CENTENARY LECTURE

Three-dimensional Structures and Chemical Mechanisms of Enzymes

By W. N. Lipscomb
DEPARTMENT OF CHEMISTRY, HARVARD UNIVERSITY, CAMBRIDGE, MASSACHUSETTS 02138, U.S.A.

1 Introduction

Three-dimensional structures, established by *X*-ray diffraction methods, are now known for over 30 proteins.* including many enzymes. Although the enzymes of known structure are mainly limited to single-chain, hydrolytic, extracellular molecules, some general qualitative conclusions can be drawn, as outlined below, from this restricted class of enzymes. Later, when more studies of three-dimensional structures of multisubunit proteins are available, additional and different general principles may be expected to emerge.

Conformation is discussed here primarily as it relates to activity of enzymes in the crystal, in solution, and in biological systems. Probable mechanisms for activity are outlined in the light of the stable, static complexes of enzymes in the crystal as deduced by *X*-ray diffraction studies. Of the several enzymes for which fairly extensive studies of such complexes have been made, those discussed below include carboxypeptidase A, chymotrypsin, ribonuclease S, and lysozyme.

Partial answers will be sought for the following questions, frequently asked when lectures on protein structures are presented. (i) Is the conformation of the enzyme in the crystal like that in solution? (ii) Are these crystalline complexes, now required to be stable for hours or days for *X*-ray work, closely related to functional enzyme–substrate complexes? (iii) How close is the functional binding complex, inferred from these studies, to the transition state of the enzyme–substrate reaction? (iv) What potentially catalytic groups from the enzyme, solution or substrate can approach readily the substrate's reactive region? (v) Is the primary bonding region for product similar to that for the

* Myoglobin, Lysozyme, Carboxypeptidase A, Ribonuclease A, Ribonuclease S, α-Chymotrypsin, Papaine, Subtilisin BPN′, Elastase, Oxyhemoglobin, Deoxyhemoglobin, Insulin, Chymotrypsinogen, Oxidized and Reduced Cytochrome *c*, Rubredoxin, Staphylococcal Nuclease, Insect Hemoglobin, Lamprey Hemoglobin, Subtilisin Carlsberg, Carbonic Anhydrase, γ-Chymotrypsin, Lactate Dehydrogenase, Cytochrome b5, High Potential Iron Protein, Trypsin, Trypsin Inhibitor (pancreatic), Carp Muscle Ca-Binding Protein, Flavodoxin, Subtilisin NOVO, Thermolysin, Concanavalin A, Malate Dehydrogenase, and Ferrodoxin. Recent studies of many of these proteins have appeared in ref. 1.

[1] Volume XXXVI of Cold Spring Harbor Symposium in Quantitative Biology, 1971, Cold Spring Harbor Laboratory, Long Island, 1972.

corresponding part of the substrate? (vi) Why are enzymes so large? (vii) Can we obtain structures of enzyme complexes of rapidly cleaved substrates? These, of course, are detailed questions related ultimately to why enzymes act so specifically and why they cause a reaction to proceed more rapidly than the uncatalysed reaction by factors sometimes approaching 10^7.

2 Conformation

The existence of nearly complete order to atomic dimensions in a protein crystal was indicated from the time of the first *X*-ray diffraction photographs of a protein, pepsin.[2] Moreover, the essential role of the mother liquor in maintaining crystalline order was also recognized,[2] and has remained a consistent feature of protein crystals studied since that time. A recent survey[3] suggests that the solvent occupies about 43% of the volume of protein crystals, but specific examples range from 27 to 65%. Some important consequences are that intermolecular contacts occur only over roughly 30% of the surface of a globular protein in a crystal, that these contacts do not usually block the diffusion of small substrates, products, or inhibitors into the active site cavity or cleft, and that most side-chains of the protein have environments in the crystal rather like those in solution. While it is probable that most peptide polymers do not have a strongly preferred configuration, the *X*-ray evidence suggests that most enzymes which show *X*-ray data to high resolution do have definite molecular structures. The evidence that these structures are not greatly changed in other environments is sometimes less direct. We turn to some examples.

The renaturation of ribonuclease,[4] and its synthesis,[5,6] with full activity provide strong evidence of the uniqueness of those conformational aspects related to activity. A comparison of the two molecules of α-chymotrypsin in totally different environments within the same crystal structure shows very few differences, mostly in flexible side-chains of the enzyme,[7] but one important, possibly functional, change occurs in the positions of the extended polypeptide chain Ser-214–Trp-215–Gly-216.[8] Other differences occur in residues 9—13 and 73—77, and appear to be due to intermolecular contacts. On slightly less secure grounds, one may compare ribonuclease A[9] and ribonuclease S,[10] or the

[2] J. D. Bernal and D. Crowfoot, *Nature*, 1934, **133**, 794.
[3] B. W. Matthews, *J. Mol. Biol.*, 1968, **33**, 491.
[4] C. J. Epstein, R. F. Goldberger, and C. B. Anfinsen, Cold Spring Harbor Symposium on Quantitative Biology, 1963, **28**, 439.
[5] B. Gutte and R. B. Merrifield, *J. Amer. Chem. Soc.*, 1969, **91**, 501.
[6] R. Hirshmann, R. F. Nutt, D. F. Veber, R. A. Vitali, S. L. Varga, T. A. Jacob, F. W. Holly, and R. C. Denkewalter, *J. Amer. Chem. Soc.*, 1969, **91**, 507.
[7] J. J. Birktoft, B. W. Matthews, and D. M. Blow, *Biochem. Biophys. Res. Comm.*, 1969, **37**, 131.
[8] A. Tulinsky, Alpbach Conference on Protein Structures, 1972.
[9] G. Kartha, J. Bello, and D. Harker, *Nature*, 1967, **213**, 862.
[10] H. W. Wyckoff, D. Tsernoglou, A. W. Hanson, J. R. Knox, B. Lee, and F. M. Richards, *J. Biol. Chem.*, 1970, **245**, 305; F. M. Richards and H. W. Wyckoff, in 'The Enzymes', ed. P. D. Boyer, Academic Press, New York, 3rd Edn., 1971, p. 647.

pair subtilisin BPN′[11] and subtilisin NOVO,[12] or the α and β chains of hemoglobin[13–15] along with myoglobin,[16,17] or α- and γ-chymotrypsins,[18,19] or the three serine enzymes α-chymotrypsin,[18] trypsin,[20] and elastase.[21,22] While there are well documented differences within these comparisons, and a few instances of partial disorder in parts of some proteins,[7,10,17] the general conclusion is that the structures are definitive and largely independent of these environmental differences.

It is more difficult to make comparisons between conformations of a molecule in solution and the crystal. The difficult experiment of locating relative heavy-atom positions within a molecule by *X*-ray scattering from solution[23] has never been attempted. However, high-resolution ^{1}H n.m.r., first applied[24] to lysozyme, ribonuclease, cytochrome *c*, myoglobin, and hemoglobin, provides information about local environments and conformational changes. It would seem clear, now that the structures of both oxidized[25] and reduced[26,27] cytochrome *c* are known, that the very large conformational differences for this substance are in contrast to that noted so far for single-chain enzymes. More recent ^{1}H n.m.r. studies have provided striking correlation of histidine and tyrosine environments in ribonuclease A[28] and in staphylococcal nuclease,[29,30] both sets of studies supporting the correlation with these aspects of structures found in the crystalline phases.[10,31,32] These comparisons extend to complexes of inhibitors at the active

[11] (*a*) J. Kraut, in 'The Enzymes', ed. P. D. Boyer, Academic Press, New York, Vol. 3, 3rd Edn., p. 547; (*b*) R. A. Alden, C. S. Wright, and J. Kraut, *Phil. Trans. Roy. Soc.*, 1970, **B257**, 119.
[12] J. Drenth, W. G. J. Hol, J. N. Jansonius, and R. Koekock, ref. 1, p. 107.
[13] M. F. Perutz, H. Muirhead, J. M. Cox, and L. C. G. Goaman, *Nature*, 1968, **219**, 5150.
[14] M. F. Perutz, *Proc. Roy. Soc.*, 1969, **B173**, 113.
[15] M. F. Perutz, *Nature*, 1970, **228**, 5273.
[16] J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, and V. C. Shore, *Nature*, 1960, **185**, 4711.
[17] H. C. Watson in 'Progress in Stereochemistry', ed. B. J. Aylett and M. M. Harris, Butterworth, London, 1969, Vol. 4, p. 299.
[18] J. J. Birktoft, D. M. Blow, R. Henderson, and T. A. Steitz, *Phil. Trans. Roy. Soc.*, 1970, **B257**, 67.
[19] D. M. Segal, G. H. Cohen, D. R. Davies, J. C. Powers, and P. E. Wilcox, ref. 1, p. 85.
[20] R. M. Stroud, L. M. Kay, and R. E. Dickerson, ref. 1, p. 125.
[21] D. M. Shotton and H. C. Watson, *Phil. Trans. Roy. Soc.*, 1970, **B257**, 111.
[22] D. M. Shotton, N. J. White, and H. C. Watson, ref. 1, p. 91.
[23] P. A. Vaughan, J. H. Sturdivant, and L. Pauling, *J. Amer. Chem. Soc.*, 1950, **72**, 5477.
[24] C. C. McDonald and W. D. Phillips, *J. Amer. Chem. Soc.*, 1967, **89**, 6332.
[25] R. E. Dickerson, T. Takano, D. Eisenberg, O. B. Kallai, L. Samson, A. Cooper, and E. Margoliash, *J. Biol. Chem.*, 1971, **246**, 1511.
[26] T. Takano, R. Swanson, O. B. Kallai, and R E. Dickerson, ref. 1, p. 397.
[27] A. G. Redfield and R. D. Gupta, ref. 1, p. 405.
[28] D. H. Meadows, G. C. K. Roberts, and O. Jardetzky, *J. Mol. Biol.*, 1969, **45**, 491.
[29] O. Jardetzky, 'Molecular Properties of Drug Receptors' (Proceedings of the Ciba Symposium, 1970), ed. R. Porter and M. O'Connor, London, 1970, p. 113.
[30] J. L. Markley and O. Jardetzky, *J. Mol. Biol.*, 1970, **50**, 223.
[31] A. Arnone, C. J. Bier, F. A. Cotton, V. W. Day, E. E. Hazen, jun., D. C. Richardson, J. S. Richardson, and A. Yonath, *J. Biol. Chem.*, 1971, **246**, 2302.
[32] F. A. Cotton and E. E. Hazen, jun., in 'The Enzymes', ed. P. D. Boyer, Academic Press, New York, 1971, Vol. 4, 3rd Edn., p. 153.

site of both of these enzymes, including the accompanying conformational changes.

Conformations in the active sites in the molecule in solution or crystal can be compared by testing for enzymatic activity. Crystalline α-chymotrypsin shows about the same activity as in solution.[33] Also, reactions with inhibitors, such as *N*-tosyl phenylalanyl chloromethyl ketone, occur[18] in a way which is thought to be productive for longer substrates in the region of Ser-214–Trp-215–Gly-216.[19] Also, activity towards smaller substrates is known for ribonuclease,[34,35] lysozyme,[36] and papaine.[33] There are, however, some situations where the activity, or reactivity of groups, is reduced in the crystal. The reactivity of ferrihemoglobin with azide ion is some 21 times lower in the crystal than in solution,[37] and the carboxylation of one histidine of myoglobin proceeds more readily in the crystal than in solution.[38] Alcohol dehydrogenase shows a thousand-fold decrease in activity as one goes from solution to crystal[39] and carboxypeptidase A_γ shows, similarly, about a 300-fold decrease.[40] The spectroscopic behaviour of arsenoazo-Tyr-248-carboxypeptidase is also interpreted[41] to indicate substantial differences between the crystal and solution states, but a re-investigation[42] of this effect shows that the behaviour is confirmed for carboxypeptidase A_γ (for which the crystal structure is unknown), but that carboxypeptidase A_α (the known structure) shows similar behaviour in the crystalline and solution states. There exists therefore the possibility that if activity is modified in one crystalline phase, it may be possible in many cases to find a different crystalline phase in which the activity is not substantially different from that in solution.

One of the strongest cautions about *X*-ray studies to date of protein structures is that atoms are located at best to only 0.4 Å or so in the enzyme, and usually somewhat less accurately in the static complexes, and yet activity can depend critically upon a finer level of atomic positions. An interesting illustration is that the active-site functional groups in α-chymotrypsin[43,44] and chymotrypsinogen[45] do not differ greatly. However, the conformations of the nearby portion 214—216 are different and the 'active site' histidine is misaligned by some 15 to 20°, but even so it is not yet clear why chymotrypsinogen is inactive. One further possibility is that chymotrypsin, but not chymotrypsinogen, can carry out a sequence of conformational changes throughout the course of the enzyme–

[33] L. A. A. Sluyterman and M. J. M. de Graaf, *Biochem. Biophys. Acta*, 1969, **171**, 277.
[34] M. Doscher and F. M. Richards, *J. Biol. Chem.*, 1963, **238**, 2399.
[35] J. Bello and E. F. Nowoswiat, *Biochim. Biophys. Acta*, 1965, **105**, 325.
[36] L. G. Butler and J. A. Rupley, *J. Biol. Chem.*, 1967, **242**, 1077.
[37] B. Chance and A. Ravilly, *J. Mol. Biol.*, 1966, **21**, 195.
[38] T. E. Hugli and F. R. N. Gurd, *J. Biol. Chem.*, 1970, **245**, 1930, 1939.
[39] H. Theorell, B. Chance and Y. Yonetani, *J. Mol. Biol.*, 1966, **17**, 513.
[40] F. A. Quiocho and F. M. Richards, *Biochemistry*, 1966, **5**, 4062.
[41] J. T. Johansen and B. L. Vallee, *Proc. Nat. Acad. Sci. U.S.A.*, 1971, **68**, 2532.
[42] F. A. Quiocho and W. N. Lipscomb, to be published.
[43] D. M. Blow, in 'The Enzymes', ed. P. D. Boyer, Academic Press, New York, Vol. 3, 3rd Edn., p. 185.
[44] T. A. Steitz, R. Henderson, and D. M. Blow, *J. Mol. Biol.*, 1969, **46**, 337.
[45] J. Kraut, in 'The Enzymes', ed. P. D. Boyer, Academic Press, New York, Vol. 3, 3rd Edn., p. 165.

substrate reaction. In spite of these qualifications, it seems probable that for most single subunit enzymes the important conformational aspects are very similar in the crystal and in solution, and hence, by inference, in the physiological medium. We now turn to our examples.

3 Carboxypeptidase A [46–48]

This enzyme cleaves, preferentially, aromatic or large aliphatic amino-acids with neutral side-chains from the C-terminal end of a polypeptide substrate. The complete polypeptide backbone of the enzyme is shown in Figures 1 and 2. The Zn^{2+} binding region is in Figure 3. A section of electron density of the enzyme–substrate complex (dotted) superimposed on the electron density of the enzyme (solid contours) is shown in Figure 4, the resulting models of the

Figure 1 *Polypeptide chain in carboxypeptidase A. The* Zn^{2+} *ion at the active site is near the centre of this drawing, and the three ligands to the* Zn^{2+} *from the protein are shown as arrows from α-carbons of His-69, Glu-72, and His-196. The α-carbons are shown as dots, and the peptide units as line segments*

[46] W. N. Lipscomb, *Accounts Chem. Res.*, 1970, **3**, 81.
[47] J. A. Hartsuck and W. N. Lipscomb, in 'The Enzymes', ed. P. D. Boyer, Academic Press, New York, Vol. 3, 3rd Edn., p. 1.
[48] F. A. Quiocho and W. N. Lipscomb, *Adv. Protein Chem.*, 1971, **25**, 1.

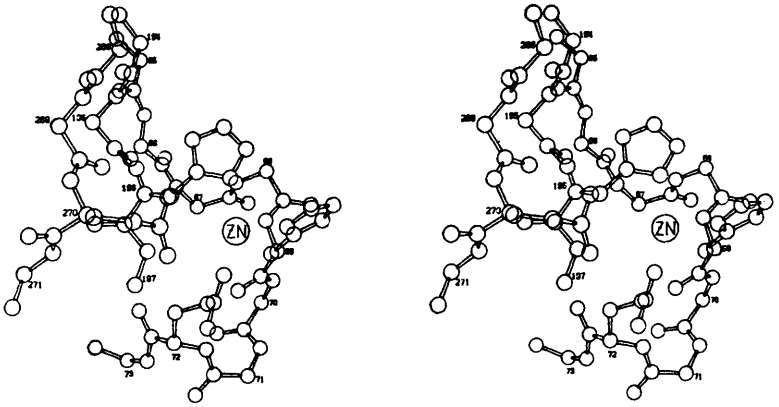

Figure 3 *Binding of* Zn^{2+} *to* N-1 *of His*-69 (*right*), *to* N-1 *of His*-196 (*above*), *and to* O *of Glu*-72. *The fourth ligand, a water molecule to be displaced by the substrate's carbonyl oxygen, is directly towards the reader from* Zn, *but is not shown in the Figure*

enzyme and enzyme–substrate complex are in Figure 5 (*facing p. 326*), and mechanistic implications are suggested in Figure 6.

The major conclusions of this study are (first) that there are very large conformational changes, especially of the side-chain of Tyr-248, when the substrate binds, and (second) that the only two side-chains of the protein which can approach within Van der Waals contact with atoms of the substrate's scissile bond are Glu-270 and Tyr-248. The positive charges of Arg-145 and of the $(ZnL_3)^+$ complex, where L is a protein ligand, serve to bind, respectively, the C-terminal carboxylate group of the substrate, and this substrate's carbonyl group, which is expected to be polarized by the $(ZnL_3)^+$ group. These binding groups, together with hydrophobic binding of the C-terminal side-chain in the enzyme's pocket, strain the peptide bond which is to be cleaved. Thus, the three-dimensional structure indicates that a mechanism of cleavage may involve, in the immediate neighbourhood, Glu-270, Tyr-248, and H_2O. In the optimum pH range just over 7, one may reasonably expect Glu-270 to be anionic, and Tyr-248 to be neutral. Hence we suggest that the proton donor is Tyr-248 or H_2O, and that the nucleophile is Glu-270 or H_2O (Figure 6).

Binding in this region of carboxypeptidase A has been studied to 2.0 Å resolution for Gly-Tyr, to 2.8 Å resolution for Phe-Gly-Phe-Gly, and at 6 Å resolution for about a dozen substrates, inhibitors, products or other analogues, but for no esters. The possibility of binding shifted to one of the five important subsites[49] of carboxypeptidase A requires that an extended series of studies be made, in order that the active site not be confused with a subsite. This is an important general qualification for *X*-ray studies of enzyme complexes.

[49] N. Abramowitz, I. Schechter, and A. Berger, *Biochem. Biophys. Res. Comm.*, 1967, **29**, 862.

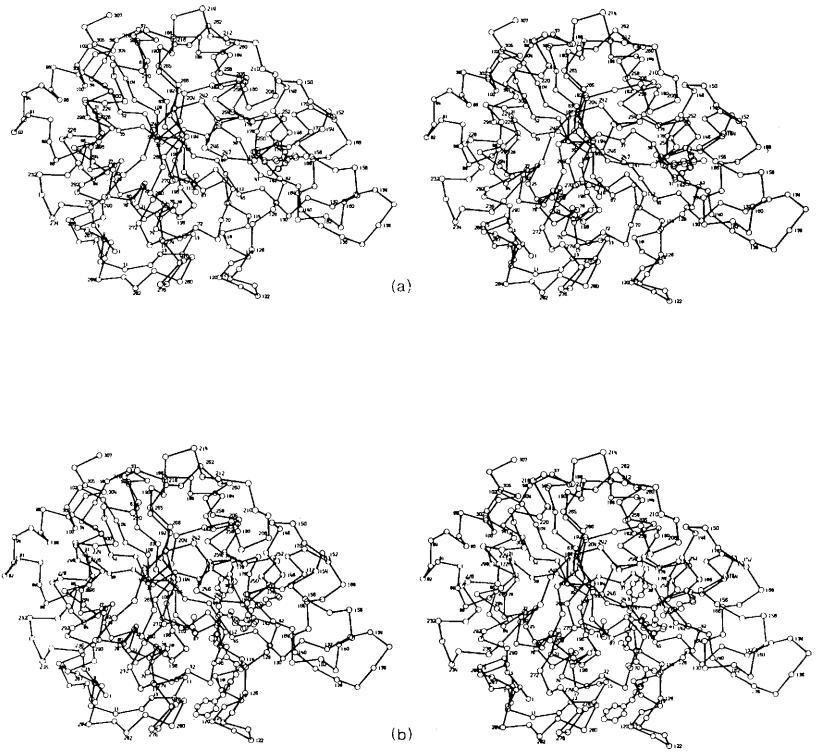

Figure 2 *Stereoview from a different angle, along the $-y$ direction from the top to the bottom of Figure 1, showing especially the large twisted pleated sheet structure in the centre of the carboxypeptidase A molecule.* (a) *Side-chains are shown for Arg*-145, *Tyr*-248, *and Glu*-270 *before substrate is added.* (b) *The same three side-chains are shown after addition of the substrate, benzoxycarbonylalanylalanyltyrosine.* (Stereoviewers may be obtained, for example, from Ward's Natural Science Establishment, Inc., Rochester, New York, Model 25, W2951.)

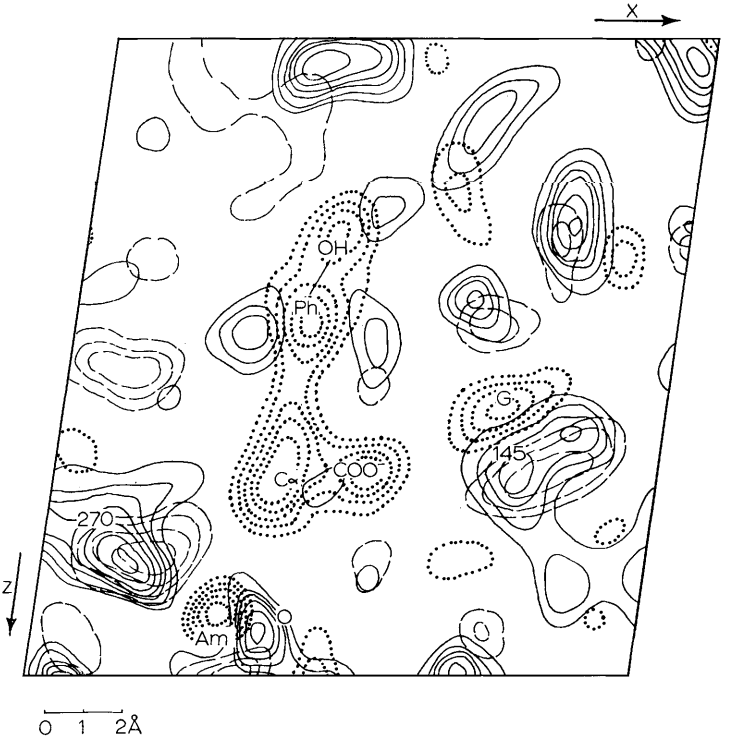

Figure 4 *A section of experimental electron density in the xy plane through part of the substrate glycyl-L-tyrosine. The phenyl and carboxylate groups of Gly-Tyr are partly in this plane. The* OH, Cα, *and probable position of the amino-group are also indicated. Movement of Arg-*145 *to the position* G *is indicated by negative contours added at the initial position, and dotted contours at the new position. Negative contours in the region of Glu-*270 *also indicate another conformational change in the enzyme when Gly-Tyr is bound*

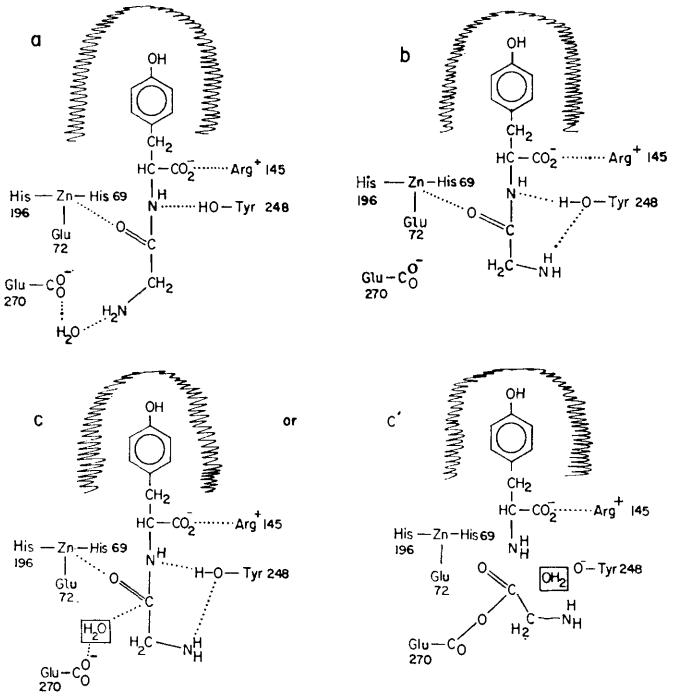

Figure 6 (a) *Binding of glycyl-*L*-tyrosine to carboxypeptidase A.* (b) *Probable change of binding from the non-productive complex to the productive complex* [(a) *to* (b)], *inferred from model building of longer substrates.* (c) *Indirect attack of Glu-*270 *promoting the attack of a water molecule on the carbon of the substrate's carbonyl group polarized by interaction with* Zn. (c') *Alternative, direct attack of Glu-*270 *on the substrate's carbonyl carbon, to be followed by hydrolysis of the resulting anhydride intermediate, which has not so far been detected*

In the absence of evidence for an intermediate, how can one test the mechanistic proposals further? If the pathway involves the acylenzyme intermediate (anhydride), and if cleavage of the anhydride occurred with some reasonable probabilities on either side of the bridging oxygen, then one could incorporate ^{18}O into the enzyme by carrying out substrate-cleavage in $H_2{}^{18}O$. Substitution of the medium by $H_2{}^{16}O$, and subsequent cleavage of a substrate in a few turnovers may then incorporate ^{18}O into the substrate. However, if no ^{18}O is in the final substrate one cannot be sure whether Glu-270 promotes the attack of H_2O or whether the anhydride is unsymmetrically cleaved. Hence, a negative result is ambiguous, but a positive result would indicate that some fraction of the pathway would involve the anhydride intermediate. This experiment has not yet been carried out.

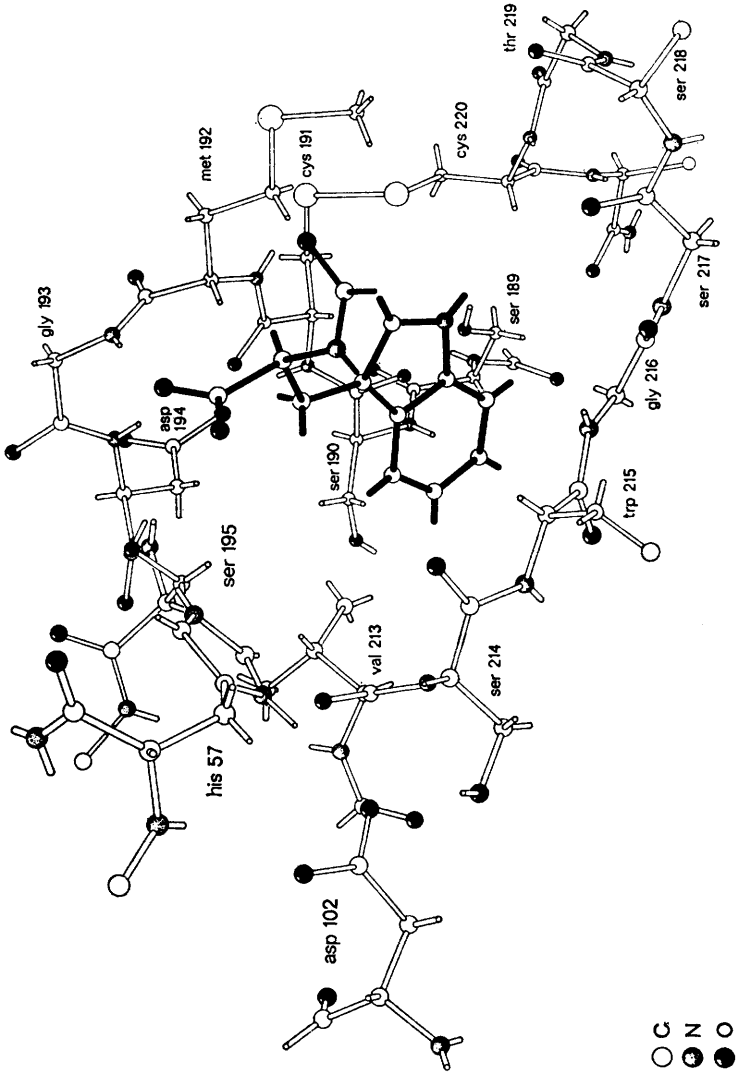

Figure 7(b) *Enlarged view of Figure 7(a) showing N-formyl-L-tryptophan as heavy lines, His-57, and Ser-195. The Trp ring is parallel to the polypeptide chain along the lower centre of the Figure.*

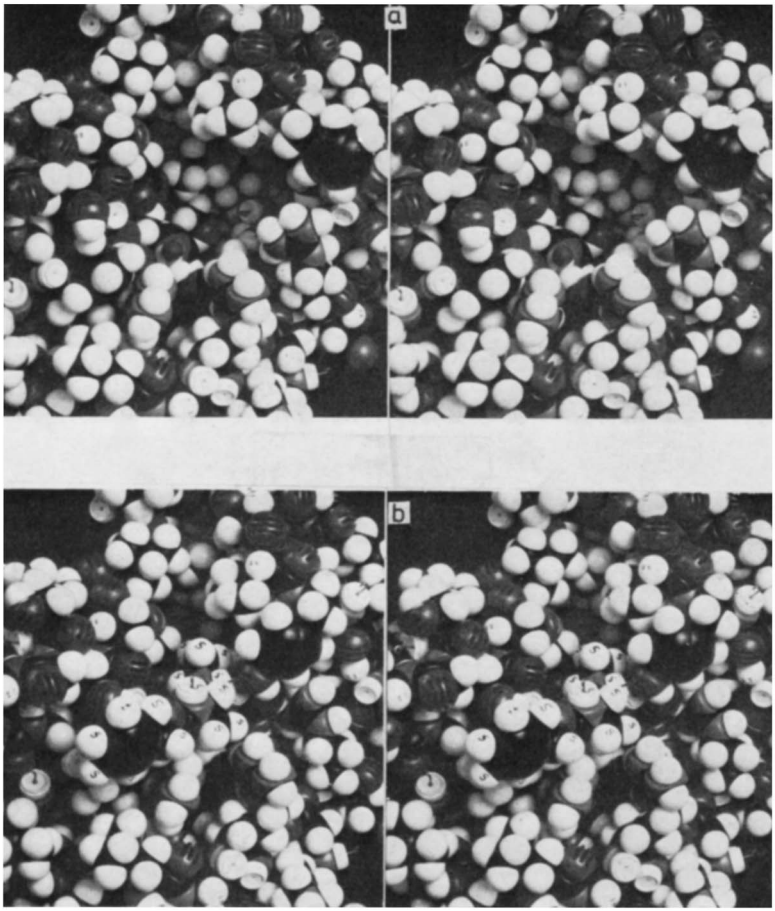

Figure 5 (**a**) *Packing model of about a quarter of the carboxypeptidase molecule, showing the active site cavity near the centre of the photograph. The* Zn *atom is represented by a small black ball at the lower centre. Tyr*-248 *is near the right edge about two-thirds of the way towards the top.* (**b**) *Packing model, showing the substrate benzoxycarbonyl-Ala-Ala-Phe in the active site. Hydrogen atoms of the substrate are marked with the label* S. *Tyr*-248 *has been moved towards the substrate* (*upper-middle, right*)

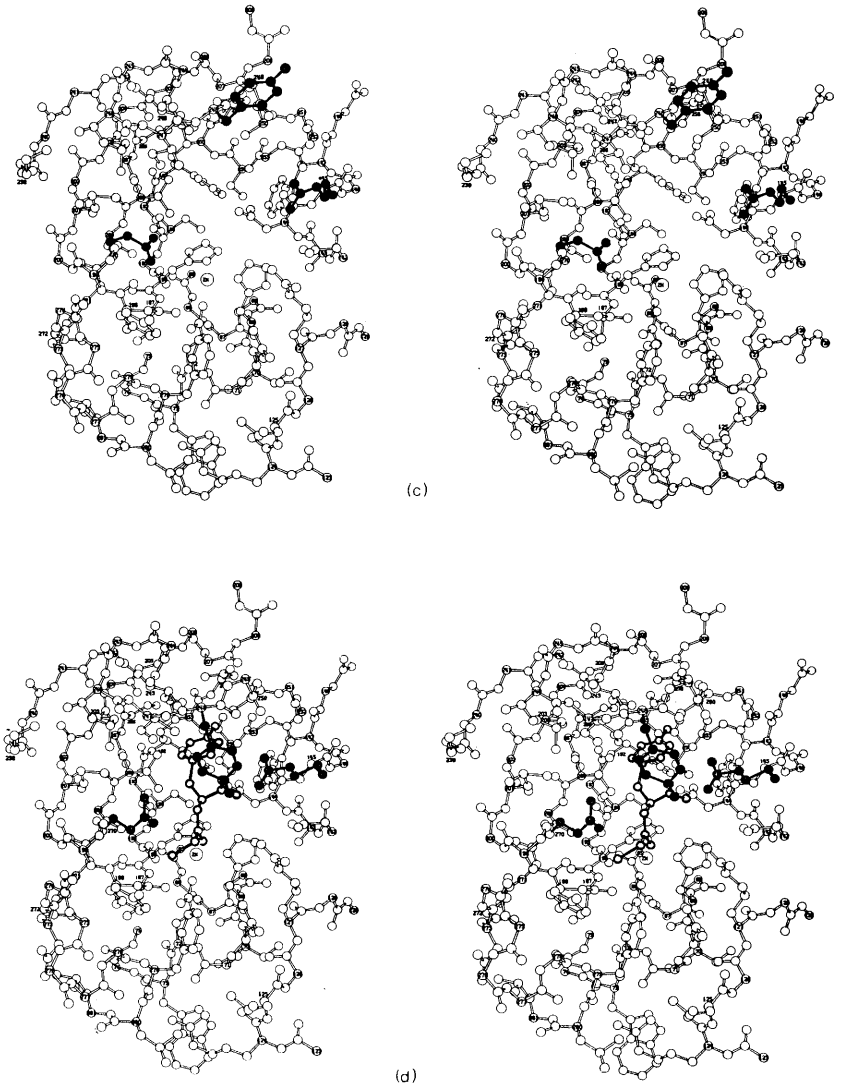

Figure 5 (c) *Stereoview along* $-y$ *of about a quarter of the carboxypeptidase A molecule, showing the cavity, the* Zn *atom, and the functional groups Arg-*145 (*right*), *Tyr-*248 (*above*), *and Glu-*270 (*left*). (d) *Stereoview of the same region, after the addition of glycyl-*L*-tyrosine* (*heavy open circles*), *showing the new positions of Arg-*145, *Tyr-*248, *and Glu-*270. *The guanidinium movement is* 2 Å, *the* OH *of Tyr-*248 *moves* 12 Å, *and the carboxylate of Glu-*270 *moves* 2 Å *when Gly-Tyr binds to the enzyme*

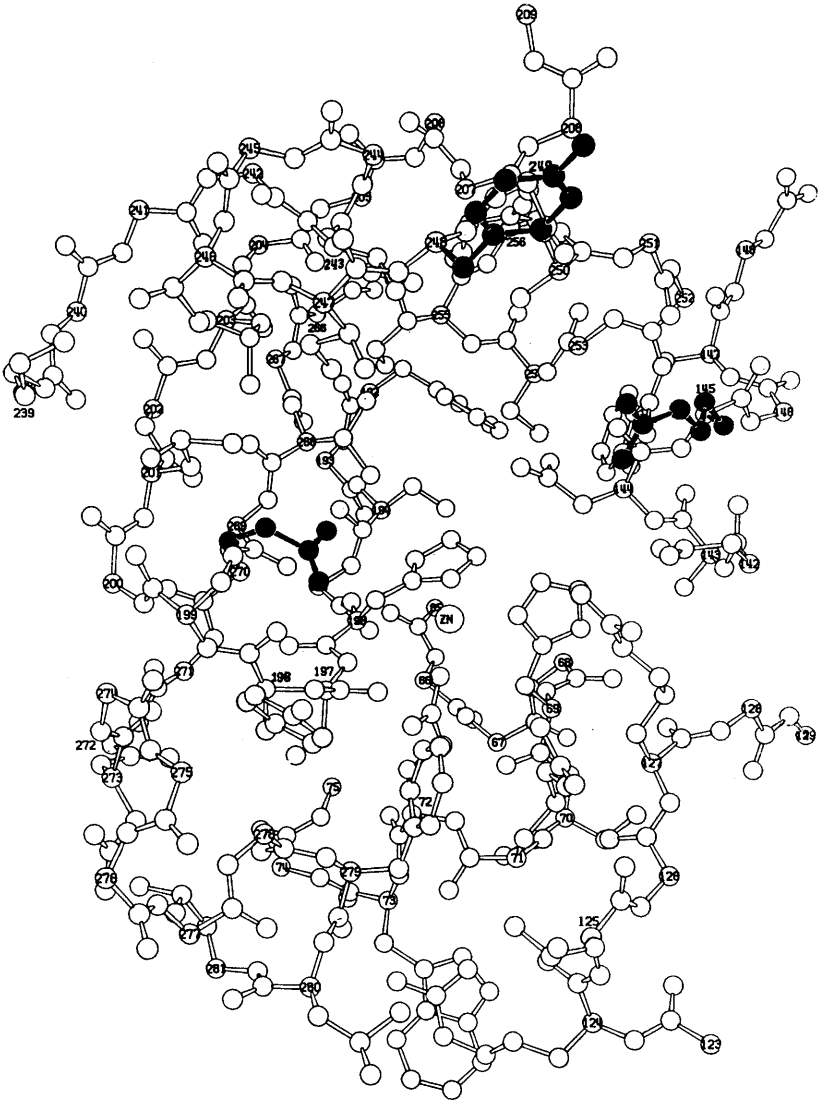

Figure 5 (e) *Enlarged view of Figure* 5(c), *showing Arg*-145 (*right*), *Glu*-270 (*left*), *and Tyr*-248 (*top*). *The* Zn *is shown as the largest circle near the centre of the drawing.*

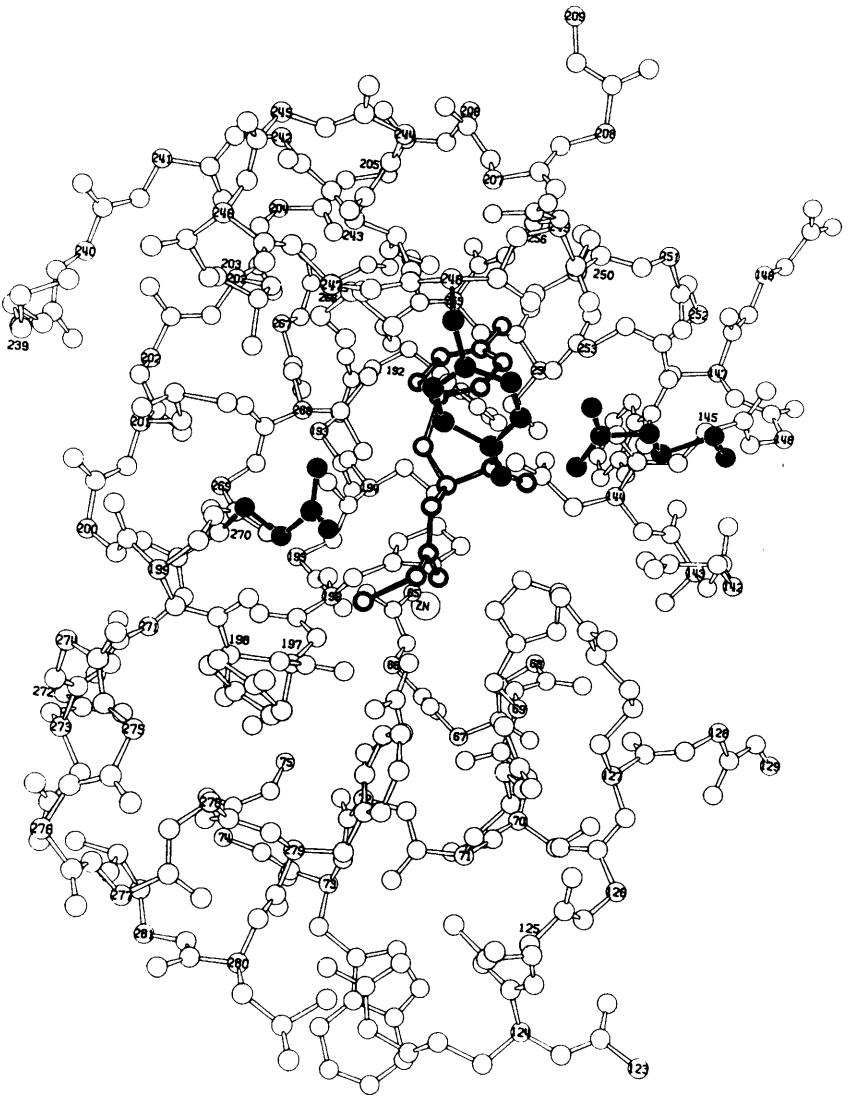

Figure 5 (f) *Enlarged view of Figure* 5(*d*), *showing the substrate Gly-Tyr as heavy open circles. Final positions after conformational changes are indicated for Arg*-145 (*right*), *Glu*-270 (*left*), *and Tyr*-248 (*top*).

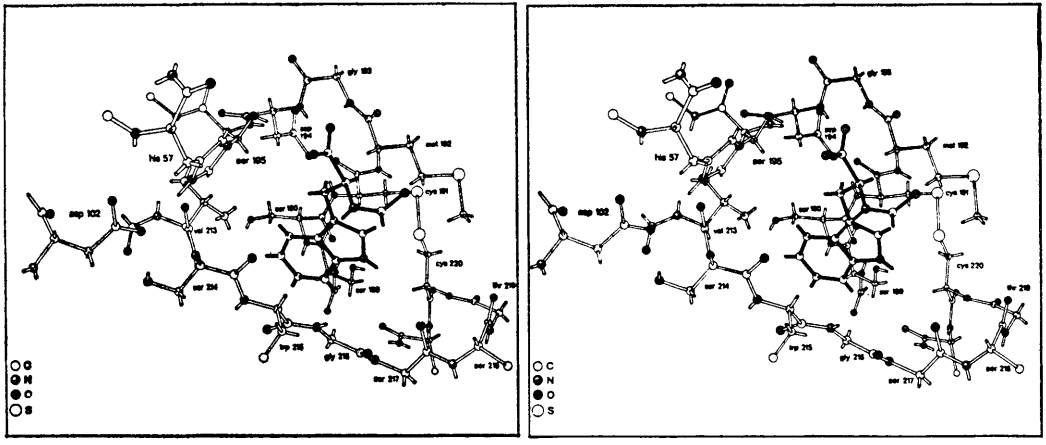

Figure 7(a) *Position of N-formyl-*L*-tryptophan (heavy lines) in the active site of* α*-chymotrypsin. The aromatic ring is in the pocket of the enzyme, and the carboxylate group of this virtual substrate is near Ser-*195 *and His-*57.

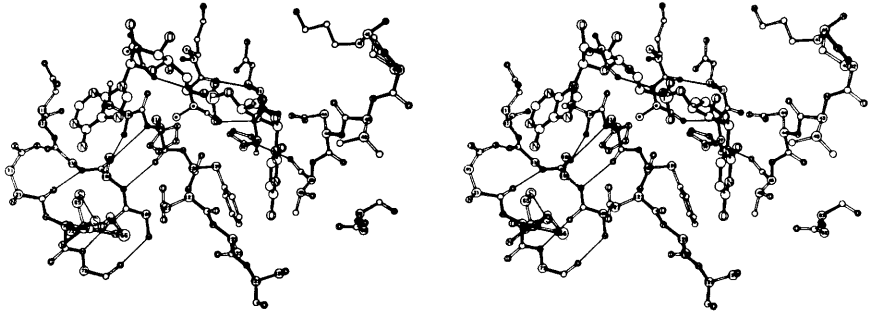

Figure 10(a) *Position of a dinucleoside phosphate analogue (heavy lines) in the active site of ribonuclease* S. *The bond to be cleaved, adjacent to the phosphorus atom in this analogue, is* P—CH_2 *instead of* P—O. *His*-112 *is in an ordered position parallel to the purine ring, and the pyrimidine ring is in the position established earlier from a simpler inhibitor.*

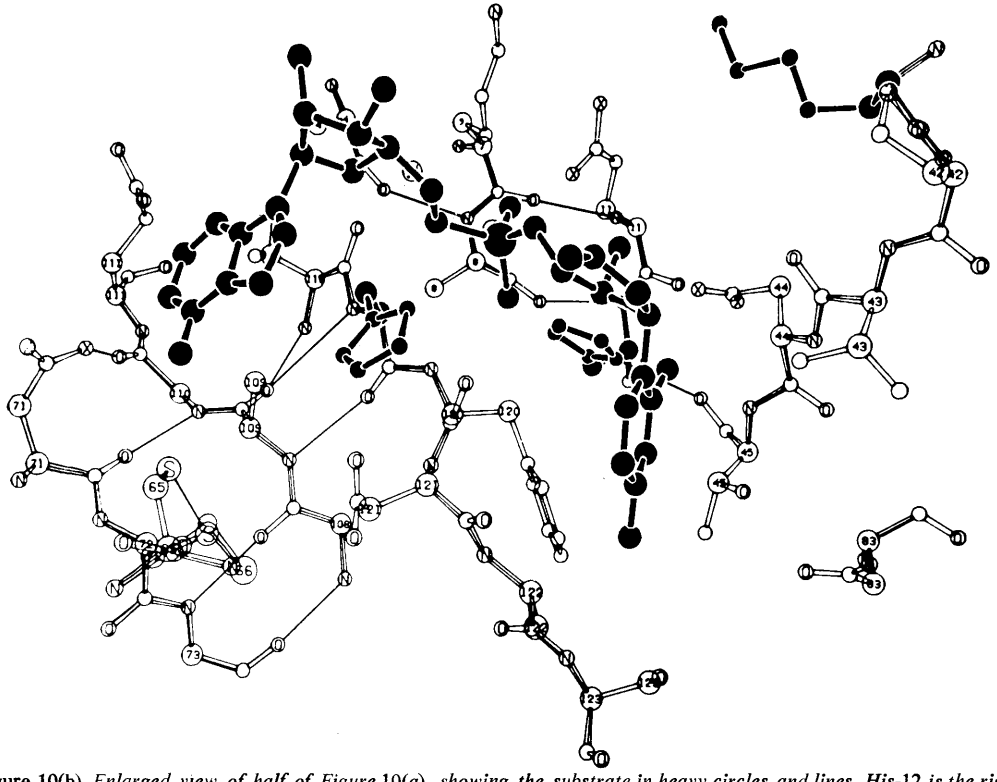

Figure 10(b) *Enlarged view of half of Figure* 10(*a*), *showing the substrate in heavy circles and lines. His*-12 *is the right. His*-119 *at the left is parallel to the purine ring. The largest atom is the phosphorus, and the bond to be cleaved is the* P—O *bond extending toward the left. This* O *atom was replaced by a* CH_2 *group in the X-ray work in order to obtain binding of a non-cleavable substrate analogue*

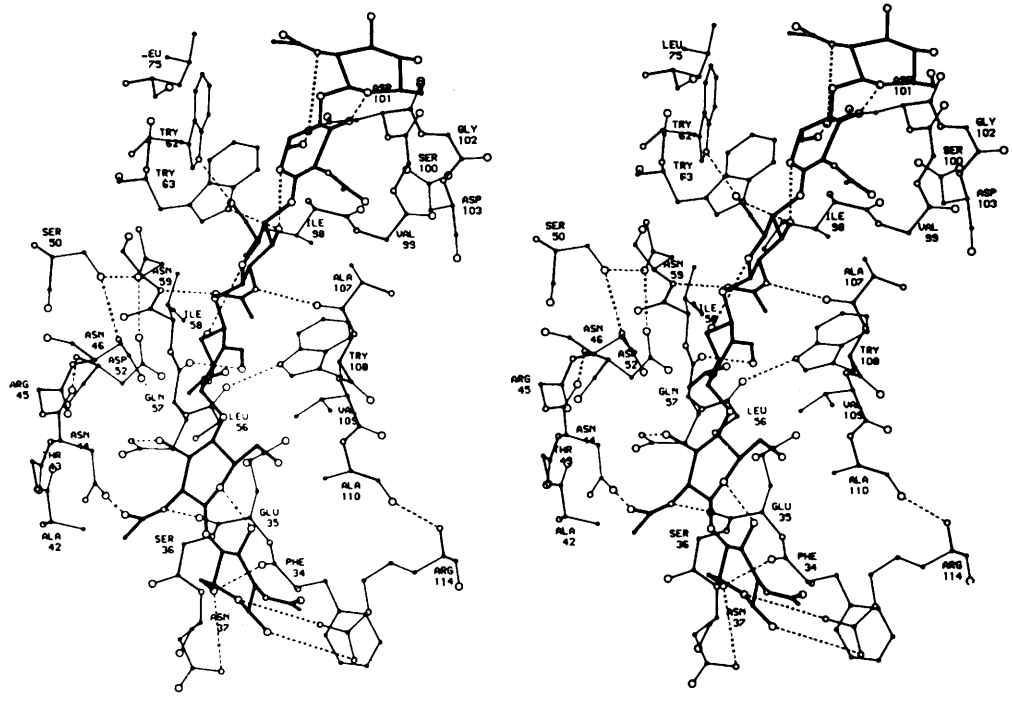

Figure 12 *Stereoview of the N-acetylglucosamine hexamer in the active site and cleft of a portion of the lysozyme molecule. Ring* D, *fourth from the top of the substrate, is shown in the distorted half-chair conformation. Rings* A, B, *and* C (*top*) *are positioned from difference electron density, while rings* D, E, *and* F (*bottom*) *are extrapolated by model building*

4 Chymotrypsin[18,43,44,50]

Before the *X*-ray study was carried out, the chemical evidence for a serine[51] (later shown from sequence studies to be Ser-195) and for a histidine[52,53] (later shown to be His-57) both at the active site was excellent. Also, for some substrates, an acylenzyme intermediate was known.[54–56]

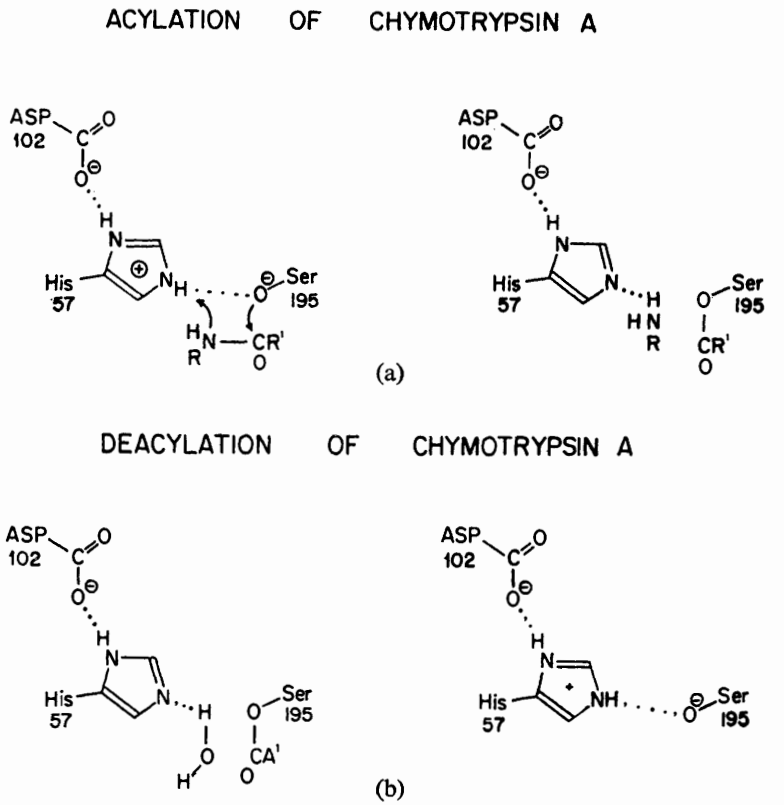

Figure 8 (a) *Acylation steps in hydrolysis of a substrate by α-chymotrypsin.* (b) *Deacylation steps are supposed to be the reverse of the acylation step, but with* H_2O *replacing the original leaving group* H_2NR. *The enzyme is returned to its original form when His*$^+$*-57 donates a proton to Ser*$^-$*-195.*

[50] G. P. Hess, in 'The Enzymes', ed. P. D. Boyer, Academic Press, New York, Vol. 3, 3rd Edn., p. 213.
[51] E. F. Jansen, M. D. F. Nutting, and A. K. Balls, *J. Biol. Chem.*, 1949, **179**, 201.
[52] B. R. Hammond and H. Gutfreund, *Biochem. J.*, 1955, **61**, 187.
[53] G. Schoellman and E. Shaw, *Biochem Biophys. Res. Comm.*, 1962, **7**, 36.
[54] A. K. Balls and H. N. Wood, *J. Biol. Chem.*, 1956, **219**, 245.
[55] M. Caplow and W. P. Jencks, *Biochemistry*, 1962, **1**, 883; B. Hartley and B. Kilby, *Biochem. J.*, 1954, **56**, 288.
[56] I. B. Wilson in 'The Mechanism of Enzyme Action', ed. W. D. McElroy and B. Glass, Johns Hopkins Press, Baltimore, Maryland, 1954.

The structure of α-chymotrypsin and of its complex with *N*-formyl-L-tryptophan[44] (Figure 7) indicate that this inhibitor, or virtual substrate, binds with its aromatic group in the pocket of the enzyme. The carbon of the carboxylate group of the *N*-formyl-L-tryptophan is close to the oxygen of Ser-195. The two oxygens of this carboxylate group appear to be hydrogen-bonded to the OH of Ser-195 and to N_3 of His-57. By contrast to carboxypeptidase A, α-chymotrypsin shows little if any change in the geometry of the active site when this inhibitor is bound. Assuming a close relationship of this complex to an active enzyme–substrate interaction, we see that the probable nucleophile Ser-195 and the probable proton donor His-57, in the acylation step (Figure 8a), are in Van der Waals contact with the atoms of the peptide bond to be cleaved. The deacylation step is just the reverse, but with a water molecule replacing the leaving group (Figure 8b).

These conclusions are based upon extrapolation of this complex to long peptide substrates. The amide of the *N*-formyl-L-tryptophan can easily be moved (0.7—0.8 Å) so that it could form a hydrogen bond to the backbone carbonyl of Ser-214, and it is proposed that binding of a longer peptide substrate be continued along this direction to form a short antiparallel pleated sheet with Ser-214, Trp-215, and Gly-216. Some evidence that this may be reasonable is the position of binding of the inhibitor shown in Figure 9. This complex was obtained by reacting the corresponding chloroketone at His-57, and then deriving the atomic positions from an *X*-ray diffraction study.[19] Although a longer peptide binds to the enzyme also in the other direction, toward the carboxylate end, neither model-building nor an *X*-ray diffraction study have yielded a satisfactory structure for this part of the enzyme–substrate complex.

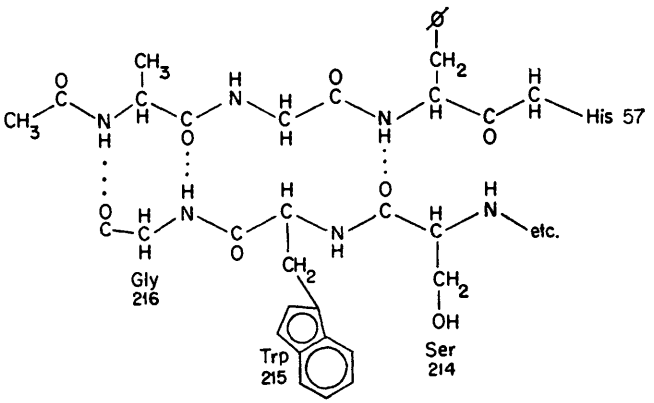

Figure 9 (a) *Antiparallel pleated sheet found in the reaction product of acetyl alanyl glycyl phenylalanyl chloromethyl ketone with His-57 of α-chymotrypsin. The protein sequence Gly-216–Trp-215–Ser-214 can also be seen in Figure 7*

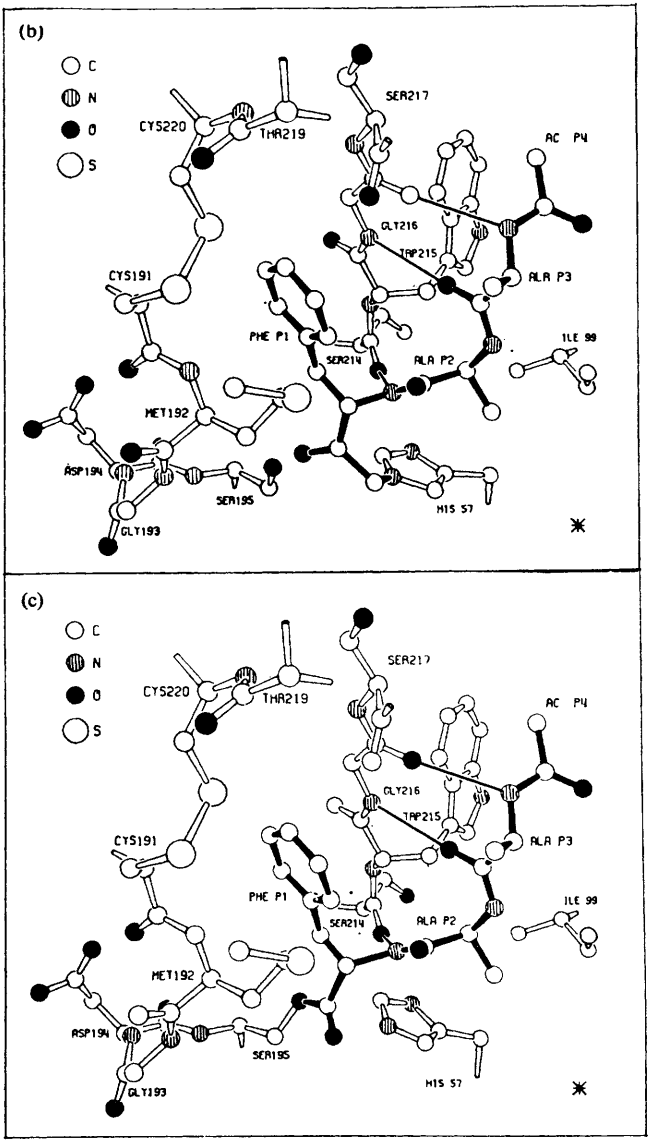

Figure 9 **(b)** *Chymotrypsin A_γ inhibited by acetyl-Ala-Ala-Phe-chloromethyl ketone after loss of* Cl *when reacted with His-57.* **(c)** *Proposed enzyme–ester intermediate showing acetyl-Ala-Ala-Phe acylating Ser-195 of chymotrypsin A_γ*

A buried charge of Asp-102, which is hydrogen-bonded to His-57 in turn hydrogen-bonded to Ser-195, may help to activate the catalytic region.[57] While it is improbable that protons can be transferred in hydrogen bonds across a large 'pH gradient', it does seem possible that this system provides the easiest route for the lines of force of this charge to escape to a region of higher local dielectric constant. This same hydrogen-bonded system has subsequently been found in other serine enzymes. Charges are also buried by the substrate in carboxypeptidase A and in lysozyme.

5 Ribonuclease A[10]

Chemical evidence for the participation of two histidines[58,59] and one lysine in the active site of ribonuclease led to an early reasonable proposal of a mechanism.[60–62] However, the detailed stereochemistry has become clear only recently, and the role of Lys-41 is still not understood.

A portion of the structure of ribonuclease S, along with a dinucleoside phosphate substrate analogue[10] is shown in Figure 10. The pyrimidine ring of this analogue is in the same position as that found earlier for several smaller inhibitors which showed occupancy curves similar to those for these inhibitors in solution.[63] The position of His-119 is now ordered, as compared with some disorder in earlier studies, and is parallel to the purine ring of this dinucleoside phosphate analogue. The only two side-chains of the protein which are in contact with the oxygen atoms of the phosphate group (if one replaces the CH_2—P of this analogue by O—P) are the rings of His-119 (N_1) and His-12 (N_3). If Lys-41 is salt-linked to this phosphate it must be hydrogen-bonded through a water molecule.

There are two steps in the hydrolysis of a dinucleoside phosphate.[64] The first step is a transesterification in which the 2′-oxygen of the ribose ring adjacent to the pyrimidine is the nucleophile. For example, if cytidylyl adenosine (CpA) is hydrolysed, this step produces cytidine-2′,3′-cyclic phosphate and adenosine. The stereochemistry of this step can clearly be carried out without pseudorotation,[65] as shown in Figure 11(a), if one starts from the binding stage shown in Figure 10 (see plates section between pp. 326 and 327).

The second step, again using CpA as an example, is the attack of water finally yielding cytidine-3′-phosphate as the product. However, the stereochemistry of this stage of the reaction depends upon the direction of attack of the water molecule on phosphorus, and cannot be derived from the *X*-ray results. The

[57] D. M. Blow, J. J. Birktoft, and B. S. Hartley, *Nature*, 1969, **221**, 337.
[58] E. A. Barnard and W. D. Stein, *J. Mol. Biol.*, 1959, **1**, 339.
[59] A. M. Crestfield, W. H. Stein, and S. Moore, *J. Biol. Chem.*, 1963, **238**, 2421.
[60] D. Findlay, D. G. Herries, A. P. Mathias, B. R. Rabin, and C. A. Ross, *Biochem. J.*, 1962, **85**, 152.
[61] A. Deavin, A. P. Mathias, and B. R. Rabin, *Nature*, 1966, **211**, 252.
[62] A. Deavin, A. P. Mathias, and B. R. Rabin, *Biochem. J.*, 1966, **101**, 14.
[63] H. W. Wyckoff, K. D. Hardman, N. M. Allewell, T. Inagami, D. Tsernoglou, L. N. Johnson, and F. M. Richards, *J. Biol. Chem.*, 1967, **242**, 3749.
[64] J. P. Hummel and G. Kalnitsky, *Ann. Rev. Biochem.*, 1964, **33**, 15.
[65] F. H. Westheimer, *Accounts Chem. Res.*, 1968, **1**, 70.

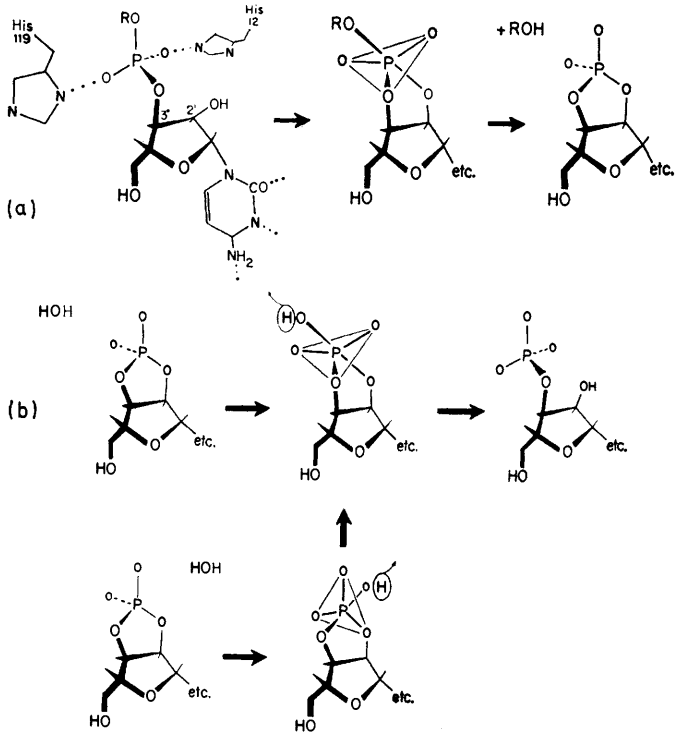

Figure 11 (a) *Formation of the cyclic intermediate occurs without pseudorotation in the five-co-ordinate intermediate.* (b) *Hydrolysis of the five-co-ordinate intermediate is shown by the upper pathway not involving pseudorotation, or by the lower pathway in which pseudorotation occurs. The choice of pathways is ambiguous in the X-ray results, but an analogous reaction in which one oxygen is replaced by sulphur, another labelled by* ^{18}O*, the third unlabelled and the last two in the sugar suggests that the pathway not involving pseudorotation is the correct one*

upper pathway of Figure 11(b) is similar to the pathway of formation, except that H_2O replaces the leaving group of Figure 11(a); no pseudorotation is involved. The pathway starting at the lower part of Figure 11(b) does involve pseudorotation. A recent experiment,[66] in which all five atoms of the five-co-ordinated intermediate are labelled, indicates that the pathway not involving pseudorotation is most probably correct: two of the oxygens at phosphorus are attached to the ribose ring in the cyclic intermediate, one other oxygen is replaced by sulphur, the fourth is ^{18}O, and the fifth is ^{16}O in this experiment.

[66] D. A. Usher, D. I. Richardson, jun., and F. Ekstein, *Nature*, 1970, **228**, 665.

6 Lysozyme[67]

The first enzyme, and the second protein, to reach atomic resolution, lysozyme is so far the clearest illustration of strain imposed upon a substrate when it is bound to an enzyme.

While the largest rates of enzymatic activity occur near pH 5, ranging from pH 3 to 7, there is also activity at higher pH's.[68] The discussion here refers to the low pH range. The binding to lysozyme of the trimer of *N*-acetylglucosamine, $(NAG)_3$, has been studied by *X*-ray diffraction methods in some detail.[67] This inhibitor fills half of the large groove or cleft in the surface of lysozyme. When these results were extended by model building to a hexamer, the fourth ring could not easily be placed in its chair form, but had to be distorted to a half-chair conformation. In Figure 12 (facing p. 327) the hexamer, $(NAG)_6$, is positioned in the cleft of lysozyme. The prominent mode of cleavage of this hexamer, or of the cell-wall type of polymer having alternating NAG and NAM (*N*-acetylmuramic acid), is cleavage between rings D and E such that the bridging oxygen is retained in the product EF. As shown in Figure 13, the only two side-

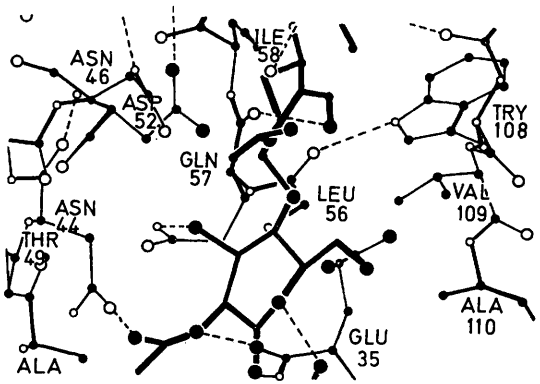

Figure 13 *Enlarged view of one-half of Figure 12 showing the substrate along the centre from top to bottom. The bond to be cleaved extends towards the left from the filled circle in the centre, which is the* O *atom of the scissile* C—O *bond. Asp*-52 *is to the upper left and Glu*-35 *is in the lower right. The optimal* pH *range is in a region where Asp*-52 *is most probably ionized, and Glu*-35 *is most probably protonated, both in the enzyme and in the enzyme–substrate complex*

chains of the enzyme in contact with atoms of the substrate's susceptible bond are Asp-52 and Glu-35, respectively assigned as ionized and un-ionized near pH 5.

In extrapolating these results to plausible mechanisms,[69] one must note that the half-chair form of ring D would favour a carbonium ion intermediate. But

[67] C. C. F. Blake, G. A. Mair, A. C. T. North, D. C. Phillips, and V. R. Sarma, *Proc. Roy. Soc.*, 1967, **B167**, 365. See also other papers in this issue pp. 349—448.
[68] J. A. Rupley. *Proc. Roy. Soc.*, 1967, **B167**, 416.
[69] L. N. Johnson, D. C. Phillips, and J. A. Rupley, 'Structure, Function and Evolution of Proteins', Brookhaven Symposia in Biology, 1968, No. 21, p. 120.

there are also nucleophiles in the immediate environment, including Asp-52, the acetamido side-chain (Figure 14), and water. The stereochemistry favours either the formation of a carbonium ion, stabilized by the negative charge of

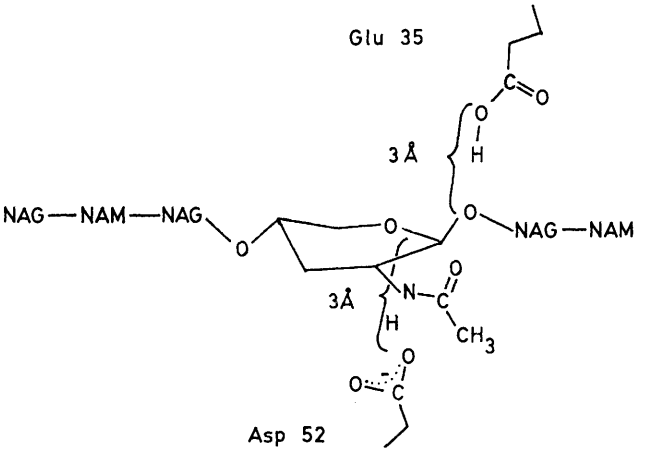

Figure 14 *Ring* D *shown in its distorted half-chair conformation, along with nearby side-chains of lysozyme. The hexamer is alternating N-acetylglucosamine and N-acetylmuramic acid, similar to material from the cell-wall polymer*

Asp-52, or a partial or complete covalent bond between the ring carbon of the substrate and an oxygen of Asp-52 (Figure 15a, 15b). Attack of the acetamido's carbonyl oxygen would require gross distortion of the binding, and hence probably does not occur at least when the substrate fragment is still bound. Attack of water produces a retention of configuration, and is favoured as a later reaction on geometrical grounds. Also, the stereochemistry favours Glu-35 as the initial proton donor, in the low pH range, rather than water.

7 The Activated Complex

The role of strain, seen most clearly in lysozyme, appears also to be of importance in the other enzymes discussed here. Concerning the role of enzymes in efficient catalysis of chemical reactions, the most perceptive of the early comments is probably that of Pauling,[70,71] 'I think that enzymes are molecules that are complementary in structure to the activated complexes of the reactions that they catalyse, that is, to the molecular configuration that is intermediate between the reacting substances and the products of reaction for these catalysed processes. The attraction of the enzyme molecule for the activated complex would thus lead to a decrease in its energy, and hence to a decrease in the energy of activation of the reaction and to an increase in the rate of the reaction.' Perhaps one

[70] L. Pauling, *Nature*, 1948, **161**, 707.
[71] L. Pauling, *American Scientist*, 1948, **36**, 51.

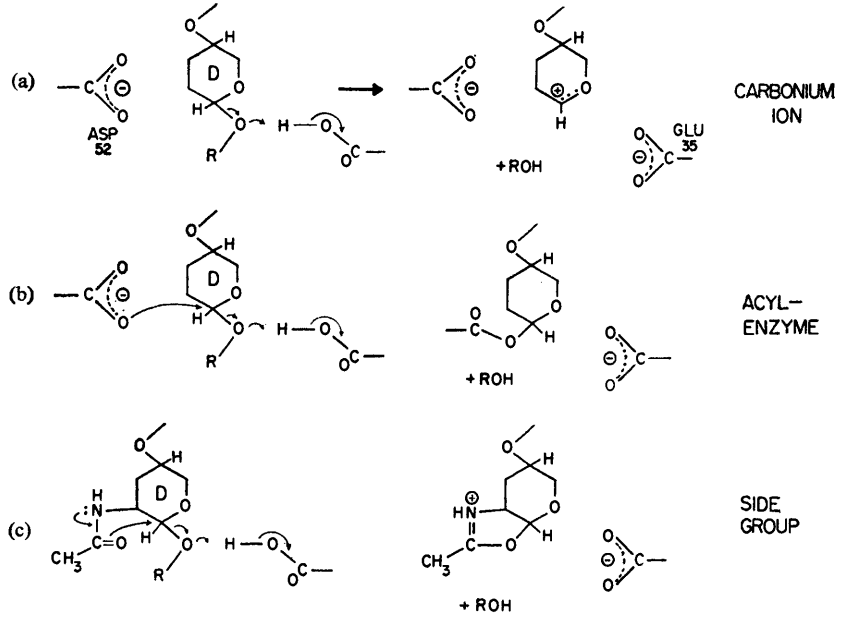

Figure 15 *Three plausible pathways for the hydrolysis of the bond between ring* D *and the bridging oxygen to ring* E. (a) *The carbonium ion mechanism is favoured.* (b) *The acylenzyme pathway requires some distortion in the active site, but cannot be eliminated as yet.* (c) *The participation of the acetamido side-chain requires large distortions in the enzyme–substrate complex, but is believed to be favoured in model reactions in solution under certain conditions*

might change the first part from 'enzymes are molecules that are complementary . . . ' to ' . . . enzymes are molecules that become complementary . . . ,' in order to place some emphasis on the flexibility of binding sites, but Pauling went on to suggest that substrate analogues which resemble the transition state may be more tightly bound to enzymes than are analogues of the substrate itself. Experiments strongly supporting these ideas[72] have recently been published.[73,74] It seems very likely that binding and strain towards the transition-state complex is a major source of the catalytic power of enzymes.

[72] J. B. S. Haldane, 'Enzymes', Longmans Green and Co., New York, 1930, p. 180; R. Lumry in 'The Enzymes', ed. P. D. Boyer, H. Lardy, and K. Myrbäck, Academic Press, New York, 1959, Vol. 1, p. 157; W. P. Jencks in 'Current Aspects of Biochemical Genetics', ed. N. O. Kaplan and E. P. Kennedy, Academic Press, New York, 1966, p. 273; W. P. Jencks, 'Catalysis in Chemistry and Enzymology', McGraw-Hill Book Co., New York, 1969, p. 282.

[73] R. Wolfenden, *Nature*, 1969, **223**, 704; *Biochemistry*, 1970, **9**, 3409; *Accounts Chem. Res.*, 1972, **5**, 10; L. N. Johnson and R. Wolfenden, *J. Mol. Biol.*, 1970, **47**, 93; B. Evans and R. Wolfenden, *J. Amer. Chem. Soc.*, 1970, **92**, 4751.

[74] G. E. Lienhard, I. I. Secemski, K. A. Koehler, and R. N. Lindquist, Cold Spring Harbor Symposium on Quantitative Biology, June 1971; I. I. Secemski and G. E. Lienhard, *J. Amer. Chem. Soc.*, 1971, **93**, 3544; K. A. Koehler and G. E. Lienhard, *Biochemistry*, 1971, **10**, 2477.

8 Factors other than Strain

The effects of local environment in the enzyme or near its surface appear to modify strongly the reactivities of various groups of the enzyme, substrate, or solvent. Examples are given above of charges which are buried by substrates or by conformational changes. One type of example is the change of a pK of a potential catalytic group. Perhaps most spectacular is the change reported by Westheimer[75] of the pK near 6 (about 4.5 units lower than 'normal') for the active lysine in acetoacetate decarboxylase. Even when structures are known for an enzyme and its complexes it is difficult to estimate the effects of nearby charges or of local dielectric constants in these effects, which nevertheless can be large. The general effect is the stabilization of states with less absolute charge when a hydrophobic environment is formed.

Other factors include the concerted character of the enzyme–substrate reaction, similar in some ways to intramolecular reactions in model compounds. Proximity effects, and the detailed geometry at the reactive atoms, must also be important in moving the reaction co-ordinate along the catalytic pathway. Geometry includes orientation effects, which are perhaps related to minimization of non-bonded repulsions at least as much as to promotion of bonded interactions, but both effects are of interest; this is a gentle comment on orbital steering.

9 What Were Those Questions Again?

(i) Not all proteins, of course, but probably enzymes which can be crystallized and which show diffraction patterns to atomic resolution will be substantially ordered to the atomic level in the crystalline state. Preliminary enzymatic activity, chemical modification, spectroscopic and especially high-resolution n.m.r. methods can be used to compare conformations in solution with the crystallographic results. The comparisons are favourable so far. (ii) Parts of the protein involved in activity can be compared in chemical modification experiments using real substrates and in three-dimensional proximity in complexes of enzymes with virtual, or slowly cleaved, substrates. The correspondence here is also excellent. (iii) If studies are made on the X-ray structures of a number of substrate analogues, inhibitors, and products as bound to the enzyme in the crystals, a pattern of binding can be deduced. It seems likely that one can find substrate analogues in which atoms are only a few to several tenths of an Ångstrom from the positions in the transition-state complex of a good substrate with most enzymes which can be studied by X-ray diffraction methods. (iv) The potential catalytic groups, side-chains of the enzyme or even of the substrate itself, are probably to be found in Van der Waals contact with the atoms of the enzymic reaction, or they can be brought into contact largely by rotations about single bonds. (v) Product binding, when studied by X-ray methods, is frequently so specific that bonding shifted to neighbouring subsites is rare. However, results based on one or a few studies can be misleading. Nevertheless, the binding of

[75] F. H. Westheimer, 'Bio-organic Chemistry and Mechanisms', The Welch Foundation Conferences on Chemical Research, XV, Houston, 1971.

products often resembles closely the binding of that corresponding part of the substrate analogue in the protein molecules in the crystalline phase.

Question (vi), 'Why are proteins so large?', requires a separate paragraph, since the answers are not all known. (*a*) Proteins of definite conformation have hydrophobic regions in the interior regions, and a hydrophilic exterior contributing to solubility. (*b*) These same problems of forming a hydrophobic interior and hydrophilic exterior apply not only to the final form of an enzyme, but also to the proenzyme, and to the portions of the initially formed protein during biosynthesis. (*c*) The hydrophobic interior may also be important in activating a charged group, either for binding or catalysis, when its environment is changed upon formation of the enzyme–substrate complex. (*d*) Flexibility is required to allow participation of catalytic groups, and to permit capture of substrate and release of products. (*e*) The hydrolytic enzymes have relatively extended binding regions, involving some five amino-acids for carboxypeptidase, seven for papaine, and six sugar sites for lysozyme. (*f*) There are two lines of evidence which suggest that enzymes are large for reasons not well understood. Modification experiments far from an active site can affect activity. The second bit of evidence refers to the very many small shifts of atoms always seen in electron density maps throughout the enzyme when a substrate analogue is bound. These shifts, generally much smaller than those in the active site, may not have a negligible cumulative effect on the activity of an enzyme. Perhaps we shall understand them some day; perhaps some protein crystallographer will describe them so that we know what they are! It is likely, then, that enzymes are evolved for binding of substrates, more especially for strain of enzyme–substrate complex towards the transition state, and for the whole sequence of conformational changes involved in the chemical stages of the enzyme–substrate reaction.

The last question also has several answers. Use of storage rings for electron beams may produce a very large increase in *X*-ray intensity for protein structure investigation. The development of television scanners with electronic image intensifiers will probably reduce the time for *X*-ray data collection by a factor of about one hundred, and thereby make possible study of complexes of more rapidly cleaved substrates than those now being investigated. Use of low-temperature methods, say about $-50\,^{\circ}C$, will also slow the enzyme substrate reaction, but one must be careful that the temperature difference does not change the rate-determining step, and that the use of a solvent other than the usual aqueous buffer does not also change the conformation and reaction. The promise of these approaches is very great, and these new methods may then allow us some freedom to give greater attention to the choice of appropriate problems of biochemical importance. For example, the further development of methods for isolation, characterization, and structure determination will yield results for enzymes and other proteins having subunit structures, from which new principles are likely to emerge.

I wish to thank D. M. Blow, D. C. Phillips and F. M. Richards for their courtesy in providing stereoviews shown herein.