So do we understand how enzymes work?
David Blow

Between 1930 and 1975 biochemical and structural analysis of enzymes led to a clear set of ideas that might form a basis for detailed understanding of enzyme action. Further development required energetic and thermodynamic analysis of enzymes in an aqueous medium, beyond the computational power then available. Structural enzymology advanced in other directions, but the fundamental questions of enzyme action must soon be re-opened.

Address: Blackett Laboratory, Imperial College, London SW7 2BZ, UK.

E-mail: d.blow@ic.ac.uk

Figure 1

Schematic representation of the mechanism of a proteolytic enzyme based on Haldane's diagrams [2]. Parts **(a)**, **(b)**, **(e)** and **(f)** closely follow Haldane's original. Parts **(c)** and **(d)** suggest how strain might affect the substrate in intermediate stages. Colours are used to indicate physical stress on the enzyme – yellow and red for tension, blue and purple for compression.
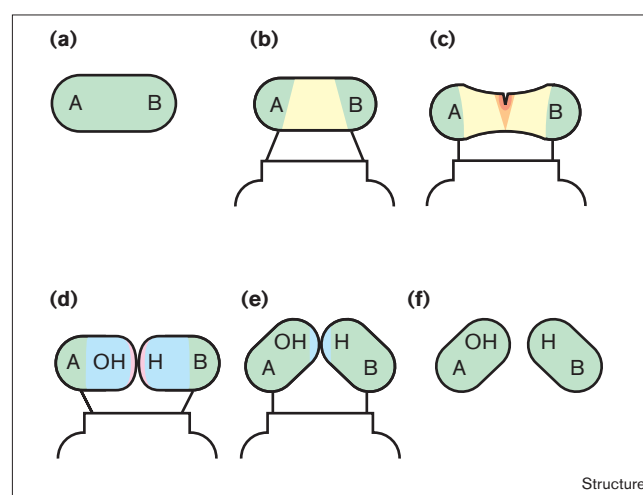
Our understanding of enzymes has developed incredibly during my lifetime. Friedrich Hunefeld crystallised haemoglobin 160 years ago. But when James Sumner crystallised jack bean urease in 1926, the biochemical establishment refused to believe that the crystals represented the enzyme. The 1929 edition of the Encyclopaedia Britannica, recognised as a classic edition for the erudition of its contributors and the accuracy of its articles, states in the article on enzymes 'The chemical structure of enzymes is not yet known. They were formerly regarded as proteins... [but are now conceived as]... a colloidal carrier and a purely chemical active group'.

This error manifests the extraordinary catalytic power of enzymes. Purified preparations containing no detectable protein (by methods then available) possessed obvious enzymatic activity. Richard Willstätter, doyen of enzyme chemists in the 1920's, concluded that the protein ('colloidal') component was a non-specific carrier of adsorbed chemical catalytic agents which were the true enzymes [1].

The crystallisation of several hydrolytic enzymes by John Northrop and Moses Kunitz in the early years of my life resolved the question beyond reasonable doubt. Enzymes are proteins. In those same years, Desmond Bernal, with Dorothy Crowfoot and Max Perutz, found they could get good X-ray diffraction patterns from protein crystals, including chymotrypsin and pepsin, if they kept them moist. The unit cells were huge.

JBS Haldane's book 'Enzymes' (1930) [2] maintains a careful neutrality about the chemical nature of enzymes, but presents thought-provoking insight into their mode of action. An enzyme might catalyse a hydrolytic reaction if the substrate components 'when combined with the enzyme, lay slightly further apart than their equilibrium distance when combined... but nearer than their equilibrium distance when free. ...Using Fischer's lock and key simile, the key does not fit the lock perfectly but exercises a certain strain on it.' Figure 1 is a slightly developed version of Haldane's stimulating diagram.

Alfred Mirsky and Linus Pauling [3] described proteins with startling modernity in 1936. A native protein molecule, they said, 'consists of one... [or more] polypeptide chains which continue without interruption throughout the molecule... folded into a uniquely defined configuration... held together by hydrogen bonds between the peptide nitrogen and oxygen atoms and also between [charged sidechains]... The denatured protein molecule we consider to be characterized by the absence of a uniquely defined conformation.' Despite the accuracy of this insight, the question as to whether a protein was a polypeptide was argued for another 20 years.

Ten years later Pauling [4] wrote reflecting Haldane's idea of enzyme action, now based on this view of protein structure. '...The only reasonable picture of the catalytic

activity of enzymes is that which involves an active region of the surface of the enzyme which is closely complementary in structure not to the substrate molecule itself in its normal configuration, but rather to the substrate molecule in a strained conformation corresponding to the 'activated complex' for the reaction catalyzed by the enzyme.'

During the 1950's the work of Pauling, Perutz, Fred Sanger, John Kendrew and their colleagues opened the door to detailed analysis of enzyme molecules by introducing techniques for chemical and crystallographic analysis of proteins. During the first 30 years of my life, enzyme molecular science had advanced from a poorly defined subject, full of amazing observation and fascinating hypothesis but confused and lacking basic data, to be ready for accurate study on the atomic scale. I should exclude from the vagueness of much enzyme study up to 1930, the successful analysis of enzyme kinetics following the chemical principles of thermodynamics and mass action, pioneered by Leonor Michaelis. There were no serious doubts that standard thermodynamics applied to enzyme catalysis. But many of us were still unable to imagine how the trick could possibly be done.

In 1960, Roberto Poljak grew beautiful crystals of lysozyme at the Massachusetts Institute of Technology and took them to David Phillips' laboratory at the Royal Institution in London. They provided the subject for the first structure determination of an enzyme. In 1965, David Phillips' group [5] revealed a structure of lysozyme at 2 Å resolution. Phillips died in 1999, and it is appropriate to recognise his contribution. Not only was this a major technical achievement in crystallography, but the clarity of the immediate interpretation was immensely important. Three important features were emphasised: an enzyme surface with a cleft shaped to accept the substrate; polar groups (in this case acid groups) ready to interact at the site of chemical catalysis; and a mode of substrate binding that favoured distortion of the substrate towards a transition configuration (in this case a hexose ring twisted towards the half-chair conformation). Even if there is still controversy surrounding the exact status of this distortion, these three features, and the clarity with which they were presented in 1965, stimulated a wide range of insightful studies into enzyme action.

In 1967, the structures of several hydrolytic enzymes were revealed by crystallography. These were structures of carboxypeptidase A (by Bill Lipscomb's group at Harvard), two independent determinations of ribonuclease A (by Dave Harker's group in New York and by Bernal and Carlisle in London) and α-chymotrypsin (by our group in Cambridge).

Because of errors in amino acid sequence determination, and difficulties in observing bound substrate analogues in crystals, it took us much longer to understand the action of chymotrypsin [6]. But by 1969, the work especially of Brian Matthews, Paul Sigler, Jens Birktoft and Richard Henderson, assisted by redeterminations of the amino acid sequence by Brian Hartley, told the story, which was complemented by the structure of subtilisin from Joe Kraut's group in Seattle [7–9]. The wide cleft at the active site is supplemented by a tight pocket mediating chymotrypsin's specificity for aromatic sidechains, and two hydrogen-bond donors positioned to hold the carbonyl oxygen of the scissile peptide bond. We noted a histidine interacting with the serine at the acylation site, polarised by a buried acid group. These features indicate a chemical mechanism for the acylation and deacylation steps that comprise the action of chymotrypsin, and position the scissile bond precisely in the position required. We saw that reaction steps through a tetrahedral intermediate require only small movements of any particle bigger than a proton [10]. There was no evidence of physical distortion of the substrate, and as there was only weak binding of the part representing the 'leaving group' it was hard to see how the scissile bond could be physically distorted.

In the early 1970's, Huber's group and our own [11,12] determined structures for specific trypsin inhibitors and their complexes with trypsin. These studies confirmed our ideas of how the substrate binds. Both groups tried hard to find evidence for distortion of the scissile bond from the normal conformation of a peptide bond, but in retrospect the evidence was weak.

This was all very well, but how did these structures explain the huge acceleration of the hydrolysis of specific peptide bonds? Comparing the unimolecular rate constant of uncatalysed hydrolysis with the bimolecular rate k_{cat}/K_M, the rate enhancement was equivalent to hydrolysis in a water concentration of 10^8 M. What does that mean? How could it be done?

Richard Henderson [13] studied a form of chymotrypsin in which the active-site histidine was methylated, preventing it from exerting its polarising effect on the active-site serine; the structure of this form was determined by Christine Wright and George Hess [14]. This modified enzyme was less efficient than chymotrypsin, but not enough. The difference would account for no more than 10^4 of the 10^8 M. Much later Charles Craik [15] studied a form of chymotrypsin in which the buried acid group (Asp102) that polarises the histidine was replaced by an uncharged asparagine sidechain. The loss of catalytic power of this enzyme was also about 10^4. On the other hand, Joe Kraut, with Steve Freer and others [16], determined the structure of the zymogen chymotrypsinogen. All the identified components for chemical catalysis were present, although part of the specific binding site was disordered. But the zymogen is not just weakly active, it is totally inactive, at

least 10^7 times less active than chymotrypsin. The hydrogen bonds that we had identified as orienting the scissile peptide bond cannot be formed. Kraut pointed out that this feature, which he termed the oxyanion hole, also assists in polarising the carbon–nitrogen bond that the enzyme is going to break [17].

Alan Fersht introduced methods to measure the rates of individual reaction steps using stopped quenched flow techniques. Using the wealth of kinetic data already available for chymotrypsin hydrolysis, a detailed picture of the rate-limiting steps and specificity-determining factors in chymotrypsin hydrolysis could be built [18].

William Lawson [19] produced a fascinating series of 'locked' substrates for chymotrypsin, based on hydrocinnamate esters. Some were excellent substrates but others, only marginally different in shape and chemistry, were several orders of magnitude less active. Phillip Rodgers [20] determined the structure of one of these compounds, and showed that a positional deviation of less than 0.2 Å of the α-carbon atom changes activity by several orders of magnitude. Dan Koshland [21] introduced the idea of 'orbital steering' to suggest that a precise orientation of a substrate to a reactive group was the key to enzyme activity, but the explanation appeared to fail. Either the achievable concentration of reactant was insufficient or the orientation demanded greater precision than is required for ordinary chemical catalysis in order to explain the enzyme's reactivity.

I was profoundly uneasy. We knew so much, but it did not add up to understanding. The many factors that could be identified were not sufficient to explain the observed catalytic rates. Then, at a Harden conference I organised in 1974, came the turning point. Bill Jencks described his work with Mike Page, which to me at least created a revelation.

Page and Jencks [22] presented their work thermodynamically, and demonstrated that the entropy loss on binding substrate to an enzyme provided sufficient free energy for a 10^8 acceleration rate. However, the essence of this thermodynamic argument can be stated in very simple terms. Haldane's diagram omits any chemical machinery for bringing about the required reaction. If this is added (Figure 2), it is obvious that the substrate must be bound in a very specific way; it needs to be immobilised with respect to the catalytic chemical groups, and oriented with great precision.

The notion of 'strain', illustrated by Haldane as physical distortion, means that the reactive groups are locally in a state of higher energy, reducing the barrier to the chemical transition. The strain does not need to be geometric, as was often assumed. In the hydrolytic enzymes it is usually a distortion of the electron distribution around the site of
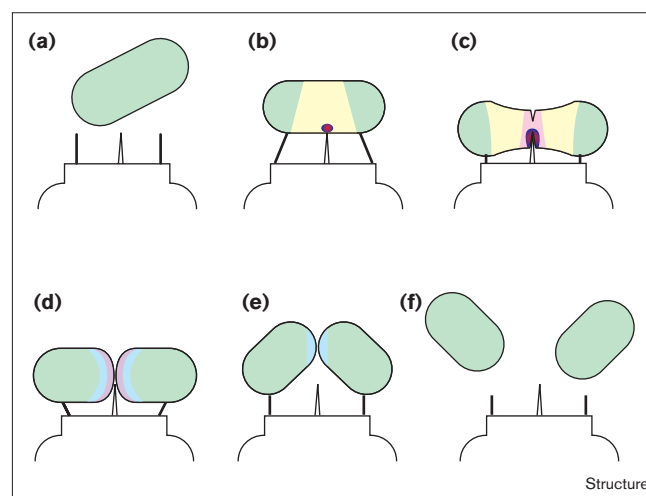
Figure 2

Figure 1 is redrawn to include the catalytic machinery for hydrolysis; **(a–f)** represent the different steps of the mechanism. There is negligible chance of the hydrolytic machinery attacking the substrate unless it is firmly bound. It is also necessary for the substrate to be precisely positioned. The additional order imposed on the system then makes the transition highly probable. Part of the free energy of substrate binding is needed to create this additional order.

catalytic change. The electron distribution is distorted towards the transition state.

Jencks [23] has summarised these discoveries as follows. 'Enzymes have evolved... to bind reacting groups... in the correct position for reaction with extraordinary accuracy. This binding brings a very large loss of entropy that corresponds to an increase in rate of up to 10^8 for 1 M reactants... Precise binding requires very strong binding... [but]... this loss of entropy must be paid for by a decrease in the observed affinity'.

To me it is amazing how little recognition is given to this turning point in understanding enzyme action (Fersht's textbooks being the important exception [18,24]). Jencks himself, having published his definitive textbook on chemical catalysis in 1969 [25], wrote the following about the 1977 enlarged edition. 'Our understanding of the advantage of intramolecularity in enzymic catalysis [has changed]. The description of this subject in Chapter 1 of the original work is completely wrong, as was most of the published work on the subject in 1969' [26]. The big biochemistry textbooks ignore this development, or worse. Lehninger [27] cites Jencks 1969 only for suggesting abzymes, and refers to the second edition as 'a new printing of an outstanding book'. Stryer [28] also quotes only the abzyme suggestion, and fails to mention the second edition at all. Matthews and van Holde [29] ignore Jencks, calling Koshland's 1958 induced-fit hypothesis 'the dominant model for enzyme catalysis' and citing chymotrypsin

as 'a second example of this after triose phosphate iso-merase.' Voet and Voet [30] do give a one-sentence explanation that an enzyme 'freezes out the relative translational and rotational motions (decreasing their entropy), thereby increasing their reactivity'. They refer, however, to the 1969 edition as 'a classic, and in many ways, still current book'.

The concepts are not easy to teach to students. To be fair, Jencks' 1975 review [31], which became the supplement to the enlarged second edition [26], is not a teaching text. But to me the work of Jencks and Page, giving for the first time a clear interpretation of the catalytic power of enzymes over those of small-molecule catalysts, was a milestone in enzymology of comparable importance to the role of quantum mechanics in chemistry.

In the space of less than 15 years, between say 1963 and 1976, enzyme catalysis had been transformed from something almost magical, because no-one had any detailed insight into how it could be done, into a system capable of rational analysis and quantitative explanation. A huge volume of experimental work on reaction rates, affinity constants, cooperativity and inhibition was ready to be brought into this scheme. I confidently expected to see experimental determinations of enzyme activity and molecular structure quickly interpreted by rigorous quantum-mechanical and thermodynamic analyses.

My expectations were incorrect. The theoretical and computational difficulties of simulating active sites in huge molecules surrounded by water were sufficient to deter most people from persevering in such a quest. Martin Karplus and his colleagues have led the field in using simulation methods to approach protein conformation, folding and dynamics, but simulation of enzyme reaction mechanism remains unsatisfactory. Work from many sources put together a much improved understanding of the properties of bulk water, but thermodynamic description of this remains difficult and controversial, blocking analysis of substrate binding in fundamental terms, although qualitative descriptions are very effective. Tremendous progress has been made in the detailed analysis of the steps of enzyme-catalysed reactions, led by Arieh Warshel [32] and with fundamental contributions from Gábor Náray-Szabó and Oliver Smart [33,34]. But the leaders of the main stream of enzymology have advanced in other directions.

In this decade structural molecular biology has become a mature branch of science, and the laborious detail of structure analysis is delegated to computers. The amazingly rapid progress has proceeded by a qualitative analysis of mechanism based on accurate structural data. This brings us towards overall insight into the whole mechanism of cell biology, from meiosis to specialised cellular functions and the control of development. There is no knowing how far this will lead, but it will certainly exceed what I can imagine today.

A Nature editorial in 1999 [35] was full of profound misunderstanding about the way physical scientists and molecular biologists interact in research, but it made an important point. 'The main method of analysis in molecular biology has been the cartoon representation of... pathways; indeed, superb papers have been written for the purpose of adding a single [curly] arrow to a cartoon. But to... understand..., one needs to have numbers attached to the arrows, and equations relating to the numbers'. Can we claim to analyse and predict reaction rates? What about the entropic factors that are hugely important in understanding them? How do we deal with water in these calculations? Can we even calculate the enthalpy barriers to individual reaction steps with useful accuracy? Or, can we foresee the effects of amino acid substitutions at the active site?

The time will soon come, when people ask for more detail and precision in the inferences drawn from complete genomes. This enzyme is very like that one, but how do its catalytic properties differ and why? These genes are controlled by that promoter, but what transcription factor does it need and under what circumstances will they be expressed? How does this cell induce changes in its neighbour, and what shape will the organism be? As computational power advances, problems that seemed insuperable in the 1990's will begin to look attractively interesting. I firmly believe that we shall come back soon to tackling the detailed and fundamental questions of enzyme mechanism.

References
1. Willstätter, R. (1927). *Problems and Methods in Enzyme Research.* Cornell University Press, Ithaca.
2. Haldane, J.B.S. (1930). *Enzymes.* Longmans Green, London, UK.
3. Mirsky, A. & Pauling, L. (1936). On the structure of native, denatured and coagulated protein. *Proc. Natl Acad. Sci. USA* **22**, 439-447.
4. Pauling, L. (1946). Molecular architecture and biological reactions. *Chem. Eng. News* **24**, 1375-1377.
5. Blake, C.C.F., *et al.*, & Phillips, D.C. (1965). Structure of hen egg-white lysozyme – a three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **206**, 757-761.
6. Blow, D.M. (1997). The tortuous story of Asp...His...Ser: structural analysis of α-chymotrypsin. *Trends Biochem. Sci.* **22**, 405-408.
7. Matthews, B.W., Sigler, P.B., Henderson, R. & Blow, D.M. (1967). Three-dimensional structure of tosyl α-chymotrypsin. *Nature* **214**, 652-656.
8. Wright, C.S., Alden, R.A. & Kraut, J. (1969). Structure of subtilisin BPN′ at 2.5 Å resolution. *Nature* **221**, 235-242.
9. Blow, D.M., Birktoft, J.J. & Hartley, B.S. (1969). Role of a buried acid group in the mechanism of action of chymotrypsin. *Nature* **221**, 337-340.
10. Blow, D.M. (1976). Structure and mechanism of chymotrypsin. *Accts. Chem. Res.* **9**, 145-152.
11. Rühlmann, A., Kukla, D., Schwager, P., Bartels, K. & Huber, R. (1973). Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **77**, 417-436.
12. Sweet, R.M., Wright, H.T., Janin, J., Chothia, C.H. & Blow, D.M. (1974). Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6 Å resolution. *Biochemistry* **13**, 4212-4228.
13. Henderson, R. (1971). Catalytic activity of α-chymotrypsin in which histidine-57 has been methylated. *Biochem. J.* **124**, 13-18.
14. Wright, C.S., Hess, G.P. & Blow, D.M. (1972). Structure of crystalline methyl-chymotrypsin. *J. Mol. Biol.* **63**, 295-303.

15. Craik, C.S., Roczniak, S., Largman, C. & Rutter, W.J. (1987). The catalytic role of the active site aspartic acid in serine proteases. *Science* **237**, 909-913.
16. Freer, S.T., Kraut, J., Robertus, J.D., Wright, H.T. & Xuong, N.H. (1970). Chymotrypsinogen: 2.5 Å crystal structure, comparison with α-chymotrypsin, and implications for zymogen activation. *Biochemistry* **9**, 1997-2009.
17. Kraut, J. (1977). Serine proteases: structure and mechanism of catalysis. *Annu. Rev. Biochem.* **46**, 331-358.
18. Fersht, A.R. (1977). *Enzyme Structure and Mechanism.* Chapter 7B, Freeman, Reading, UK.
19. Hayashi, Y. & Lawson, W.B. (1969). Stereochemistry of the active site of α-chymotrypsin; binding geometry of tryptophan derivatives. *J. Biol. Chem.* **244**, 4158-4167.
20. Rodgers, P.S., Goaman, L.C.G. & Blow, D.M. (1976). On the correlation between three-dimensional structure and reactivity for a series of locked substrates of chymotrypsin. *J. Am. Chem. Soc.* **98**, 6690-6695.
21. Storm, D.R. & Koshland, D.E., Jr. (1970). A source for the special catalytic power of enzymes. *Proc. Natl Acad. Sci. USA* **66**, 445-452.
22. Page, M.I. & Jencks, W.P. (1971). Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc. Natl Acad. Sci. USA* **68**, 1678-1683.
23. Jencks, W.P (1997). From chemistry to biochemistry to catalysis to movement. *Annu. Rev. Biochem.* **66**, 1-18.
24. Fersht, A.R. (1999). *Structure and Mechanism in Protein Science.* Freeman, New York.
25. Jencks, W.P. (1969). *Catalysis in Chemistry and Enzymology.* McGraw Hill, London, UK**.**
26. Jencks, W.P. (1977). Foreword. In *Catalysis in Chemistry and Enzymology.* Enlarged edition, Dover, New York.
27. Lehninger, A.L., Nelson, D.L. & Cox, M.M. (1993). *Principles of Biochemistry.* 2nd Edition, p. 225, Worth, New York .
28. Stryer, L. (1995). *Biochemistry.* 4th Edition, pp. 200-201, Freeman, New York.
29. Matthews, C.K. & van Holde, K.E. (1996). *Biochemistry.* 2nd Edition, pp. 368-369, Benjamin/Cummings, Wokingham, UK.
30. Voet, D. & Voet, J.G. (1995). *Biochemistry.* 2nd Edition, p. 379, Wiley, Chichester, UK.
31. Jencks, W.P. (1975). Binding energy, specificity, and enzymic catalysis: the Circe effect. *Adv. Enzymol.* **43**, 219-410.
32. Warshel, A. (1991). *Computer Modelling of Chemical Reactions in Enzymes.* Wiley, Chichester, UK.
33. Náray-Szabó, G. & Warshel, A. (1997). *Computational Approaches to Biochemical Reactivity.* Kluwer, Dordrecht, Netherlands.
34. Smart, O.S. (1994). A new method to calculate reaction paths for conformational transitions of large molecules. *Chem. Phys. Lett.* **222**, 503-512.
35. Editorial. (1999). Can physics deliver another biological revolution? *Nature* **397**, 89.