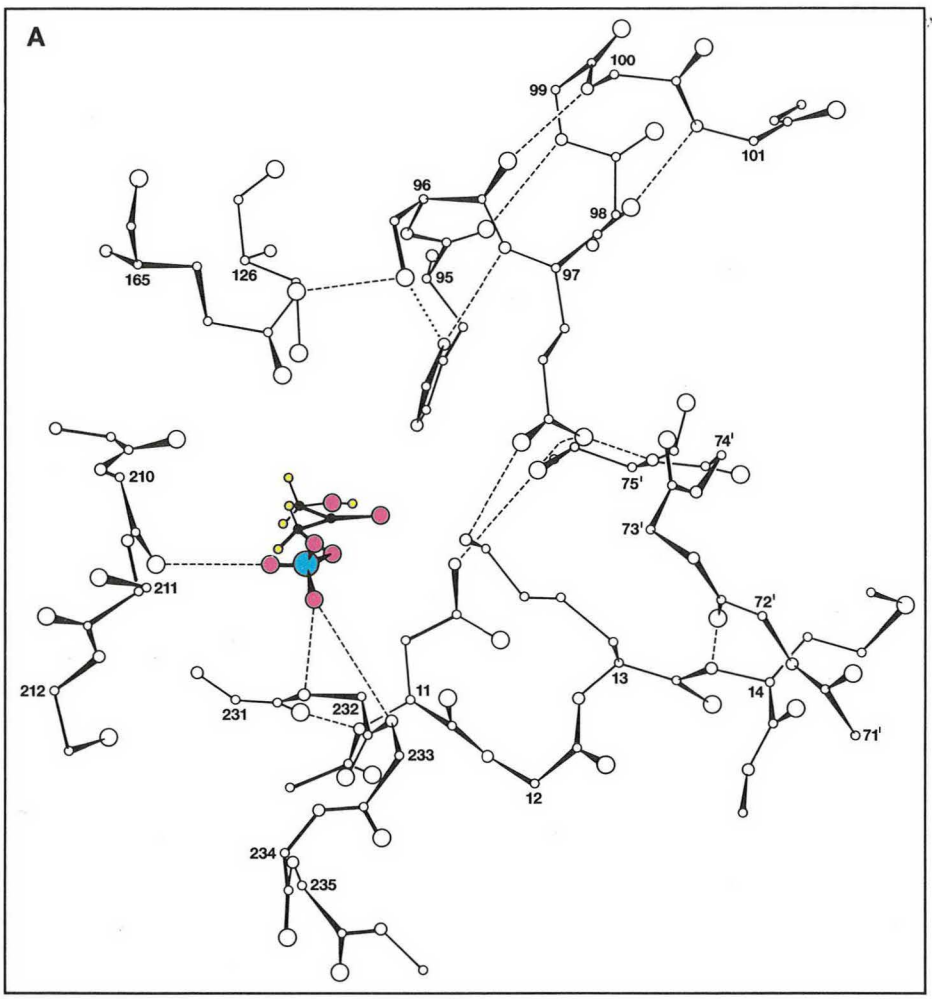


Figure 6: Examples of rate acceleration through a high effective concentration of neighboring group. Since a first-order reaction is being compared with a second-order reaction (with units s^{-1} and $M^{-1}s^{-1}$, respectively), the ratio gives the "effective concentration". For a detailed list of such phenomena, see KIRBY, A.J.: Adv. Phys. Org. Chem. 17, 183 (1980).



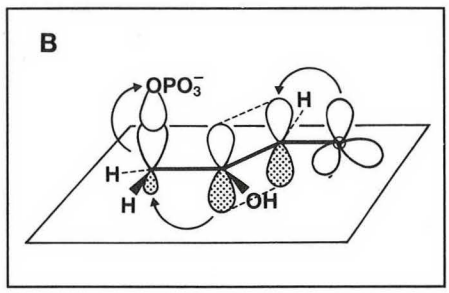


Figure 7: In the case of the enzyme triosephosphate isomerase the substrate is extended. (A) Model structure of the substrate dihydroxyacetone phosphate in the binding pocket. When the ene-diol intermediate (see Figure 5 A) is generated, it would have to change to a conformation in which the bond with the bridging oxygen of the phosphate and the π-electrons of the double bond are anti-periplanar in order to be able to eliminate phosphate (according to the theory of stereoelectronic effect). This is schematically shown in (B). In the binding pocket this is apparently prevented and the undesired phosphate elimination does not occur in the enzymatic reaction.
Adapted from: ALBERT, T. et al.: Phil. Trans. R. Soc. Lond. B **293**, 159 (1981).

2.5 Structural complementarity of the active site to the transition state

This theory and its background have already been discussed in the preceding two sections; as further illustration two examples ar now presented. The protease subtilisin is a typical serine protease[2]. A serine residue in the active center is acylated (Fig. 4) to give an acyl-enzyme intermediate, promoted by general base catalysis by a histidine residue. The protonated histidine is in turn stabilized by an aspartate residue. Furthermore, the structure of the active site is such that it is complementary to the tetrahedal adduct being formed as an intermediate on the serine residue[17]. When all three catalytic residues (Ser, His and Asp) are converted to alanine by site-directed mutagenesis, the "residual enzyme" is still a catalyst, albeit with markedly reduced efficiency[18].

The second example serves to illustrate that in a few enzymes this type of catalysis may be the most important factor contributing to rate acceleration. Tyrosyl-t-RNA-synthetase catalyzes, in the first step, the formation of tyrosyl-AMP from tyrosine and ATP. Careful examination of the structure and mechanism by site-directed mutagenesis has shown[19] that the acceleration the reaction rate is due not to acid/base catalysis but to specific interactions which **only** occur in the transition state of the reaction between substrate and protein.

Enzymes which catalyze complex multistep reactions must, under certain circumstances, undergo conformational changes in order to create an optimum environment for each individual step[20].

2.6 Prevention of side reactions

A chemical reaction may often proceed via an intermediate which may react further in several directions. This creates problems for the organic chemist when the reaction he desires is not the preferred one. An enzyme can often control the reaction pathway through appropriate stereochemistry in the active center[21] (Fig. 7). This may involve not only catalyzing the desired elementary step but also preventing other steps from occuring.

3. The size of enzymes

Different enzymes in fact use various combinations of these mechanisms. A synthetic catalyst which utilizes only one of these mechanistic devices will generally not

be capable of achieving the same rate acceleration. The question arises again and again as to whether there might not be other additional forces which only enzymes can exploit. For instance, enzyme flexibility and size[22] has often been discussed in the context of catalytic efficiency[23]. Also, a variety of "unconventional" theories has been suggested, for instance, that enzymes may direct the thermal energy of the solvent into targeted vibrations on the bond which is being cleaved[23]. Evidence for this idea is scant. If contributions of this type were to be of general significance, it should be impossible to synthesize small organic molecules having the efficiency of enzymes[24]. Although model enzymes exhibiting true, efficient "turnover" (i. e. **true catalysis** and not simply a stoichiometric reaction) are very rare, there are a number of highly efficient ones among the stoichiometric enzyme models[25] and true catalysts[26]. This indicates that there is no reason to assume that enzymes operate by virtue of some secret forces. Rather, they have simply evolved to combine several efficient mechanistic devices well known from physical organic chemistry (which sometimes, of course, are associated with conformational changes).

It is intriguing, nevertheless, to note how large enzymes are in relation to their substrates. We know much too little about the evolutionary origin of protein structures and mechanisms of protein folding[27] for us to judge whether or not a given protein structure is required to bring the few amino acids at the active site exactly into the right position. There is also, of course, the problem of regulation (e. g. allosteric affects) and multifunctionality[2] which increase the required enzyme size. Furthermore, many authors assume that the **main chain** (i. e. the structure created by the folding topology of the polypeptide backbone) might be involved in substrate binding as such, at least in some cases. The dipoles created by whole helices could, for instance, be utilized[28]; then, a part of the protein structure necessary for function is already given. None of this implies, however, that a newly constructed enzyme (when only its catalytic function in vitro is considered) must necessarily be so large. It is clear that the requirements for an optimal protein catalyst are different in the test tube than in the cell[29].

4. The rate of enzymes

Some enzymes have perfected rate acceleration to a degree that the reaction simply could not proceed any faster[30]. A reaction cannot proceed infinitely fast, but is limited by the rate at which new substrate can be supplied to the active site of the enzyme by diffusion ("diffusion-controlled reaction"). In an enzyme of this type, there is no longer any evolutionary pressure to further improve catalysis, since substrate diffusion limits the reaction rate.

Other enzymes work at a more leisurely pace. Perhaps this is because there is no particularly urgent need for the synthesis of a little-needed metabolite and the enzyme consequently does not need improving and no evolutionary pressure acts on it. Also, some chemical reactions simply cannot be speeded up any more, not even in the enzyme[30, 31].

Before proceeding to consider ways to new enzymes, one more point needs to be clarified. How good is a catalyst in comparison to the uncatalyzed reaction, i. e. how exactly can the noncatalyzed reaction be compared to the catalyzed one (Fig. 8)? Even with only a single substrate reacting in water in a base-catalyzed reaction, this question is not entirely trivial. The reaction of the substrate with water is a pseudo-first-order reaction (since there is virtually no change in the concentration of water) and is characterized by the value k_{uncat}, which depends, predominantly on pH and temperature. The enzyme-catalyzed reaction, on the other hand, comprises (at least) three steps: (bimolecular) binding, catalysis (in this example, a pseudo-first-order reaction) and (unimolecular) dissociation of the product. Kinetically comparable to the uncatalyzed reaction is only the "productive decay" of the enzyme-substrate complex (ES complex) into enzyme and product – also a pseudo-first-order process. This "productive decay rate" for the ES complex is equal to the turnover number k_{cat}, i. e. to the maximum reaction rate when the enzyme is completely saturated (expressed in mol substrate per mol enzyme and time, i. e. in the unit time^{-1}). In both cases (since we are examining a base-catalyzed reaction) the same pH and temperature must be maintained. Should this be impossible on account of the uncatalyzed reaction proceeding too sluggishly, an extrapolation must be carried out. This procedure contains the implicit assumption of a constant reaction mechanism at various pH and temperature values, an assumption which may well be wrong. The ratio k_{cat}/k_{uncat} then gives the factor by which the reaction in the enzyme is faster than in solution.

How does one go about comparing the efficiency of an enzyme with that of a low-molecular weight catalyst C? In such a case the non-enzymatic reaction rate is directly proportional to the concentration of C, i. e. it is second-order. It is now appropriate to compare the bimolecular reaction between enzyme and substrate to the bimolecular rate between substrate and catalyst C. This corresponds then to the enyzmatic reaction rate extrapolated to infinite dilution of the substrate, and its rate constant is k_{cat}/k_M, and it is also second-order (Fig. 8).

Rate accelerations achieved by enzymes are very variable and can be extremely high. *J. P. Guthrie* estimated[32] that between the k_{cat} value of alkaline phosphatase and the uncatalyzed first-order hydrolysis rate for methyl phosphate under the same reaction conditions there are approximately 17 orders of magnitude. Between the second-order rate constant for the attack of water and that of the enzyme nucleophile on the substrate there are as many as 21 orders of magnitude! Most enzymes, however, do not achieve such enormous rate accelerations. Moreover, comparisons of this type are mostly academic since, in practice, general acid/base catalysis takes place through the buffer, so that the "non-enzymatic" reaction is rarely an "uncatalyzed" one.

5. Strategies for obtaining new enzymes

Following on from these general considerations, we can now examine the various ways to new enzymes. At present 4 strategies can be distinguished:

1. Screening for new enzyme activities in

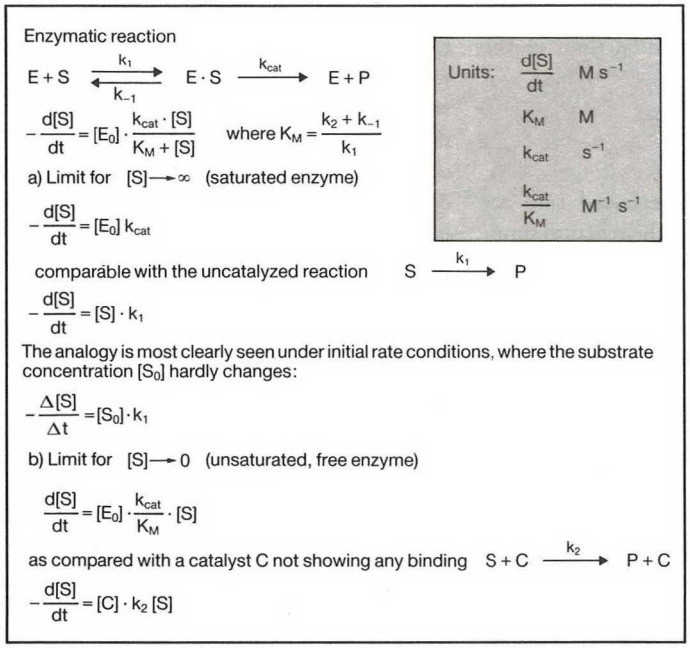


Figure 8: Comparison of the kinetics of enzymatic and nonenzymatic reactions.

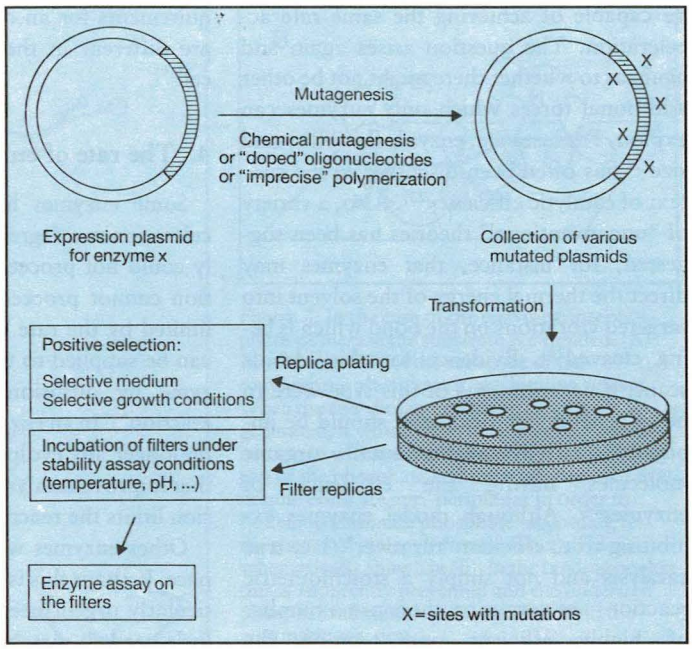


Figure 9: Strategy for random mutagenesis of a cloned gene.

species (mostly microorganisms) which have not previously been investigated.

2. Modification of an existing enzyme by random mutagenesis and screening for improved properties.

3. (More or less) rational engineering of an enzyme, with the aim of improving its properties.

4. Eliciting antibodies with catalytic activity.

The possibilities of and problems associated with these strategies are discussed and compared below.

6. Screening of microorganisms

The traditional way of obtaining new enzyme activities has been to search for new microorganisms which, because of their wide range of different growth conditions, are capable of producing many enzymes not found elsewhere[33]. Even today, this approach constitutes the most important practical method, as all the other methods are still in their early infancy. However, the strategy does require an efficient assay system for raw extracts, since it is not possible to perform enzyme purifications from a large number of strains based simply on suspicion and without a knowledge of the proteins being involved. The problem of finding a convenient assay is not a simple task. The desired activity may be masked by a variety of phenomena. For example, the product might go on to react with a different enzyme or the substrate might react considerably faster in a different reaction than with the enzyme under investigation. Additionally, the enzyme may only be present in small quantities and thus remain undiscovered, although it might have easily been cloned and overexpressed.

The enormous diversity and adaptability of microbial metabolism is reflected in a correspondingly vast number of enzymes which act on a wide variety of substrates and which have become optimally adapted to a broad spectrum of living conditions[2, 33]. Consequently, microorganisms are also frequently suitable sources of enzymes of specificities found elsewhere, but which have, for example, become adapted to high salt concentrations (halophiles) or high temperatures (thermophiles)[34].

Enzymologists hope that by studying these enzymes they will come to understand the mechanisms by which nature has enabled enzymes to adapt to adverse conditions and that, by using protein engineering methods, it will be possible in the long term to rationally adapt other enzymes to new reaction conditions[35]. Taking the enzymes from thermophiles as an example, the reasons for their increased stability are exceedingly complex and still poorly understood. Comparing the same enzyme from mesophilic and thermophilic organisms often discloses significant differences in sequence. Many of the differences found are due to normal genetic variation or drift. They don't influence the function of the enzyme and, either by chance or for other reasons, have become preferred in the organism. The goal, therefore, is to elucidate the differences leading to stability at high temperatures. This is a difficult task, however, since whether or not a residue or a loop in a protein has a stabilizing or destabilizing effect depends on the context in which the residue finds itself. Very likely, this question can only be solved by an analytical approach based on protein engineering methods.

In extreme cases, the search for intrinsically stable enzymes in thermophilic organisms can be disappointing. Some cases are known in which the enzyme isolated

from a thermophilic organism[36] is no more stable in isolated form than the analogous variant of a mesophilic species[34]. Stability of such enzymes is then due to the special conditions in the intracellular environment of the thermophilic organism[34, 36] (for example, salt concentration, metal ions, or special low-molecular weight compounds which stabilize the protein). Whether or not these reaction conditions are acceptable to an organic chemist intending to use an enzyme of this type will depend on the particular circumstances.

7. Random mutagenesis of a cloned enzyme and screening for improved properties

A logical extension of the screening method from the preceding paragraph has been made possible by the advent of gene technology. The gene from the enzyme of interest is subjected to random mutagenesis and mutants which exhibit improved properties are screened for. This might entail, for instance, a change in substrate specificity or stability. In theory, the following strategy could be used in any microorganism[37], though in practice it is convenient only with cloned genes in genetically well characterized host organisms such as *E. coli, Bacillus subtilis* and yeast. The gene encoding the enzyme is cloned and inserted into a plasmid, which is then selectively mutagenized. This involves, for instance, treating the plasmid with a high dose of mutagenic substances[38] or hybridizing a single-strand form of the plasmid with oligonucleotides which, through their synthesis, carry a certain number of random base changes[39] (Fig. 9). From the progeny, those clones must be found which carry the desired properties. This is generally the most difficult problem. The yield of mutants with the desired properties is very small already for theoretical reasons, since only a minute fraction of all conceivable changes will be beneficial. Thus, a vast number of colonies needs testing, requiring a growth or color assay to allow a decision to be made at colony level. Such a test should detect mutants which are able to react with a derivative of the substrate, or possibly others which are more stable, depending on the goal of the experiment.

The crux of the problem is the development of a reliable assay which will work at colony level. Subtilisin, a protease secreted by *Bacillus subtilis,* was subjected to such random mutagenesis and screening[40]. First, filter replicas were prepared from the Petri dishes. On the filters, the secreted protein localized within a halo around the colonies. After the filters had been incubated at the desired temperature or subjected to other adverse conditions (e. g. alkaline pH), a color assay was used for the remaining protease activity so as to detect mutants, which were still active following incubation, directly on the filter (see also ref.[41]). A similar procedure has been used to detect mutants that can utilize a modified substrate[42].

This method is only likely to succeed if there exists a variant that has the requisite properties and differs by only very slight modification from the starting molecule (generally 1 to 2 amino acid substitutions). A change in substrate specificity will thus necessarily be very small and any gain in stability rather moderate. Theoretically, this method can be used repeatedly on the improved variant. Only the future can tell whether success with this strategy is the exception or the rule.

8. Protein engineering

Because of space limitations, this article can only provide a rough outline of the current state of the art in protein engineering[43].

Central to any rational change in the sequence of a protein is a precise knowledge of its three-dimensional structure. X-ray crystallography is generally the method of choice[44]. Recently, however, the determination of structures in solution by NMR has made rapid progress[45]. While no crystals are required, the necessary technical sophistication is on a par with X-ray crystallography. Furthermore, additional efforts (e. g. biosynthetic labeling with stable isotopes) are required for determining the structure of a protein with more than 100 amino acids. The accuracy of the structure obtained from NMR is only rarely equal to that of a highly resolved X-ray structure since usually the conformation of only part of the amino acid side chains can be determined. It is possible, on the other hand, to derive other information, e. g. on dynamic processes in the protein, from NMR investigations. Moreover, the structure is not influenced by contacts between neighboring molecules in the crystal. The two methods are therefore complementary rather than competing.

Gene technology may be already helpful in the initial stages of solving a structure by facilitating the production of the enzyme in large quantities, and this methodology may make it possible to produce only a part of the protein (e. g. one domain), which is sometimes easier to crystallize than the complete protein.

8.1 Potential and limitation of theory

By closely inspecting three-dimensional structures, working hypotheses can sometimes already be formulated, so that an idea for achieving an effect on the protein's function may be tested through amino acid substitution (by methods of gene technology). This might involve a change in substrate specificity or pH optimum or stability. At this point, one must consider the effort involved in testing the hypothesis (through immediate production and characterization of the suitably modified proteins) and compare it to the work involved in a more rigorous theoretical analysis. It is clear that the requirements of a pragmatically minded chemist whose main interest is in an improved enzyme, differ from those involved in basic research in this area, who need to establish the fundamental basis of the effects observed in order to elaborate more rational approaches in the long term.

Modern graphics software can carry out an amino acid substitution on the screen within seconds but, initially at least, the image is a product of pure fantasy. The side chain of any amino acid is more or less free to rotate through a variety of torsion angles. What conformation will the new side chain and its old neighbors adopt? Worse still, it is not even sure that an amino acid substitution is a purely local phenomenon. It is perfectly conceivable (and there are well-

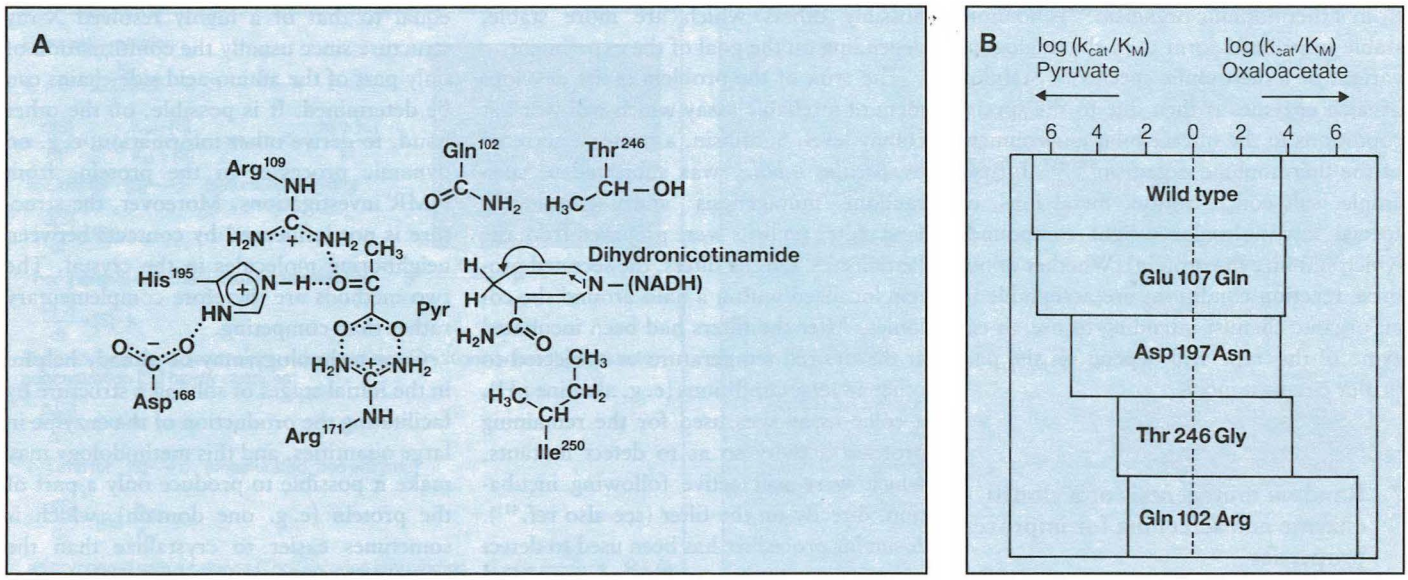


Figure 10: (A) Schematic representation of the binding pocket of lactate dehydrogenase. (B) The effect of individual mutations on the bimolecular rate constant k_{cat}/K_M for the reduction of pyruvate and oxaloacetate.

known examples) that the conformation of other, more remote side chains may also be modified or, in extreme cases, the conformation of the main chain itself may be altered.

A particular problem arises when the length of the chain of the "mutated" protein is different from that of the wild-type, whose structure is known. In most cases, this means that a loop in the chain of the new protein is now of a different length than the original one. What the conformation of this new loop looks like is a question which can only be definitively answered through X-ray structure analysis of the **modified** protein. The quality of the X-ray structure analysis is of enormous importance. Subtle changes can only be discerned (and the appropriate working hypotheses elaborated) when the structure has been determined to sufficiently high resolution. This also means that, ideally, one should have available a crystal structure of the exact protein destined for modification. One can, of course, through repeated use of "graphic" substitution, construct a model of a related protein. However, the cumulative effect of errors in fixing side chain conformers can rapidly lead to loss of quality as one progresses away from the sequence of the protein with known structure. Nevertheless, models of this type have been used in planning informative modifications[43]: one simply needs to appreciate the problems involved.

This state of affairs poses a challenge to theoretical chemists. One of the responses has been the development of empirical force fields; a potential energy can then be assigned to any conformation[46]. The forces which act on the atoms of a protein are made up of numerous components. Each chemical bond is of a certain equilibrium length and any deviation from this state elicits a force to return it to this length. Exactly the same principle applies to the bond angles and also, albeit for rather complicated theoretical reasons, to the torsion angles. In the protein, there are attractive and repulsive electrostatic forces at work which also have an effect on the position of every atom. In addition, all atoms are mutually attractive (*van der Waals*' forces) but repel one another when they come too close. Several research groups have in recent years formulated empirical force fields in an attempt to describe all these forces.

How can such a force field be used to obtain information about a protein structure? The first possible strategy is to start from a modelled structure (e. g. following an amino acid substitution) and to modify the structure according to one of several well-known algorithms until the potential energy has reached a minimum. This procedure is known as energy minimization. All known algorithms have the disadvantage of being "trapped" in the first minimum they encounter. This minimum ist almost certainly not the "global minimum". Considering that a molecule as complex as a protein has a vast number of degrees of freedom, there will in **every** case be a minimum of potential energy close to the initial structure; therefore after energy minimization the protein looks virtually as it did before.

Another method, developed independently by a number of research groups, involves defining that part of the protein structure one wishes to vary. One then obtains, either systematically or by using a random number generator, a large number of possible conformations[47]. The potential energy of all these is then calculated and the conformation with the lowest energy is selected. The question must always be posed, however, as to whether an adequate number of conformations was investigated.

A third method, which is again based on these empirical force fields but too complex to be discussed here in detail, is "molecular dynamics"[48]. The atoms are initially assigned random velocities and their movements are calculated as a function of

time. The movement of all atoms can be described by *Newton*'s law of motion, force = mass times acceleration. A simulation of motion is thus created for the protein. This has the advantage of directly testing the mobility of certain residues. This method requires supercomputer resources and, even then, can only simulate a few 10^{-11} sec in the life of a protein. Nothing of any direct interest to the chemist occurs during such a very brief time span, but tremendous energy is currently being devoted to extend the observable time span[49].

What all these methods have in common is a force field the quality of which is unknown and of which a number of variants exists. Since there has been far too little comparison of calculations with experimental data (and there is little agreement on what exactly should actually be compared), it is not possible at present to judge the quality of these calculations. Credit must be given to the theoreticians involved for their pioneering work, but these strategies must still be regarded as research projects with an uncertain outcome rather than as established predictive methods.

The conclusion to be drawn for all protein engineering projects is that there will always be a degree of uncertainty regarding the actual structure of the modified protein until it has been determined experimentally. Various research groups have now begun to determine a large number of structures of variants of the same protein by crystallographic methods[50, 51] so as to be able to discern systematic effects.

8.2 Some case studies

As an analytical method, protein engineering, i. e. the targeted modification of a protein of known structure (generally through the methods of gene technology), has already become firmly established in protein research and enzymology[43]. From the related literature, which has undergone explosive growth, a few examples are now selected and discussed. While being studied for the purpose of basic research, these examples do point the way to applications.

The first example deals with the deliberate modification of substrate specificity of an enzyme. *J. J. Holbrook* et al.[52] described the successful conversion of a lactate dehydrogenase into a malate dehydrogenase (Fig. 10). The active site of the enzyme had to be modified so that the carboxymethyl group (in malate or oxaloacetate) would be preferred to the methyl group (in lactate and pyruvate). Enzymatic activity of the modified protein depends on the unmodified part of the substrate being bound at the same position, so that the reaction with the coenzyme NAD can still take place as before. In separate experiments, two acidic residues located nearby (Glu107 and Asp107) were exchanged for the corresponding amide moieties (Gln and Asn), in an attempt to prevent a possible repulsion of the negatively charged side chain of malate. The small intrinsic malate dehydrogenase activity of lactate dehydrogenase was not increased thereby, but the lactate dehydrogenase activity was merely lowered! Similar results were obtained after the exchange of Thr246 for Gly, in an attempt to create more room for the bulky carboxymethyl group. It was only when Gln102 was exchanged for Arg that a breakthrough was achieved, presumably because the charge of the carboxymethyl group can now be complemented by the guanidinium group.

It is too early to deduce any general theories for engineering changes in substrate specificity from this one example. The change was only minor (the introduction of an additional carboxyl group), but it served to show that such an approach is possible in principle.

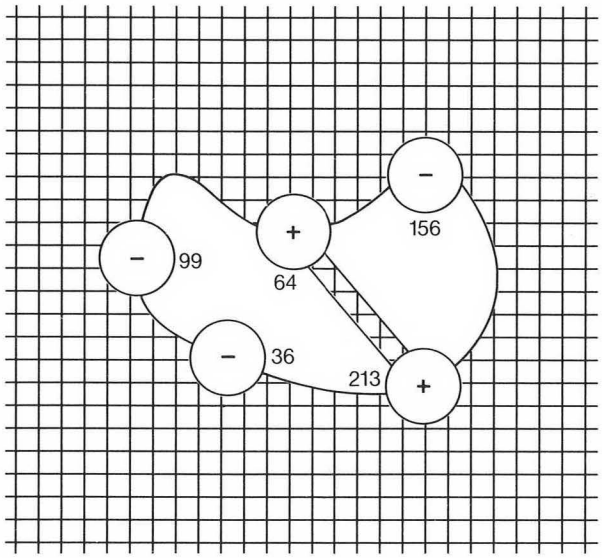
The second example concerns changing the pH optimum for an enzyme. A modification of this type might be useful, for example, when, in coupled enzymatic reactions, a common optimum pH must be found for several enzymes. Model studies for changing a pH optimum have been reported for the protease subtilisin[53]. This serine protease has a catalytically essential histidine residue in position 64 (Fig. 4 B). A titration curve of activity against pH reflects the pK_a of this imidazole ring. *Fersht* and coworkers expected that through electrostatic effects (elicited by changes in surface charges on the protein), it might be possible to influence the ease with which His64 can be protonated. This was indeed confirmed experimentally (Fig. 11). Only such polar residues were exchanged on the surface of the protein which were expected to be of no consequences either for the structural integrity of the enzyme or for substrate binding, but which make contact with the surrounding water. One interesting finding from this work is that the apparent dielectric constant within the protein is unexpectedly high, at around 50[53].

8.3 Engineering high stability

We now move on to the problem of creating more stable proteins through protein engineering. In order to understand the possible strategies, we need first to discuss in more detail the phenomenon of protein stability. This phenomenon is exceedingly complex and currently a topic of intensive research. Consequently, we can in this discussion only deal with the matter in very basic terms.

The native state in most proteins is only about 5 to 15 kcal/mol more stable than the unfolded state[27]. Although a vast number of interactions contribute to the stability of the native protein structure, virtually all of the amino acids involved in intramolecular interactions in the native state may interact with the solvent in the unfolded state. Also, in the unfolded state the entropy of the protein chain is far greater than in the native structure while the entropy of the solvent, through the larger hydrophobic surface accessible in the unfolded state, is lower[27]. The sum of all these numerous interactions within the protein and between protein and solvent and between solvent molecules must be compared for the folded and unfolded state. The difference is the free energy of stabilization for the native state. It is a difference of large numbers and it is very small indeed.

There are, however, many well-known exceptions. Phospholipase A_2, for example, a protein with approximately 120 amino acids and (generally) 7 disulfide bridges, can be subjected to prolonged boiling and storage in organic solvents without impairment of its specific activity[54]. Unusual stability is also seen in superoxide dismutase, which exhibits enzyme activity in the presence of normally denaturing detergents (e. g. SDS) or in denaturants such as 6 M urea[55].



Changes in the protein	Mean distance to nitrogen atoms of His 64 (Å)	ΔpK_a
Asp → Ser^{99}	12.6	−0.40
Glu → Ser^{156}	14.4	−0.38
Ser → Lys^{99}	15.0	(−0.25)
Ser → Lys^{156}	16.5	(−0.25)
Lys → Thr^{213}	17.6 (?)	+0.08
Asp → Gln^{36}	15.1	−0.18
Asp → Lys^{99}	(13.8)	−0.64
Gly → Lys^{156}	(15.5)	−0.63
Asp → Ser^{99} and Glu → Ser^{156}	(13.5)	−0.63
Asp → Lys^{99} and Glu → Lys^{156}	(14.7)	−1.00

Figure 11: Schematic drawing of the position of important ionic groups in subtilisin and their effect on the pK_a of His 64.

Neither of these proteins, incidentally, is derived from a thermophilic organism, and the list of such highly stable proteins can be extended. These examples show that proteins can be dramatically stabilized, a prospect which initially raises great hopes.

Nevertheless, the practical chemist is not interested in the free energy of stabilization of the native state. He is far more interested in the lifetime of the enzyme under reaction conditions (or perhaps even in the shelf life). This is not necessarily the same and the causes of both phenomena may in fact be very dissimilar.

Under "denaturing conditions" (i. e. at an elevated temperature or at inappropriate salt concentrations, extremes of pH or in the presence of denaturing agents such as urea or guanidinium hydrochloride), the native structure passes through a series of intermediates into disordered forms of the chain[27]. A folding intermediate along this path can now react further in a variety of ways. Only under a narrow set of conditions (mostly: low protein concentration, "correct" pH, "correct" salt conditions, low urea or guanidinium hydrochloride concentrations to avoid aggregation reactions) can an intermediate refold to the native state. Under most conditions something else will occur (Fig. 12): the intermediate is chemically inactivated[56], it aggregates, adsorbs onto the surface of the vessel, or folds into a form different from the native state. In these cases, inactivation is irreversible.

From this consideration, two points become immediately apparent. First, to successfully stabilize an enzyme the reason for loss of activity must be found[56]. Only by removing the true cause of the enzym's facile denaturation, can stability be raised. Second, there are two points at which the problem can be approached: at the reversible equilibrium between the native structure and a critical intermediate or at the subsequent irreversible step that is relevant to the enzyme.

Any attempt to stabilize the reversible steps, here referred to as "conformational stabilization", is hampered by a general lack of understanding about protein folding and protein structures. Nevertheless, through the efforts of various research groups at least a few important aspects have been identified:

1. Optimum packing within the hydrophobic core of a protein[57]; neither mutual steric hindrances nor cavities must be present.
2. Electrostatic effects, such as charged amino acid side chains interacting with helix dipoles and thus stabilizing the protein[58].
3. Networks of hydrogen bonds[59].
4. The effect of conformational entropy.

B. W. Matthews and co-workers[60] postulated that an amino acid with many conformational degrees of freedom in the unfolded state loses **more** entropy in folding than an amino acid which has fewer torsional degrees of freedom accessible in the unfolded state. He proposed, with experimental data supporting this idea, that the exchange of glycine for alanine or alanine for proline can have a stabilizing effect. The only requirement would be that there are no enthalpic reasons to the contrary, i. e. if the new residue were to collide with other parts of the protein.

A test of this hypothesis in the author's laboratory[61] may serve to illustrate the inherent problems. The model chosen was the enzyme glyceraldehyde phosphate dehydrogenase (GAPDH, a homo-tetramer). In separate experiments, all glycine residues occurring in helices were exchanged for

alanine residues. Only **one** exchange of this type brought about notable stabilization, both in irreversible denaturation experiments (i. e. measuring the half-life at high temperatures) and in urea-induced reversible unfolding and folding experiments. Exact analysis showed that the loss of activity does not correlate with unfolding of any helix. Rather, imperfect packing of the hydrophobic core of the wild-type subunits seemed to have been ameliorated through this exchange. This example shows that careful analysis is needed in order to gather information for further rational approaches.

The irreversible steps have likewise been a focus of protein engineering efforts. Subtilisin, for example, which is remarkably sensitive to oxidation of the Met222 residue, can be rendered far more robust through its substitution[62]. The Genentech group, approaching the problem pragmatically, substituted all the other 19 amino acids at this position and tested the activities and stabilities of the mutant enzymes. Numerous suitable substitutions were discovered in these experiments.

Also from the Genentech group comes an intriguing experiment on the question of the mechanism by which disulfides influence the stability of a protein[63]. Stabilizing disulfides were incorporated in various mutants of T4 lysozyme that differ in their stability because of different mutations elsewhere in the protein. It was demonstrated that the **reversible** unfolding of the mutant proteins is not at all affected by the presence of the disulfide bonds, but that the S-S bonds apparently prevent aggregation or misfolding of the partially unfolded intermediates and thus prevent their **irreversible** loss. The caveat of these experiments is that the conclusions may be valid only for T4 lysozyme. Disulfides, both intramolecular[64] and intermolecular[65], have since been incorporated into numerous proteins for stabilization purposes.

This brief summary is intended to outline the current state of protein engineering and to illustrate the possibilities which exist for obtaining new or improved enzymes. Though no modified proteins are as yet ready for the market, the rapid pace of progress in this area means that they might be in the foreseeable future. It remains doubtful, however, whether the first successful products will be the result of truly "rational" planning. The way to the routine use of these engineering methods is long and will still necessitate massive efforts in basic research.

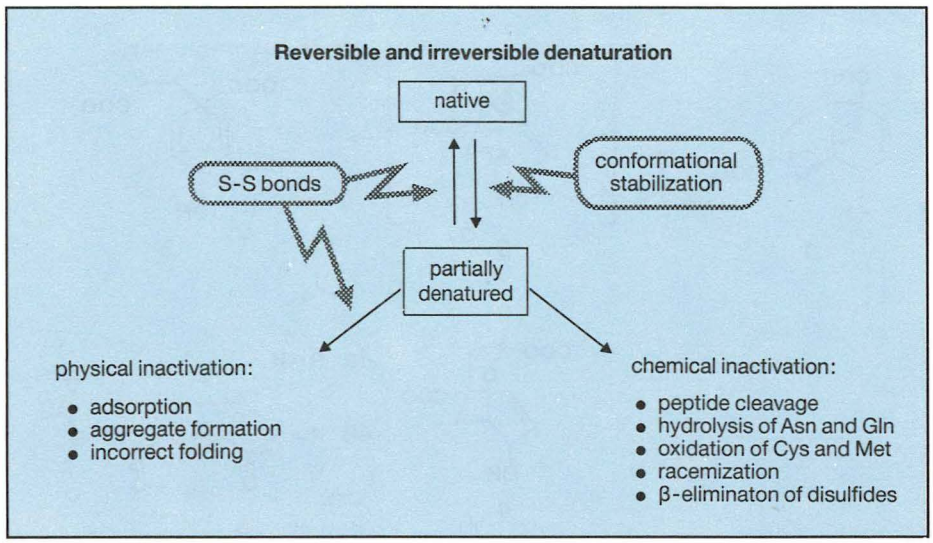


Figure 12: Schematic diagram of possible steps in the inactivation of an enzyme. The relative significance of these steps differs from protein to protein.

9. Catalytic antibodies

A fourth method of obtaining new enzymatic activities might be to start building from scratch. We are not considering science fiction here, however, and at the moment the prospect of "designer enzymes" is nothing other than that. Rather, we will concentrate on an entirely empirical strategy mentioned at the beginning of this article: the use of antibodies for catalysis.

The idea was first committed to writing by *W. P. Jencks* in 1969[66] (interestingly enough, in a textbook). If an enzyme has a structure that is truly complementary to the transition state of a reaction, *W. P. Jencks* surmised, then it should be possible to reverse the argument. Any protein having such a complementary structure should then be able to catalyze a similar reaction. The immune system is able, in a first approximation, to produce antibodies against any chemical substance and should thus permit the production of antibodies against transition state analogs. The question was, would such an antibody have any catalytic activity?

Just a couple of years later a number of research groups, working independently, tested this proposition, but achieved only moderate success[67]. The observable catalytic effects were generally only slight or, in some cases, not even measurable since the intrinsic rate acceleration caused by the antibodies was too small. In polyclonal antiserum, even after immunization of the animal, specific antibodies make up only a small fraction of the immunoglobulins. Additionally, some of the initial experiments were over-ambitious and aimed at overcoming tremendous energy barriers. Consequently, moderate rate acceleration would not have been discovered, because the reaction would have still proceeded far too sluggishly. The breakthrough came with the availability of monoclonal antibodies[68]. Only with these was it possible to achieve protein concentrations high enough to detect small catalytic activities. Monoclonal antibodies against transition state analogs have been produced since 1986, e. g. in the laboratories of *R. Lerner* and *P. G. Schultz*[69]. At the same time methods were developed in the author's laboratory for making the antibody molecule itself more easily amenable to modification by protein

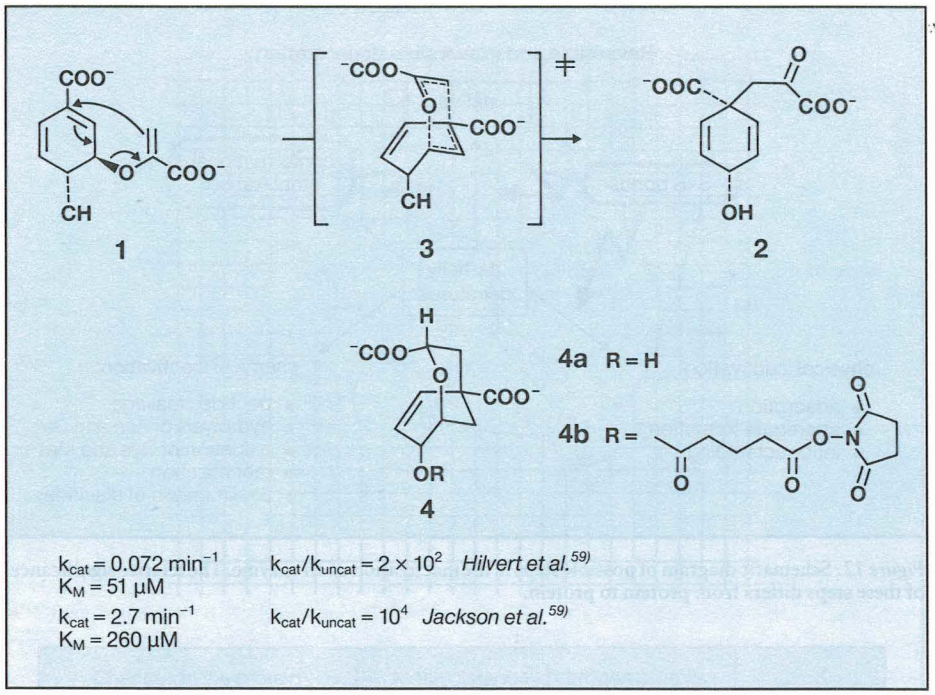


Figure 13: Reaction of chorismate mutase. Chorismate (1) rearranges to prephenate (2) via a chair-like transition state (3). The transition state analog was used in the free form (4a) to determine binding constants to the catalytic antibody. To derivatize an immunogenic protein it was used with a suitable spacer (4b).

engineering, thus expanding the potential for producing catalytic antibodies[70-72].

The strategy of producing a catalytic antibody by immunizing a mouse shall be illustrated by a number of examples. Two research groups simultaneously produced antibodies which catalyze a *Claisen* rearrangement[73]. Both groups chose the rearrangement of chorismate to prephenate (Fig. 13), which is catalyzed by the enzyme chorismate mutase, and is part of the pathway of the synthesis of aromatic amino acids in bacterial and plant cells[74]. The mechanism of the non-enzymatic reaction has been studied and it is known that the transition state passes through a chairlike geometry. In the transition state, the C-O bond is mostly broken before the formation of the new C-C bond. The enzymatic reaction (approximately 10^6 times faster) also proceeds via a chairlike transition state. proceeds via a chairlike transition state. Both research groups[73] therefore synthesizstate structure (Fig. 13) and coupled it, via a spacer, to an immunogenic protein. (It is generally not possible to elicit antibodies against a small molecule without coupling it to a macromolecule). This transition state analog inhibits chorismate mutase with a dissociation constant of approximately 0.15 µM, while the substrate binds with only approximately 41 µM. Both research groups were able to find an antibody which not only binds the antigen but which also accelerates the rearrangement, albeit at lesser efficiency than chorismate mutase. This presents strong evidence that the model of transition state complementarity is correct[4, 8]. Particularly interesting is the fact that the activation **enthalpy** is only slightly lower when compared to the uncatalyzed reaction (from 20.7 kcal/mol to 18.3 kcal/mol), while the activation **entropy** undergoes a much greater reduction, from -12.85 calK^{-1}mol^{-1} to -1.2 calK^{-1}mol^{-1} (see ref.[69]). The absence of solvent isotope effects is consistent with this antibody accelerating the reaction **only** through the binding site being structurally complementary to the transition state. The enzyme probably uses covalent catalysis[74] and can achieve greater rate accelerations.

A slightly modified strategy was used to elicit an antibody which catalyzes a β-elimination[75] (Fig. 14). In enzymatic reactions such processes are mostly base-catalyzed. The antigen should thus elicit antibodies that carry, in the desired position, an amino acid capable of functioning as a general base catalyst at neutral pH (e. g. glutamate or aspartate). To this end, an ammonium ion was incorporated into the antigen in order to create charge complementarity in the antibody at precisely the required position. This stategy led to moderate but measurable catalysis.

A third example is intended to show that by suitable design of the immunogen it is possible to catalyze even more demanding reactions, e. g. cleavage of a peptide bond. *B. L. Iverson* and *R. A. Lerner*[76] elicited an antibody that binds a metal ion adjacent to the peptide bond to be cleaved. For this purpose, an antigen in the form of a tetrapeptide derivative was synthesized (Fig. 15) that forms a stable complex with cobalt-"triene" (triene = triethylenetetramine) via an amine and a carboxyl group. The tetrapeptide substrate, a separate triene molecule, and various metal ions together with the antibody were used in the actual cleavage reaction. The basic idea was to get the antibody to form a binding pocket both for the peptide and for the metal-triene complex. The metal was thus to be placed in the vicinity of the bond being cleaved and act either as a *Lewis* acid polarizing a carbonyl group or, as *Brønstedt* base deprotonating a water molecule, which can then attack the peptide bond as a hydroxide ion. Indeed, this strategy produced an antibody which cleaves a peptide bond with a turnover number of 10^{-4} sec^{-1}.

This last example was a great pioneering achievement. It also serves, however, to show how far removed this technique is from "designer enzymes". The examples cited are only a selection (for a more recent review article see, e. g., ref.[69]), but they do illustrate the potential which exists for achieving new activities through immunization. In particular, a specific binding protein can be created without any need for a knowledge of protein folding, since the immunological approach is entirely empirical.

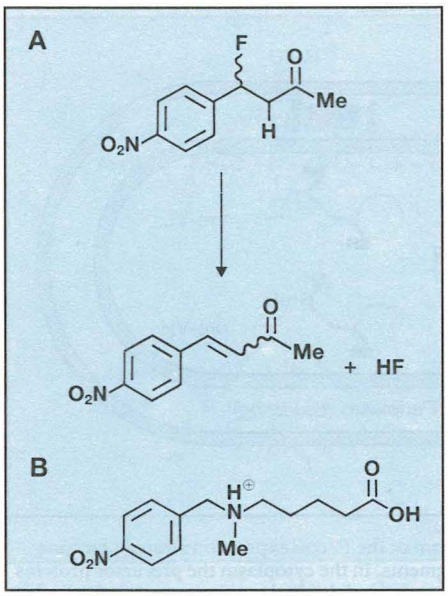


Figure 14: (A) Antibody-catalyzed elimination by an antibody elicited with the immunogen shown in (B).

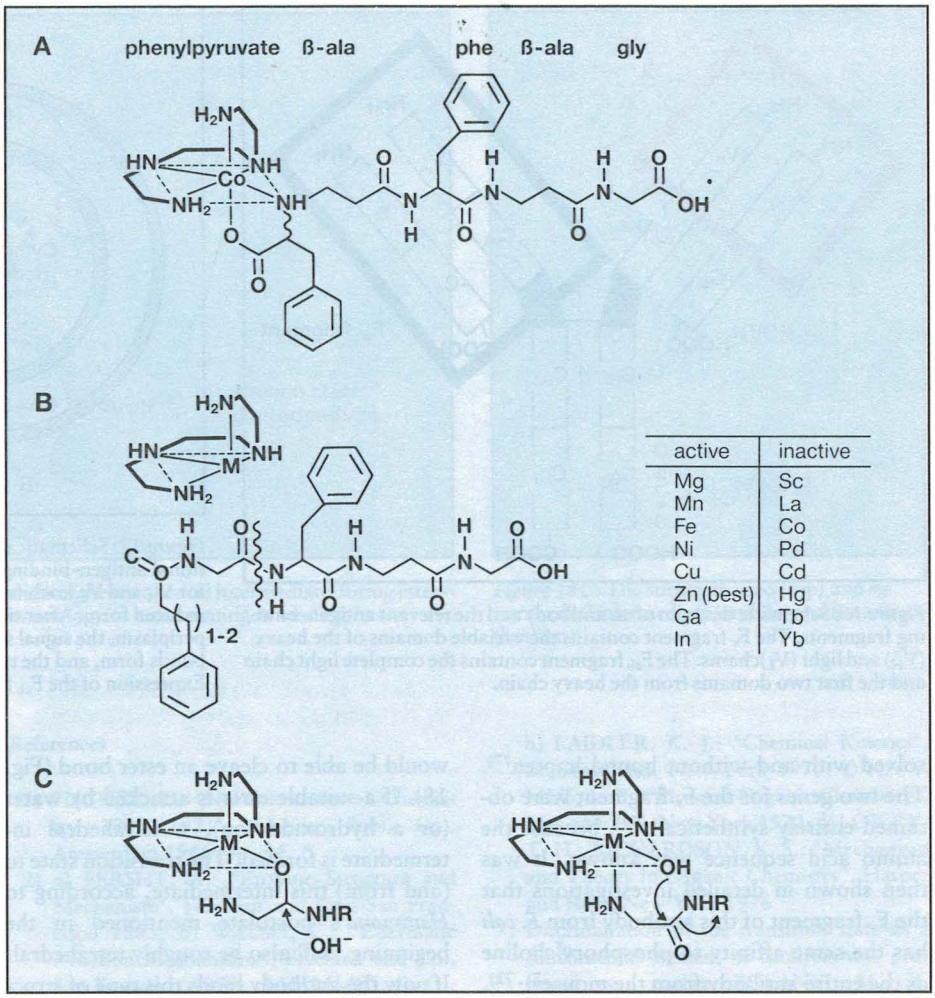


Figure 15: (A) As immunogen, the tetrapeptide was derivatized with phenylpyruvate and reduced, and combined with a stable cobalt-trien complex. (The covalent binding to an immunogenic protein is not shown here.) (B) The tetrapeptide and various metal-triene complexes were then used as substrates for the peptide-cleaving antibody thus obtained. (C) The two principal mechanisms in which a metal ion can accelerate peptide hydrolysis in the binding pocket; left, as a *Lewis* acid: right, as a general base.

While the broad spectrum of reactions already catalyzed raises hopes, the potential of this strategy still needs critical appraisal. Undoubtedly, potential catalysts for a large number of reactions and substrates may be developed by this approach due to the large available antibody repertoire. It remains to be seen, however, whether the activities achieved until now can be significantly improved. The immune response is not a selection for nucleophiles, but the antibodies are selected in the animal solely for their antigen **binding** affinity. Furthermore, the binding of metal ions to natural antibodies not subjected to protein engineering can only be achieved by a chelate molecule being part of the immunogen during immunization and the chelate then being a co-substrate. The optimum reaction rate is obtained when the pK_a of a catalytic group is approximately equal to the pH of the reaction: for enzymes this usually means close to neutrality. Because of charge complementarity, however, **strong** acids and bases are preferred in the antibody, which are less suitable as general acid/base catalysts.

An additional problem caused by the modest activities of catalytic antibodies is the difficulty in detecting catalytic activity when traces of enzymes which catalyze the same reaction are present in the antibody producing cell or the supernatant. This problem is particularly cumbersome in the case of nucleases and proteases.

Out of these considerations, methods were developed in the author's laboratory aimed at facilitating the modification of the catalytic antibody itself through the methods of protein engineering[70-72]. While the methods for modifying DNA sequences had been well established, the expression (i. e. the biosynthesis from a recombinant gene) of genetically engineered antibodies could in the past only be achieved with large effort. A system was developed to permit the production of fully functioned antibody F_v or F_{ab} fragments in bacteria (*Escherichia coli*) (Fig. 16, 17)[70-72]. The method is based on the expression of both chains in the same cell and the secretion of both proteins into the periplasmic space between the two membranes. There the disulfide bonds form in an oxidizing environment and the two domains V_L and V_H assemble. The functional protein can be purified by affinity chromatography with immobilized hapten (antigen) in a single step.

For these investigations, the phosphorylcholine binding antibody with the designation McPC603 was used. Its main attraction was that its three-dimensional structure was

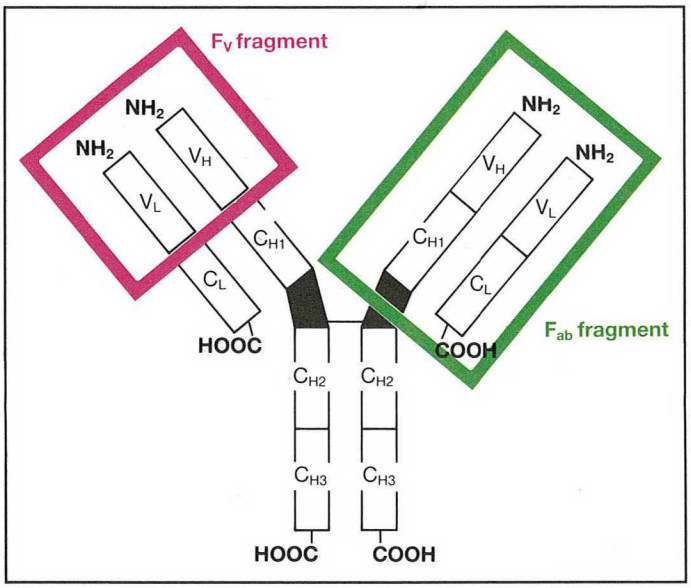


Figure 16: Schematic diagram of an antibody and the relevant antigen-binding fragments. The F_v fragment contains the variable domains of the heavy (V_H^v) and light (V_L)chains. The F_{ab} fragment contains the complete light chain and the first two domains from the heavy chain.

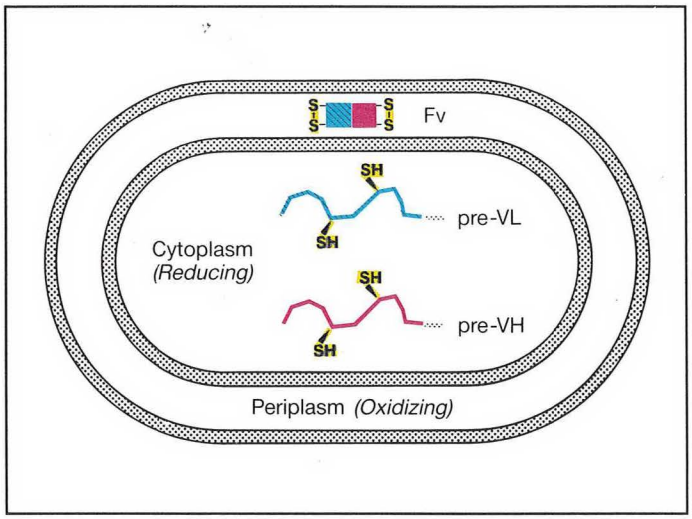


Figure 17: Schematic diagram of the *E. coli* expression strategy for functional antigen-binding fragments. In the cytoplasm the precursor proteins for V_H and V_L, each fused to a bacterial signal sequence, are synthesized in reduced form. After translocation through the inner membrane into the periplasm, the signal sequences are cleaved, the domains fold, the disulfide bonds form, and the two chains assemble into the functional F_V fragment. Expression of the F_{ab} fragment is entirely analogous.

solved with and without bound hapten[77]. The two genes for the F_v fragment were obtained entirely synthetically[70], because the amino acid sequence was known. It was then shown in detailed investigations that the F_v fragment of this antibody from *E. coli* has the same affinity to phosphorylcholine as the entire antibody from the mouse[71, 72]. Thus the protein required to bind the antigen can be drastically reduced in size.

Since this antibody binds phosphorylcholine, it seemed reasonable to suppose that it would be able to cleave an ester bond (Fig. 18). If a suitable ester is attacked by water (or a hydroxide ion), a tetrahedral intermediate is formed. The transition state to (and from) this intermediate, according to *Hammond*'s postulate mentioned in the beginning, will also be roughly tetrahedral. If now the antibody binds this type of structure preferentially, i. e. better than the substrate, it should catalyze this hydrolysis.

It was indeed shown that the recombinant F_v fragment from *E. coli* is able to do this[72] as had been found for related antibodies obtained from mouse[78]. Though the observed catalysis is only moderate, this model system opens up interesting perspectives. First, it is now possible to make any desired modification to the sequence of this antibody, permitting a systematic investigation of structural effects on catalysis. Second, the structure of the binding site is known[77] and the recombinant V_L domain produced in *E. coli*[79] was recently crystallized and its structure determined[79], so that

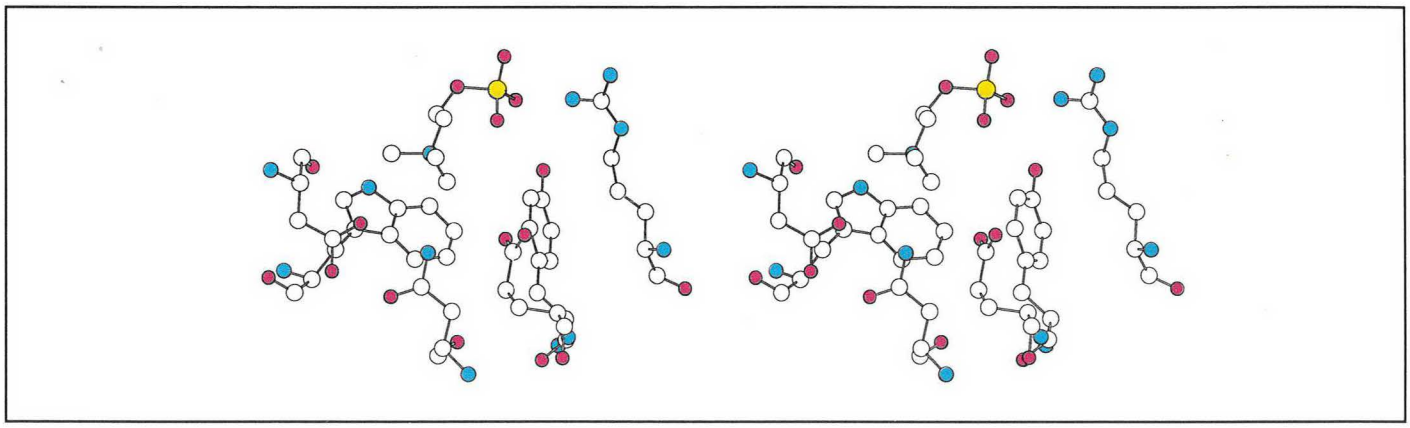


Figure 18A: Stereo diagram of the binding pocket of antibody McPC603 with bound hapten phosphorylcholine. Atoms are shaded according to atom types, with oxygen in red, nitrogen in blue, carbon in white and phosphorus in yellow. The residues shown are, from left to right: AspL97, TrpH107, AsnH101, phosphorylcholine, GluH35, TyrH33 and ArgH52.

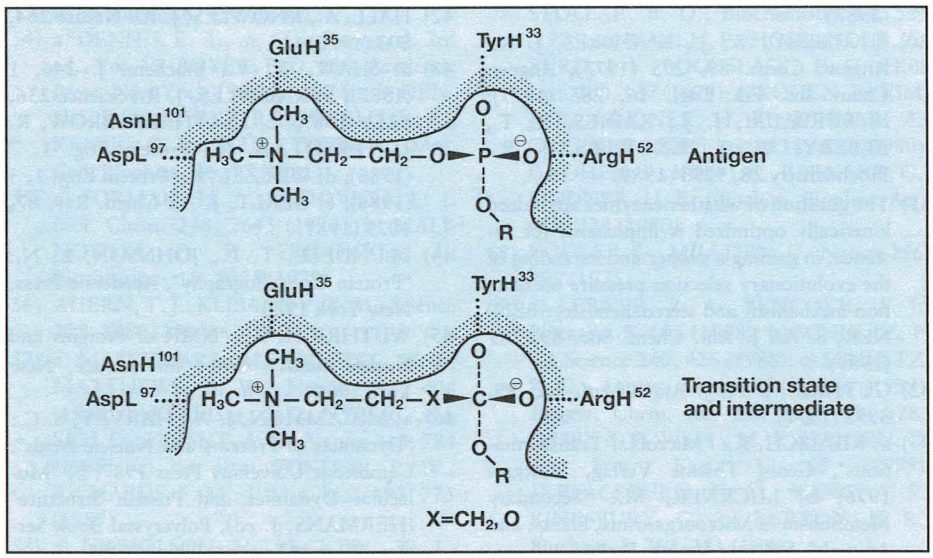


Figure 18B: Schematic diagram showing the complexation of the tetrahedral intermediate during ester-hydrolysis in the binding pocket of the antibody and its analogy to normal antigen binding.

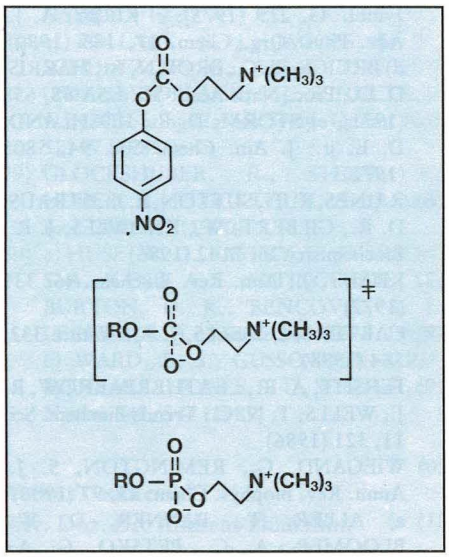


Figure 18C: The substrate used (top) and its tetrahedral intermediate (middle) and the antigen (bottom).

information about the structure of the modified fragments is now available. Third, the fragment is of a size that makes it amenable to structure analysis by NMR. Last, expression in the native, functional state as the prerequisite for metabolic selection or screening has been achieved. Many of these findings and methods will be generally applicable to catalytic antibodies. This bacterial system may eventually even be used to express libraries from the entire immunological repertoire of mouse or man, and there are encouraging results toward this goal[80]. It might then be possible one day to select catalytic antibodies without the need for mouse immunization.

10. Prospects

It is probable that, along the way to new enzymes, all of these strategies will need to be combined. The interdisciplinary character of this research, in which enzymology, gene technology, immunology, organic chemistry, theoretical chemistry and, in particular, structural research come together, is apparent. Enzymes have by no means given up all their secrets – just a few of them. Without doubt, enzyme based catalysts will strongly influence the chemistry of the future.

References

1) FISCHER, E.: "Nobel Lectures – Chemistry 1901–1921", Elsevier Publ. Co., Amsterdam 1966, pp. 34
2) a) FERSHT, A.: "Enzyme Structure and Mechanism", Freeman and Co., San Francisco 1985; b) WALSH, C.: "Enzymatic Reaction Mechanisms", Freeman and Co., San Francisco 1979; c) HAMMES, G. G.: "Enzyme Catalysis and Regulation", Academic Press, Orlando 1982; d) JENCKS, W. P.: "Catalysis in Chemistry and Enzymology", McGraw Hill, New York 1969; e) GANDOUR, R. D., SCHOWEN, R. L. (ed.): "Transition States of Biochemical Processes", Plenum Press, New York 1978; f) BENDER, M. L., BERGERON, R. J., KOMIYAMA, M.: "The Bioorganic Chemistry of Enzymatic Catalysis", Wiley and Sons, New York 1984; g) DUGAS, H.: "Bioorganic Chemistry", Springer Verlag, Berlin 1981; h) KRAUT, J.: Science **242**, 533 (1988)
3) FISCHER, E.: Ber. Dtsch. Chem. Ges. **27**, 2985 (1894)
4) HALDANE, J. B. S.: "Enzymes", Longman, Green and Co., London 1930; MIT Press, Cambridge, Mass. 1965
5) a) PELZER, H., EIGNER, E.: Z. Phys. Chem. **B 15**, 445 (1932); b) EYRING, H.: Chem. Rev. **17**, 65 (1935); c) EVANS, M. G., POLANYI, M.: Trans. Faraday Soc. **31**, 875 (1935)
6) a) MOORE, W. J.: "Physical Chemistry", Prentice Hall, Englewood Cliffs 1972; b) LAIDLER, K. J.: "Chemical Kinetics", McGraw Hill, New York 1970; c) HAMMET, L. P.: "Physical Organic Chemistry", McGraw Hill, New York 1970; d) LOWRY, T. H., RICHARDSON, K. S.: "Mechanism and Theory in Organic Chemistry", Harper and Row, New York 1976
7) Discrimination of the transition state as a fundamental element of catalysis was proposed even earlier: POLANYI, M.: Z. Elektrochem. **27**, 143 (1921)
8) PAULING, L.: Chem. Eng. News **24**, 1375 (1946)
9) WOLFENDEN, R.: Annu. Rev. Biophys. Bioeng. **5**, 271 (1976)
10) HAMMOND, G. S.: J. Am. Chem. Soc. **77**, 334 (1955)
11) a) BLACKLOW, S. C., HERMES, J. D., KNOWLES, J. R.: 39th Colloquium Mosbach 1988, WINNACKER, E.-L., HUBER, R. (eds.), p. 59 ff.; b) ALBERY, W. J., KNOWLES, J. R.: Biochemistry **25**, 2572 (1986) as well as the immediately preceding six articles by KNOWLES, J. R., et al.
12) EIGEN, M.: Angew. Chem. **75**, 489 (1963)
13) a) WARSHEL, A., ÅQUIST, J., CREIGHTON, S.: Proc. Natl. Acad. Sci. USA **86**, 5820 (1989); b) for a contrary view, see: DEWAR, M. J. S., STORCH, D. M.: Proc. Natl. Acad. Sci. USA **82**, 2225 (1985)
14) BÜRGI, A. B., DUNITZ, J. D.: Acc. Chem. Res. **16**, 163 (1983)
15) a) MENGER, F. M.: Acc. Chem. Res. **18**, 128 (1985); b) JENCKS, W. P.: Adv. En-

zymol. **43**, 219 (1975); c) KIRBY, A. J.: Adv. Phys. Org. Chem. **17**, 183 (1980); d) BRUICE, T. C., BROWN, A., HARRIS, D. O.: Proc. Natl. Acad. Sci. USA **68**, 658 (1971); e) STORM, D. R., KOSHLAND, D. E. jr.: J. Am. Chem. Soc. **94**, 5805 (1972)

16) RAINES, R. T., SUTTON, E. L., STRAUS, D. R., GILBERT, W., KNOWLES, J. R.: Biochemistry **25**, 7142 (1986)

17) KRAUT, J.: Ann. Rev. Biochem. **46**, 331 (1977)

18) CARTER, P., WELLS, J. A.: Nature **332**, 564 (1988)

19) FERSHT, A. R., LEATHERBARROW, R. J., WELLS, T. N. C.: Trends Biochem. Sci. **11**, 321 (1986)

20) WIEGAND, G., REMINGTON, S. J.: Annu. Rev. Biophys. Chem. **15**, 97 (1986)

21) a) ALBER, T., BANNER, D. W., BLOOMER, A. C., PETSKO, G. A., PHILLIPS, D., RIVERS, P. S., WILSON, I. A.: Philos. Trans. R. Soc. London, Ser. B **293**, 159 (1981); b) GORENSTEIN, D. G.: Chem. Rev. **87**, 1047 (1987)

22) HUBER, R., BENNETT, W. S.: Biopolymers **22**, 261 (1983)

23) a) KELL, D. B.: Trends Biochem. Sci. **7**, 349 (1982); b) WELCH, G. R., SOMOGYI, B., DAMJANOVICH, S.: Progr. Biophys. Mol. Biol. **39**, 109 (1982); c) SRERE, P. A.: Trends Biochem. Sci. **9**, 387 (1984)

24) a) BRESLOW, R.: Science **218**, 532 (1982); b) BRESLOW, R.: Adv. Enzymol. **58**, 1 (1986); c) SCHMIDTCHEN, F. P.: Top. Curr. Chem. **132**, 101 (1986); d) REBEK, J.: Science **235**, 1478 (1987); e) CRAM, D. J.: Angew. Chem. **100**, 1041 (1988); Angew. Chem. Int. Ed. Engl. **27**, 1009 (1988); f) LEHN, J. M.: Angew. Chem. **100**, 92 (1988); Angew. Chem. Int. Ed. Engl. **27**, 89 (1988); g) DIEDERICH, F.: Angew. Chem. **100**, 372 (1988); Angew. Chem. Int. Ed. Engl. **27**, 362 (1988)

25) This is actually a contradiction in terms. Though many enzyme models do not catalyze a true turnover, they do in fact provide important insights into enzyme mechanisms and are known in the literature as model enzymes. Hence the same name here.

26) MOCK, W. L., IRRA, T. A., WEPSIEC, J. P., ADHYA, M.: J. Org. Chem. **54**, 5302 (1989)

27) a) ANFINSEN, C. B.: Science **181**, 223 (1973); b) CREIGHTON, T. E.; Progr. Biophys. Mol. Biol. **33**, 231 (1978); c) JAENICKE, R.: Progr. Biophys. Mol. Biol. **49**, 117 (1987)

28) HOL, W. G. J.: Progr. Biophys. Mol. Biol. **45**, 149 (1985)

29) SRIVASTA, D. K., BERNHARD, S. A.: Ann. Biophys. Biophys. Chem. **16**, 175 (1987)

30) a) ALBERY, W. J., KNOWLES, J. R.: Angew. Chem. **89**, 295 (1977); Angew. Chem. Int. Ed. Engl. **16**, 285 (1977); b) BURBAUM, J. J., RAINES, R. T., ALBERY, W. J., KNOWLES, J. R.: Biochemistry **28**, 9293 (1989)

31) The question of whether enzymes are indeed kinetically optimized is important, for instance, in gaining a proper understanding of the evolutionary selection pressure on reaction mechanism and stereochemistry: BENNER, S. A.: J. Am. Chem. Soc. **89**, 789 (1989)

32) GUTHRIE, J. P.: J. Am. Chem. Soc. **99**, 3991 (1977)

33) a) KIESLICH, K.: "Microbial Transformations", Georg Thieme Verlag, Stuttgart 1976; b) LUCKNER, M.: "Secondary Metabolism in Microorganisms, Plants, and Animals", Springer Verlag, Berlin 1984

34) a) JAENICKE, R.: Ann. Rev. Biophys. Bioeng. **10**, 1 (1981); b) JAENICKE, R.: Forum Mikrobiol. **10**, 435 (1988)

35) Regarding the stability of proteins, see also: BALDWIN, R. L., EISENBERG, D., in: "Protein Engineering" (OXENDER, D. L., FOX, C. F. eds.), Alan Liss, New York 1987, pp. 127 – 148

36) a) "Microbes in extreme environments" (HERBERT, R. A., CODD, G. A. eds.), Academic Press, London 1986; b) "Thermophiles" (BROCK, T. D. ed.), Wiley-Interscience, New York 1986

37) a) MATSUMURA, M., AIBA, S.: J. Biol. Chem. **260**, 15298 (1985); b) CLARKE, P. H., DREW, R.: Bioscience Rep. **8**, 103 (1988); c) EVNIN, L. B., CRAIK, C. S.: Ann. N. Y. Acad. Sci. **542**, 61 (1988)

38) a) MILLER, J. H.: "Experiments in Molecular Genetics", Cold Spring Harbor, New York 1972; b) SMITH, M.: Ann. Rev. Genet. **19**, 423 (1985)

39) a) GOFF, S. A., SHORT-RUSSEL, S. R., DICE, J. F.: DNA **6**, 381 (1987); b) HÜBNER, P., IIDA, S., ARBER, W.: Gene **73**, 319 (1988); c) NER, S. S., GOODIN, D. B., SMITH, M.: DNA **7**, 127 (1988); d) HERMES, J. D., PAREKH, S. M., BLACKLOW, S. C., KÖSTER, H., KNOWLES, J. R.: Gene **84**, 143 (1989)

40) CUNNINGHAM, B. C., WELLS, J. A.: Protein Eng. **1**, 319 (1987)

41) a) LIAO, H., McKENZIE, T., HAGEMAN, R.: Proc. Natl. Acad. Sci. USA **83**, 576 (1986); b) ALBER, T., WOZNIAK, J. A.: Proc. Natl. Acad. Sci. USA **82**, 747 (1985); c) BRYAN, P., ROLLENCE, M., WOOD, J., DODD, S., WHITLOW, M., HARDMAN, K., PANTOLIANO, M.: J. Cell Biochem. **13A**, 66 (1989); d) PAKULA, A. A., SAUER, R. T.: Proteins **5**, 202 (1989)

42) HALL, A., KNOWLES, J. R.: Nature **264**, 803 (1976)

43) a) SHAW, W. V.: Biochem. J. **246**, 1 (1987); b) KNOWLES, J. R.: Science **236**, 1252 (1987); c) LEATHERBARROW, R. J., FERSHT, A. R.: Protein Eng. **1**, 7 (1986); d) WETZEL, R.: Protein Eng. **1**, 3 (1986); e) GERLT, J. A.: Chem. Rev. **87**, 1079 (1987)

44) BLUNDELL, T. L., JOHNSON, L. N.: "Protein Crystallography", Academic Press, New York 1976

45) WÜTHRICH, K.: "NMR of Proteins and Nucleic Acids", Wiley Interscience, New York 1986

46) a) McCAMMON, J. A., HARVEY, S. C.: "Dynamics of Proteins and Nucleic Acids", Cambridge University Press 1987; b) "Molecular Dynamics and Protein Structure" (HERMANS, J. ed), Polycrystal Book Service, Western Springs, Illinois 1985

47) a) CHOTHIA, C., LESK, A. M., LEVITT, M., AMIT, A. G., MARIUZZA, R. A., PHILLIPS, S. E. V., POLJAK, R. J.: Science **233**, 755 (1986); b) FINE, R. M., WANG, H., SHENKIN, P. S., YARMUSH, D. L., LEVINTHAL, C.: Proteins **1**, 342 (1986); c) SHENKIN, P. S., YARMUSH, D. L., FINE, R. M., WANG, H., LEVINTHAL, C.: Biopolymers **26**, 2053 (1987); d) BRUCCOLERI, R. E., KARPLUS, M.: Biopolymers **26**, 137 (1987)

48) a) van GUNSTEREN, W. F.: Protein Eng. **2**, 5 (1988); b) KOLLMAN, P., van GUNSTEREN, W. F.: Methods Enzymol. **154**, 430 (1987)

49) These brief simulations may well be sufficient to achieve an approximate assessement of the difference in free energy between two (very similar) states, e. g. the differences in binding energy between two very similar variants of an inhibitor or of an enzyme. Various approaches are summarized in: van GUNSTEREN, W. F., WEINER, P. K.: "Computer Simulations of Biomolecular Systems", Escom Science Publishers, Leiden 1989

50) BOTT, R., ULTSCH, M., WELLS, J., POWERS, D., BURDICK, D., STRUBLE, M., BURNIER, J., ESTELL, D., MILLER, J., GRAYCAR, T.: ACS Symp. Series **334**, 139 (1987)

51) a) MATTHEWS, B. W.: Biochemistry **26**, 6885 (1987); b) ALBER, T.: Ann. Rev. Biochem. **58**, 765 (1989)

52) WILKS, H. M., HART, K. W., FEENEY, R., DUNN, C. R., MUIRHEAD, H., CHIA, W. N., BARSTOW, D. A., ATKINSON, T., CLARKE, A. R., HOLBROOK, J. J.: Science **242**, 1541 (1988)

53) a) RUSSELL, A. J., THOMAS, P. G., FERSHT, A. R.: J. Mol. Biol. **193**, 803 (1987); b) RUSSELL, A. J., FERSHT, A.

R.: Nature **328**, 496 (1987)
54) a) DENNIS, E. A., in: "The Enzymes", 3rd ed., Vol. 16 (BOYER, P. D. ed.), Academic Press, New York 1983, p. 307 ff.; b) DENNIS, E. A., DARKE, P. L., DEEMS, R. A., KENSIL, C. R., PLÜCKTHUN, A.: Mol. Cell. Biochem. **36**, 37 (1981)
55) a) FORMAN, H. J., FRIDOVICH, I.: J. Biol. Chem. **248**, 2645 (1973); b) MALINOWSKI, D. P., FRIDOVICH, I.: Biochemistry **18**, 5055 (1979)
56) AHERN, T. J., KLIBANOV, A. M.: Science **228**, 1280 (1985)
57) a) MATSUMARA, M., BECKTEL, W. J., MATTHEWS, B. W.: Nature **334**, 406 (1988); b) KELLIS jr., J. T., NYBERG, K., SALI, D., FERSHT, A. R.: Nature **333**, 784 (1988); c) see also: PONDER, J. W., RICHARDS, F. M.: J. Mol. Biol. **193**, 775 (1987)
58) a) NICHOLSON, H., BECKTEL, W. J., MATTHEWS, B. M.: Nature **335**, 651 (1988); b) SALI, D., BYCROFT, M., FERSHT, A. R.: Nature **336**, 740 (1988)
59) ALBER, T., DAO-PIN, S., WILSON, K., WOZNIAK, J. A., COOK, S. P., MATTHEWS, B. W.: Nature **330**, 41 (1987)
60) a) MATTHEWS, B. W., NICHOLSON, H., BECKTEL, W. J.: Proc. Natl. Acad. Sci. USA **84**, 6663 (1987); b) see also: HECHT, M. H., STURTEVANT, J. M., SAUER, R. T.: Proteins **1**, 43 (1986)
61) GANTER, C., PLÜCKTHUN, A.: Biochemistry, in press
62) ESTELL, D. A., GRAYCAR, T. P., WELLS, J. A.: J. Biol. Chem. **260**, 6518 (1985)
63) a) WETZEL, R., PERRY, L. J., BAASE, W. A., BECKTEL, W. J.: Proc. Natl. Acad. Sci. USA **85**, 401 (1988); b) WETZEL, R.: Trends Biochem. Sci. **12**, 478 (1987)
64) a) MITCHINSON, C., WELLS, J. A.: Biochemistry **28**, 4807 (1989); b) WELLS, J. A., POWERS, D. B.: J. Biol. Chem. **261**, 6564 (1986); c) PANTOLIANO, M. W., LADNER, R. C., BRYAN, P. N., ROLLENCE, M. L., WOOD, J. F., POULOS, T. L.: Biochemistry **26**, 2077 (1987); d) PERRY, L. J., WETZEL, R.: Science **226**, 555 (1984)
65) a) GLOCKSHUBER, R., MALIA, M., PFITZINGER, I., PLÜCKTHUN, A.: Biochemistry **29**, 1326 (1990); b) SAUER, R. T., HEHIR, K., STEARMAN, R. S., WEISS, M. A., JEITLER-NILSSON, A., SUCHANEK, E. G., PABO, C. O.: Biochemistry **25**, 5992 (1986)
66) JENCKS, W. P.: "Catalysis in Chemistry and Enzymology", McGraw Hill, New York 1969, p. 288
67) a) SLOBIN, L. I.: Biochemistry **5**, 2836 (1966); b) RASO, V., STOLLAR, B. D.: Biochemistry **14**, 584 (1975); c) RASO, V., STOLLAR, B. D.: Biochemistry **14**, 591 (1975); d) KOHEN, F., HOLLANDER, Z., BURD, J. F., BOGUSLASKI, R. C.: FEBS Lett. **100**, 137 (1979); e) KOHEN, F., KIM, J. B., LINDNER, H. R., ESHHAR, Z., GREEN, B.: FEBS Lett. **111**, 427 (1980); f) KOHEN, F., KIM, J. B., BARNARD, G., LINDNER, H. R.: Biochim. Biophys. Acta **629**, 328 (1980)
68) KÖHLER, G., MILSTEIN, C.: Nature **256**, 495 (1975)
69) a) LERNER, R. A., BENCOVIC, S. J.: Bioessays **9**, 107 (1988); b) SCHULTZ, P. G.: Science **240**, 426 (1988); c) SCHULTZ, P. G.: Angew. Chem. **101**, 1336 (1989); Angew. Chem. Int. Ed. Engl. **28**, 1283 (1989); d) POWELL, M. J., HANSEN, D. E.: Prot. Eng. **3**, 69 (1989); e) BLACKBURN, G. M., KANG, A. S., KINGSBURY, G. A., BURTON, D. R.: Biochem. J. **262**, 381 (1989)
70) PLÜCKTHUN, A., GLOCKSHUBER, R., PFITZINGER, I., SKERRA, A., STADLMULLER, J.: Cold Spring Harbor Symp. Quant. Biol. **52**, 105 (1987)
71) a) SKERRA, A., PLÜCKTHUN, A.: Science **240**, 1038 (1988); b) PLÜCKTHUN, A., SKERRA, A.: Methods Enzymol. **178**, 497 (1989)
72) PLÜCKTHUN, A., GLOCKSHUBER, R., SKERRA, A., STADLMULLER, J.: Behring Inst. Mitt. in press
73) a) JACKSON, D. Y., JACOBS, J. W., SUGASAWARA, R., REICH, S. A., BARTLETT, P. A., SCHULTZ, P. G.: J. Am. Chem. Soc. **110**, 4841 (1988); b) HILVERT, D., CARPENTER, S. H., NARED, K. D., AUDITOR, M. T. M.: Proc. Natl. Acad. Sci. USA **85**, 4953 (1988)
74) a) COPLEY, S. D., KNOWLES, J. R.: J. Am. Chem. Soc. **107**, 5306 (1985); b) ADDADI, K., JAFFE, E. K., KNOWLES, J. R.: Biochemistry **22**, 4494 (1983); c) COATES, R. M., ROGERS, B. D., HOBBS, S. J., PECK, D. R., CURRAN, D. P.: J. Am. Chem. Soc. **109**, 1160 (1987); d) SOGO, S. G., WIDLANSKI, T. S., HOARE, J. H., GRIMSHAW, C. E., BERCHTOLD, G. A., KNOWLES, J. R.: J. Am. Chem. Soc. **106**, 2701 (1984); e) GUILFORD, W. J., COPLEY, S. D., KNOWLES, J. R.: J. Am. Chem. Soc. **109**, 5013 (1987)
75) SHOKAT, K. M., LEUMANN, C. J., SUGASAWARA, R., SCHULTZ, P. G.: Nature **338**, 269 (1989)
76) IVERSON, B. L., LERNER, R. A.: Science **243**, 1184 (1989)
77) a) SEGAL, D. M., PADLAN, E. A., COHEN, G. A., RUDIKOFF, S., POTTER, M., DAVIES, D. R.: Proc. Natl. Acad. Sci. USA **71**, 4298 (1974); b) SATOW, Y., COHEN, G. H., PADLAN, E. A., DAVIES, D. R.: J. Mol. Biol. **190**, 593 (1986)
78) a) POLLACK, S. J., JACOBS, J. W., SCHULTZ, P. G.: Science **234**, 1570 (1986); b) POLLACK, S. J., SCHULTZ, P. G.: Cold Spring Harbor Symp. Quant. Biol. **52**, 97 (1987)
79) GLOCKSHUBER, R., STEIPE, B., HUBER, R., PLÜCKTHUN, A.: J. Mol. Biol. **213**, 613 (1990)
80) a) HUSE, W. D., SASTRY, L., IVERSON, S. A., KANG, A. S., ALTING-MEES, M., BURTON, D. R., BENCOVIC, S. J., LERNER, R. A.: Science **246**, 1275 (1989); b) WARD, E. S., GÜSSOW, D., GRIFFITHS, A. D., JONES, P. T., WINTER, G.: Nature **341**, 544 (1989)

Author's address:
Priv. Doz. Dr. Andreas Plückthun
Gen-Zentrum University of Munich
Max-Planck-Institut for Biochemistry
Am Klopferspitz
D-8033 Martinsried/München
FRG