Structure and Mechanism of Chymotrypsin

David M. Blow

MRC Laboratory of Molecular Biology, Hills Road, Cambridge, United Kingdom

Received December 27, 1974

Our understanding of enzyme catalysis has advanced steadily since x-ray diffraction results gave the first detailed information about the architecture of an enzyme molecule 10 years ago. The mechanism of chymotrypsin is probably understood in more detail than any other enzyme at the present time. As enzymes go, it is a very simple one, and it has proved accessible to study by a wide range of different techniques. Although it is 7 years since the first molecular model of chymotrypsin was built,[1] it is only recently that I have begun to believe that the major effects are qualitatively understood. A quantitative assessment, which will require a quantum mechanical description of the active site and a detailed thermodynamic analysis of each step of the reaction, is still probably several years away.

Chymotrypsin catalyzes the hydrolysis of peptide bonds of protein foods in the mammalian gut. It is secreted in the pancreas as chymotrypsinogen, a single-chain protein of 245 amino acids,[2] and is activated to chymotrypsin by the hydrolysis of a single peptide bond, catalyzed by trypsin.[3]

Amino acid sequence analysis has shown that enzymes which have a close homology to chymotrypsin are widespread. In addition to chymotrypsin B, trypsin, and elastase, all secreted in the pancreas, there are several similar enzymes in the blood-clotting system, including thrombin and blood-clotting factor X, and in the complement-fixing system.[4–6] A trypsin-like enzyme occurs in the Pacific dogfish and the sea anemone, and the enzyme cocoonase used by the pupa of the silkworm to digest its cocoon is closely related.[7] Proteolytic enzymes in bacteria mostly show no sign of genetic relationship to chymotrypsin. Subtilisin, found in *B. subtilis* and *B. amyloliquefaciens*, has a completely unrelated structure,[8] though it uses an almost identical catalytic mechanism.[9] Chymotrypsin-like enzymes have, however, been found in some species of soil bacteria, including *Myxobacter 495*. In these enzymes, large tracts of the amino acid sequence, including the catalytically important residues, show extensive homology with chymotrypsin, and there can be little doubt that the basic molecular architecture is similar.[10] Possibly these enzymes have developed through the protozoa from a primitive ancestral gene, but it seems more likely that some kind of genetic transfer has occurred from higher organisms.

Chymotrypsin is a monomeric enzyme which exhibits no allosteric effects. The conformational changes which occur in the process of normal proteolysis seem to be very small. One of the substrates in proteolysis is water, which has to be present in all experiments on the enzyme. These features have greatly simplified study of the catalytic mechanism.

Scheme I

$$ECH_2OH + RCOX \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \underset{\text{Michaelis complex}}{ECH_2OH:RCOX} \xrightarrow{k_2}$$

$$ECH_2OCOR \xrightarrow[H_2O]{k_3} ECH_2OH + RCOOH$$

$$\text{acyl enzyme} + HX$$

The Acyl-enzyme Mechanism

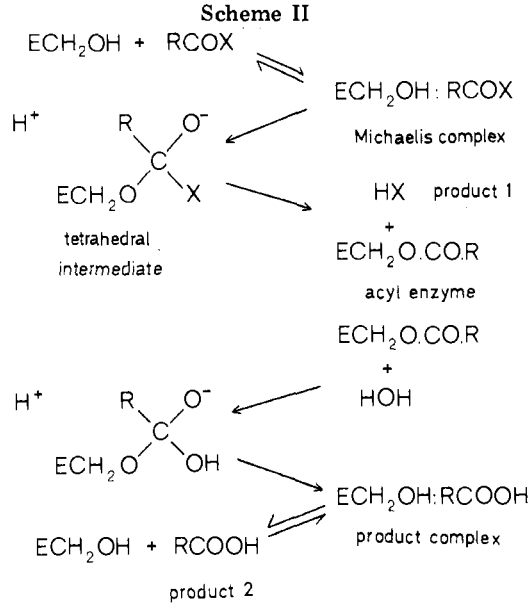
Chymotrypsin has a uniquely reactive serine, Ser-195, which may be acylated by a variety of agents. When *p*-nitrophenyl acetate is used, a rapid "burst" of 1 mol of *p*-nitrophenol/mol of enzyme is produced, followed by a slower steady release of nitrophenol as the acetyl-chymotrypsin is hydrolyzed.[11] Both steps are slowed down at low pH, and at pH 4 acetyl-chymotrypsin is sufficiently stable to be isolated.[12] Since the pK_a of the reaction is about 6.7, it was early suggested that histidine was involved in the reaction.[13]

Many substrates give similar kinetic behavior, and all esters of the same acyl group show the same steady-state rate of hydrolysis. This suggested that the same acyl-enzyme intermediate was formed in every case, and that hydrolysis of the common acyl-enzymes was the rate-limiting step. Amides, on the other hand, are hydrolyzed at slower rates, which depend on the amide group involved.[14] This is consistent with Scheme I, in which CH_2OH represents the

David M. Blow studied natural sciences at Corpus Christi College, Cambridge, and studied for his Ph.D. under M. F. Perutz in the Medical Research Council's Unit for the study of molecular biological systems in the Cavendish Laboratory, Cambridge. After 2 years in the United States, he returned to Cambridge and has been on the staff of the Laboratory of Molecular Biology since it was established in 1962.

(1) B. W. Matthews, P. B. Sigler, R. Henderson, and D. M. Blow, *Nature (London)*, **214**, 652–656 (1967).
(2) B. S. Hartley, *Nature (London)*, **201**, 1284–1287 (1964).
(3) P. Desnuelle, *Enzymes, 1st Ed.*, **4**, Chapter 5 (1960).
(4) B. S. Hartley, *Symp. Soc. Gen. Microbiol.*, **24**, 151–182 (1974).
(5) E. W. Davie and E. P. Kirby, *Curr. Top. Cell. Regul.*, **7**, 51–86 (1973).
(6) T. Barkas, G. K. Scott, and J. E. Fothergill, *Biochem. Soc. Trans.*, **1**, 1219–1220 (1973).
(7) H. Neurath, R. A. Bradshaw, and R. Arnon in "Structure–Function Relationships of Proteolytic Enzymes", P. Desnuelle, H. Neurath, and M. Ottesen, Munksgaard, Copenhagen, 1970, pp 113–134.
(8) F. S. Markland and E. L. Smith, *J. Biol. Chem.*, **242**, 5198–5211 (1967).
(9) J. D. Robertus, R. A. Alden, J. J. Birktoft, J. Kraut, J. C. Powers, and P. E. Wilcox, *Biochemistry*, **11**, 2439–2449 (1972).
(10) M. O. J. Olson, N. Nagabhushan, M. Dzwinich, L. B. Smillie, and D. R. Whitaker, *Nature (London)*, **228**, 438–442 (1970); L. T. J. Delbaere, W. L. B. Hutcheon, M. N. G. James, and W. E. Thiessen, *Nature (London)*, **257**, 758–763 (1975).
(11) B. S. Hartley and B. A. Kilby, *Biochem. J.*, **50**, 672–678 (1952).
(12) A. K. Balls and J. H. Wood, *J. Biol. Chem.*, **219**, 245–256 (1956).
(13) L. Cunningham, *Compr. Biochem.*, **16**, 85–188 (1965).
(14) M. L. Bender and F. J. Kezdy, *J. Am. Chem. Soc.*, **86**, 3704–3714 (1964).

Scheme II



Ser-195 side-chain of the enzyme E, and RCOX is an amide or ester substrate.

Here k_1/k_{-1} gives the binding constant K_s for formation of the Michaelis complex; k_2, the acylation rate constant, is rate limiting for amide hydrolysis; and k_3, the deacylation rate constant, is rate limiting for ester hydrolysis.

Even quite recently, experimental data have been published which seemed to contradict the acyl-enzyme mechanism,[15-17] but further analysis has shown a particular disturbing effect in each case[18-20] and the acyl-enzyme mechanism may be taken as established.

The Tetrahedral Intermediate

Formation and hydrolysis of an ester bond, as indicated by steps 2 and 3 of Scheme I, are supposed to proceed by a mechanism in which the trigonally coordinated C is attacked by the approach of a fourth ligand, approximately perpendicular to the plane of coordination.[21-23] This approach provides a fourth ligand to C, which moves slightly out of the plane of the three original ligands to form a tetrahedrally coordinated form. The tetrahedral form may collapse to a more stable state, either by repelling the attacking ligand and returning to its original state or by repelling another ligand and forming a new three-coordinated state, by moving into the plane which the new ligands define (Scheme II; Figure 1).

Due to the relative instability of the tetrahedral form, it is not present in sufficient abundance to have a detectable effect on the kinetics of a normal hydrolytic reaction. Until recently, only indirect evidence[22] for the existence of such a form had been obtained, but two crystallographic studies of complexes of trypsin with protein trypsin inhibitors have shown that these are stabilized in the tetrahedral form.[24,25]

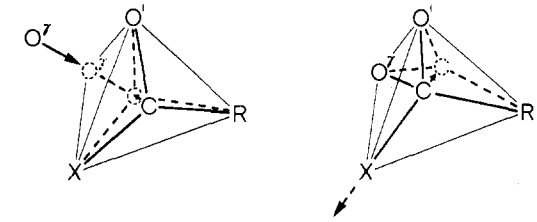


Figure 1. The carbon atom C moves between one planar trigonal state and another by passing through the center of the tetrahedron.

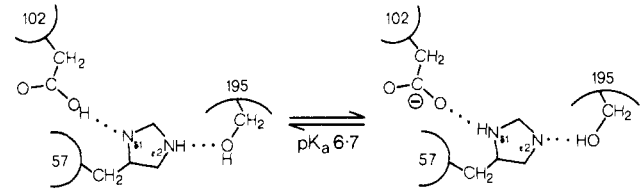


Figure 2. The charge relay system, indicating the transition between the state characterizing the active enzyme and the (inactive) form which exists below pH 6.7, according to Hunkapiller et al.[31]

The Charge-Relay System

The first interpretable electron-density map of chymotrypsin[1] showed, with the aid of the amino acid sequence, that the side chain of His-57 was lying close to Ser-195. This gave satisfying agreement with the supposed involvement of histidine from the pH–activity curve. It was only when a more accurate electron density map was obtained, after further crystallographic work, and a correction to the amino acid sequence, that a far more interesting situation was revealed[26] (Figure 2).

Ser-195 lies in a shallow depression on the enzyme surface, with its O^{γ} at hydrogen-bond distance from one of the ring nitrogens ($N^{\epsilon 2}$) of His-57. The imidazole ring of His-57 has its edge at the surface of the molecule, so that $N^{\epsilon 2}$ is accessible from the solvent; the other nitrogen ($N^{\delta 1}$) is buried. Immediately behind $N^{\delta 1}$, and completely buried from contact with the solvent, is the side chain of Asp-102, with its carboxyl group at hydrogen-bonding distance from $N^{\delta 1}$. Two water molecules are buried close to the carboxyl group, but not hydrogen bonded to it;[27] otherwise the environment of the carboxyl group is completely hydrophobic.

The unique juxtaposition of a buried acid group with the unusually polarizable system of the imidazole ring, which in its uncharged form can carry a proton on either of the two ring nitrogens, seemed at first to be the key to the activity of the trypsin family of enzymes. This view was strengthened when it turned up again and again, not only in the crystal

(15) R. M. Epand, *Biochem. Biophys. Res. Commun.*, **37**, 313–318 (1969).
(16) S. E. Bresler, G. P. Yeasov, and V. M. Krutyakov, *Mol. Biol. (Moscow)*, **3**, 15–28 (1969).
(17) G. P. Hess, J. McConn, E. Ku, and G. McConkey, *Philos. Trans. R. Soc. London, Ser. B*, **257**, 89–104 (1970).
(18) I. V. Berezin, N. F. Kazanskaya, A. A. Klyasov, and V. K. Svedas, *Eur. J. Biochem.*, **38**, 529–536 (1973).
(19) A. R. Fersht and Y. Requena, *J. Mol. Biol.*, **60**, 279–290 (1971).
(20) J. Fastrez and A. R. Fersht, *Biochemistry*, **12**, 2025–2034 (1973).
(21) L. P. Hammett, "Physical Organic Chemistry", McGraw-Hill, New York, N.Y., 1940, pp 355–357.
(22) M. L. Bender, *J. Am. Chem. Soc.*, **73**, 1626–1629 (1951).
(23) H. B. Bürgi, J. D. Dunitz, and E. Shefter, *J. Am. Chem. Soc.*, **95**, 5065–5067 (1973). The initial approach makes an angle of 109° to the C–O direction.
(24) R. Huber, D. Kukla, W. Bode, P. Schwager, K. Bartels, J. Deisenhofer, and W. Steigemann, *J. Mol. Biol.*, **89**, 73–101 (1974). Refined, unpublished coordinates show a tetrahedral conformation around the carbonyl carbon, but the distance to the serine O^{γ} is 2.6 Å.[50]
(25) R. M. Sweet, H. T. Wright, J. Janin, C. H. Chothia, and D. M. Blow, *Biochemistry*, **13**, 4212–4228 (1974).
(26) D. M. Blow, J. J. Birktoft, and B. S. Hartley, *Nature (London)*, **221**, 337–340 (1969).
(27) J. J. Birktoft and D. M. Blow, *J. Mol. Biol.*, **68**, 187–240 (1972).

structures of elastase[28] and of trypsin,[29] but also in the structure of subtilisin.[30]

We proposed[26] that the function of the buried acid group was to polarize the imidazole ring, since the buried negative charge would induce a positive charge adjacent to it. The exposed nitrogen, $N^{\epsilon 2}$, would carry an excess negative charge, strengthening the hydrogen bond between Ser-195 and His-57. This gave a possibility of proton transfer along hydrogen bonds, which could possibly allow the hydroxyl proton to be transferred to His-57 from Ser-195. A deprotonated O^{γ} would be highly reactive to attack the scissile peptide bond. There were, however, obvious difficulties about a scheme which involved proton transfer. The pK_a of $-CH_2OH$ is about 15, so that the amount of proton transfer from serine to histidine at pH 7 is likely to be very small. Also, the system clearly has a pK_a at about 6.5 to 7, which is absolutely normal for a histidine, while this particular histidine is in an extremely abnormal environment. In our 1969 paper,[26] therefore, the issue of proton transfer was carefully left open.

These problems seem to have been resolved by the experiments of Hunkapiller et al.[31] on the α-lytic protease of *Myxobacter 495*. Although little is known of the structure of this protein at present, it is assumed that the homologies of active-site sequences imply that there is a similar Asp-His-Ser system in the active site. The important feature of the enzyme is that it contains only one histidine residue, so that the ^{13}C magnetic resonance of the one $C^{\delta 2}$ histidine atom can be clearly identified. The pH-dependent changes of chemical shift and coupling constant of this resonance were carefully analyzed, and they indicate that the pK_a of the histidine residue is about 4.2. The change at pH 6.5 to 7 is ascribed to protonation of the carboxylate of Asp-102 by these workers (although others have criticized this interpretation[32]).

This means that the buried environment of Asp-102 perturbs its pK_a strongly, making it much more difficult to ionize than a normal carboxyl group. His-57 is strongly perturbed by the adjacent protonated carboxyl group, and also becomes much more difficult to ionize than normal. In terms of the proton transfer system, the new result is that below pH 7 the proton between His-57 and Asp-102 is localized on Asp-102 (Figure 2). Above pH 7 it is transferred to $N^{\delta 1}$, but the histidine remains electrically neutral by losing the proton on $N^{\epsilon 2}$. The transition at pH 7 thus reverses the polarization of the imidazole ring. $N^{\delta 1}$ shows the anomaly that it is deprotonated below pH 6.7, but protonated above pH 7.

In another series of NMR measurements, the proton resonance due to the proton between Asp-102 and His-57 of chymotrypsin has been observed. This shows a transition near pH 7, corresponding to the transfer of the proton.[33]

In the active form of the enzyme, the polarization of the imidazole ring above pH 7 creates an environment which can readily accommodate the proton of Ser-195 when this attacks the scissile bond. The repeated occurrence of the charge-relay system in neutral proteases leaves little room for doubt that this constitutes the proton-transfer system required to implement Scheme II. But the discovery that the charge-relay system exists in identical form in the "inactive" precursor chymotrypsinogen[33,34] showed that this was not enough to make an effective enzyme. Trypsinogen and chymotrypsinogen do exhibit enzymatic activity, but at such a low level ($\sim 10^{-7}$ that of trypsin) that it can only be detected under special conditions.[35,36]

The Substrate Binding Site

Close to the active serine of chymotrypsin is a deeply invaginated pocket, large enough to accommodate an indole or a toluene molecule. When an aromatic group occupies this pocket, it is sandwiched between peptide bonds 190–191 and 191–192 on one side and 215–216 on the other. A hydrophobic group so accommodated can make numerous hydrophobic contacts, especially involving Ser-190, Cys-191–220, Val-213, Trp-215, and Tyr-228[37] (Figure 3). This pocket does not exist in chymotrypsinogen, but is formed by the rearrangement of residues 191–194 which occurs on activation.[34]

The narrowness of the pocket defines the plane of the aromatic side chain of a Tyr, Phe, or Trp residue, including the C^{β} atom. The shape of the mouth of the pocket leaves little freedom for the C^{α} position. The whole conformation of the amino acid side chain in the primary specificity pocket is thus rather closely defined. The shape of the pocket corresponds exactly with the range of amino acids which define the primary specificity of the enzyme.

The substrate binding pocket does not restrict the conformational angle χ_1 which defines rotation about the C^{α}–C^{β} bond and which thus controls the position of adjacent atoms of the polypeptide chain (Figure 4). Possible values for χ_1 are partly restricted by the shape of the molecular surface outside the substrate binding pocket. A conveniently placed hydrogen-bond acceptor, the CO group of Ser-214, close to the mouth of the pocket, further controls χ_1. When the NH group of an amino acid in the primary binding site is directed toward this acceptor, its carbonyl carbon is brought close to the side chain of Ser-195.

If the above interactions are made, the orientation of the scissile bond depends on the conformational angle ψ, which defines rotation about the C^{α}–C bond (Figure 4). Once again, the adjacent surface of the enzyme restricts the possibilities, and permits two ranges for this angle, about 180° apart. One of these brings the carbonyl oxygen of the scissile bond close to two hydrogen-bond donors which are directed away from the enzyme surface. These are the peptide NH groups of Gly-193 and Ser-195.[37,38] (The alterna-

(28) D. M. Shotton and H. C. Watson, *Nature (London)*, **225**, 811–816 (1970).
(29) R. M. Stroud, L. M. Kay, and R. E. Dickerson, *J. Mol. Biol.*, **83**, 185–208 (1974).
(30) C. S. Wright, R. A. Alden, and J. Kraut, *Nature (London)*, **221**, 235–242 (1969).
(31) M. Hunkapiller, S. H. Smallcombe, D. R. Whitaker, and J. H. Richards, *J. Biol. Chem.*, **248**, 8306–8308 (1973).
(32) G. Robillard and R. G. Shulman, *J. Mol. Biol.*, **86**, 519–540 (1974).
(33) G. Robillard and R. G. Shulman, *J. Mol. Biol.*, **71**, 507–511 (1972).
(34) S. T. Freer, J. Kraut, J. D. Robertus, H. T. Wright, and N. H. Xuong, *Biochemistry*, **9**, 1997–2009 (1970).
(35) J. Kay and B. Kassell, *J. Biol. Chem.*, **246**, 6661–6665 (1971).
(36) P. H. Morgan, N. C. Robinson, K. A. Walsh, and H. Neurath, *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 3312–3316 (1972).
(37) T. A. Steitz, R. Henderson, and D. M. Blow, *J. Mol. Biol.*, **46**, 337–348 (1969).

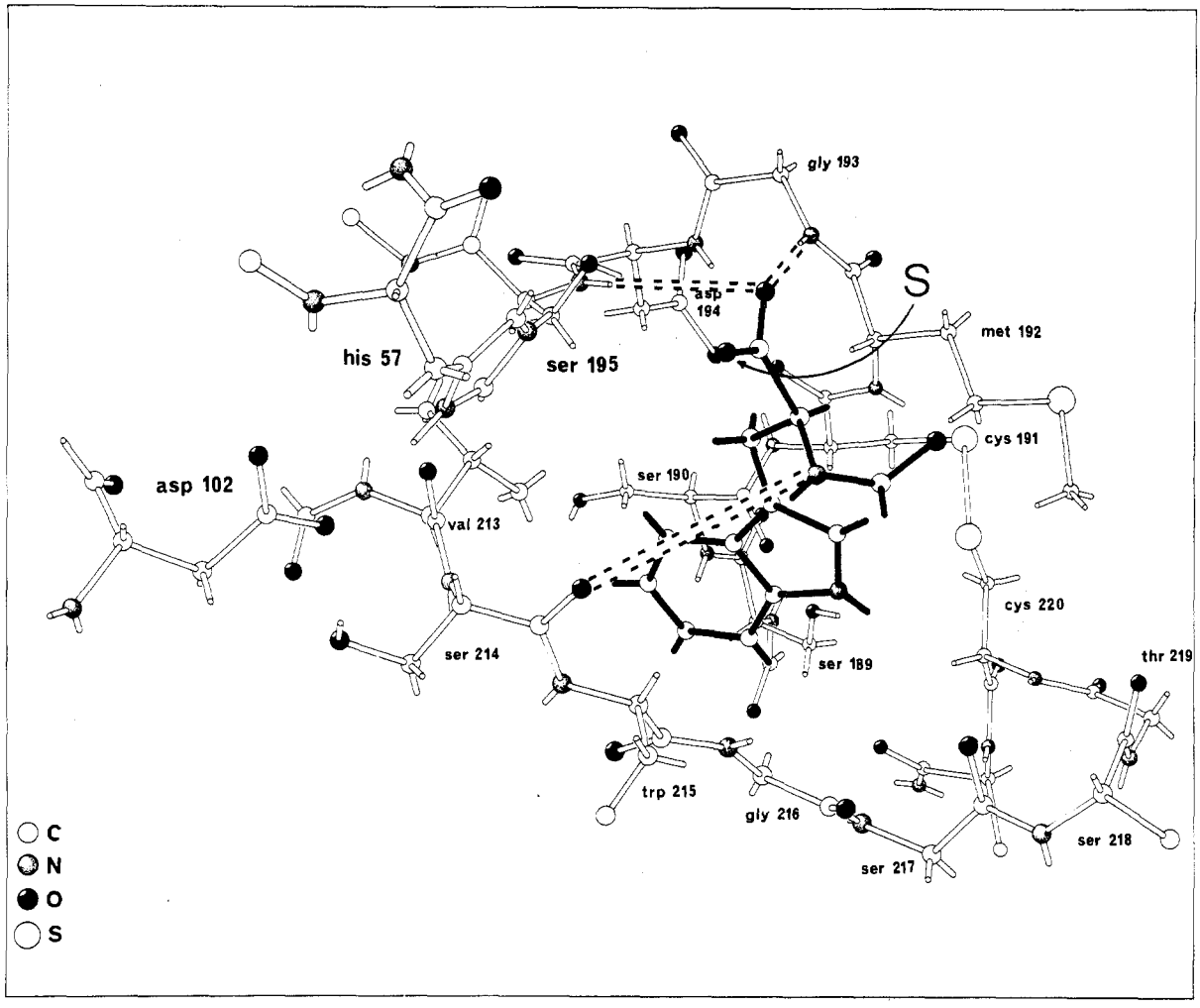


Figure 3. Complex of chymotrypsin with formyl-L-tryptophan, as determined crystallographically by Steitz et al.[37] In a substrate, the scissile bond would be adjacent to the carboxylate oxygen S.

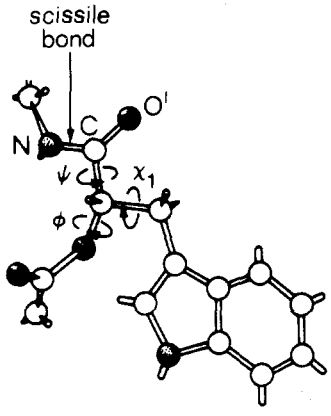


Figure 4. The conformational angles χ_1, ϕ, and ψ which determine the orientation of the scissile bond when an aromatic side chain is located in the primary substrate binding pocket.

tive range of ψ, which brings the N of the scissile bond into this position, is unfavorable for hydrogen bonding.) If the carbonyl oxygen accepts these hydrogen bonds, the NH of the scissile bond is brought close to His-57 and Ser-195. The scissile peptide bond lies in a plane almost perpendicular to the plane of the charge relay system (defined by the carboxyl oxygen, the center of the imidazole ring, and O^{γ}(195)).

These interactions are sufficient to define a conformation for a small molecular substrate, for example *N*-acetyl-L-tryptophan amide (Figure 4). For trypsin substrates, an exactly similar mode of binding exists, in which a lysine or arginine side chain is stabilized by an acid group (Asp-189) which is at the bottom of the substrate binding pocket in trypsin.[29,37] A much smaller pocket exists in elastase, due to the presence of two bulky side chains (Val-215 and Thr-226) which replace glycine in trypsin and chymotrypsin.[28] These will fill up most of the pocket, which can accept the methyl side chain of an alanine residue, but little more. This small binding pocket does not completely define the direction of the C^{α}–C^{β} bond, nor does it give enough binding energy by comparison with other modes of binding. Secondary substrate interactions are far more important for elastase, which functions most efficiently with favorable *polypeptide* substrates.[39-41]

An extremely similar binding site exists for subtilisin, except that the primary binding pocket is less well defined, being more clearly described as a shallow groove.[9]

The highly developed primary binding site made it relatively easy to study the binding of substrates to

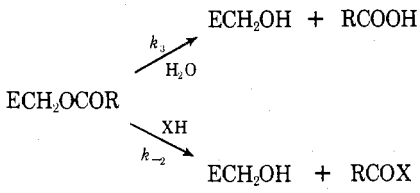
(38) R. Henderson, *J. Mol. Biol.*, **54**, 341–354 (1970).

(39) D. Atlas, S. Levit, I. Schechter, and A. Berger, *FEBS Lett.*, 11, 281–283 (1970).

(40) A. Gertler and T. Hofmann, *Can. J. Biochem.*, 48, 384–386 (1970).

(41) R. C. Thompson and E. E. Blout, *Proc. Natl. Acad. Sci. U.S.A.*, **47**, 1341–1355 (1970).

the enzyme on one side of the scissile bond. The binding of the other part of the substrate, X, has to be studied by less direct means. Scheme I shows the course of a reaction in which the solvent contains only water. If the solution is enriched in XH, the acyl-enzyme can break down in one of two ways:



and the relative efficiencies of various XH as deacylating agents can be compared.[42] Glycinamide and alaninamide were shown to be 10–100 times more potent deacylating agents than water, while hydrazine, a very reactive nucleophile in nonenzymatic reactions, was only as effective as water. This indicated a preferred mode of binding of a peptide group X to the enzyme, believed to involve hydrophobic contacts with the cystine bridge 42–58. There was no evidence, however, of strong binding of alaninamide to the acyl-enzyme. The binding energy was thus responsible for accelerating the reaction, rather than assisting with binding the product XH.

Two crystallographic studies of complexes of trypsin with protein trypsin inhibitors[24,25] confirmed these ideas in a striking way. The inhibitors bind very much like a peptide substrate, but the deacylation step is prevented (or diminished) by the exclusion of water from the active site and by the high stability of the complex. Both the studied inhibitors, although completely different molecules, make the anticipated interactions on both sides of the scissile bond. But in order to make all these interactions simultaneously, the atomic arrangement at the active site is forced into a very crowded situation. Both studies showed that this crowding results in stabilization of the tetrahedral intermediate (Scheme II), in which four atoms are all within covalent bond distance of the carbonyl carbon.[24] In this conformation, the carbonyl oxygen is optimally aligned toward the two hydrogen-bond donors NH-193 and NH-195. The partial negative charge which it probably carries will strengthen these bonds and help in delocalizing the buried negative charge.

It has been found that many inhibitors bind about equally strongly to anhydrotrypsin, a form of trypsin in which Ser-195 is converted to dehydroalanine.[43,44] In this case no chemical bond can be made between enzyme and inhibitor; but the removal of the serine O^{γ} relieves the overcrowding at the active site, and in other respects the close fit between the surfaces of enzyme and inhibitor is not disturbed.[25] Crystallographic analysis confirms this.[45] The chemical changes at the active site, stabilizing the tetrahedral intermediate, involve little *net* energy change, since the affinities of inhibitors for trypsin and anhydrotrypsin are similar.[43,44] The energy gained from the two polarized hydrogen bonds from NH-193 and NH-195, and from the delocalization of the buried negative charge, evidently balances the energy of formation of the strained bond between enzyme and inhibitor (see below).[25]

Transition-State Theory[46–49]

In each step of a chemical reaction, reactants pass from one relatively stable state to another through a state of higher energy. The state of highest energy through which the reactants must pass, to get from one stable state to another, is the transition state. The additional free energy needed to reach the transition state is the activation energy of the reaction.

Part of the activation energy is enthalpic: for example, the work which needs to be done to bring two atoms close enough that a covalent bond begins to be formed between them. Another part is entropic, and expresses the additional order which needs to be imposed on the system to reach the transition state: for example, the chance that two atoms would approach sufficiently close to react. Obviously the entropic component for a unimolecular reaction of two atoms which can collide by bond rotation is less than it is for a bimolecular reaction of two atoms floating freely in a box.

The rate of a chemical reaction depends on the activation energy, according to the Arrhenius equation. Although an absolute rate can be calculated in some idealized cases, these are rare. We do not know enough about the modes of vibration of the reactants, nor about the range of possible configurations which will still allow the reaction to occur. In enzyme reactions the correct orientation of the substrate can greatly reduce the entropy of activation.[46] The charge distribution at the active site and the geometry of the binding site reduce the enthalpy of the transition state.

The appearance of a tetrahedral adduct in the two crystalline trypsin–inhibitor complexes[24,25] gives clear evidence about the catalytic mechanism of the enzyme. The tetrahedral form, which in normal hydrolytic reactions exists too briefly to be observed, has been stabilized to exist permanently in crystals. Huber's results[24] show that the stabilized form has an abnormal bond length—a form which might be assumed to be even less stable. The observed conformation gives us some idea of how the reduction in the activation energy has been achieved, as detailed at the end of the last section.

The energy profile as the tetrahedral complex develops must be rather flat. Probably the O^{γ}–C bond can only shorten to a normal value of 1.5 Å in the tetrahedral intermediate by straining adjacent bond angles. In the trypsin–inhibitor complexes, where numerous secondary interactions are optimal, the complex is actually stabilized by 17 kcal with its abnormally long O^{γ}–C bond. Small substrates of chymotrypsin, which make a few favorable secondary interactions, need 10–15 kcal of activation energy for acyl-

(42) A. R. Fersht, D. M. Blow, and J. Fastrez, *Biochemistry*, **12**, 2035–2041 (1973).
(43) H. Ako, R. J. Foster, and C. A. Ryan, *Biochemistry*, **13**, 132–139 (1974).
(44) J.-P. Vincent, M. Peron-Renner, J. Pudles, and M. Lazdunski, *Biochemistry*, **13**, 4205–4211 (1974).
(45) R. Huber, W. Bode, D. Kukla, U. Kohl, and C. A. Ryan, *Biophys. Struct. Mech.*, 1, 1–13 (1975).
(46) M. I. Page and W. P. Jencks, *Proc. Natl. Acad. Sci. U.S.A.*, **68**, 1678–1683 (1971).
(47) R. Wolfenden, *Acc. Chem Res.*, **5**, 10–18 (1972).
(48) G. E. Lienhard, *Science*, **180**, 149–154 (1973).
(49) A. R. Fersht, *Proc. Roy. Soc. London, Ser. B*, **187**, 397–407 (1974).

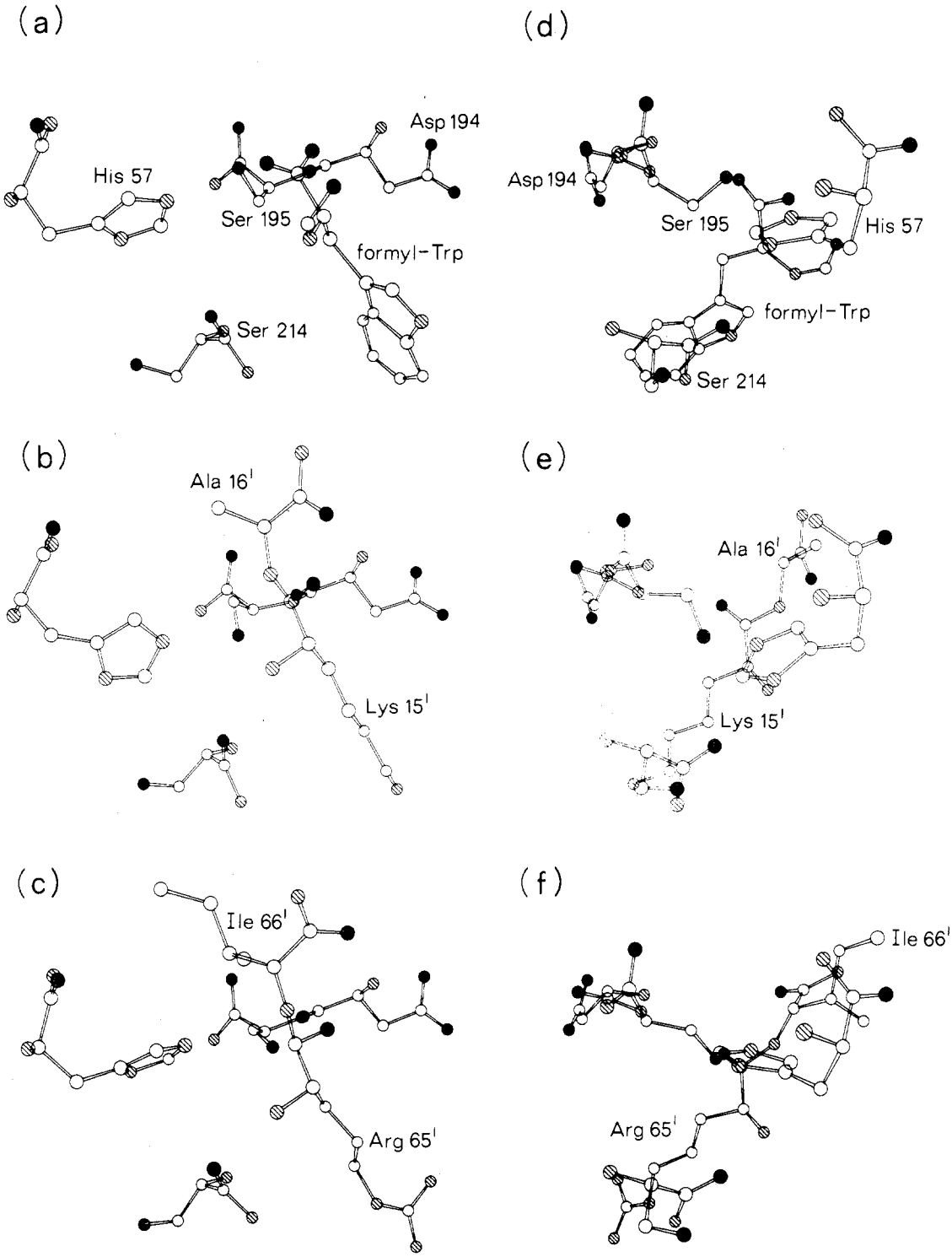


Figure 5. The relative positions of key amino acids at the active site of (a) a complex of bovine chymotrypsin with the substrate analogue formyl-L-tryptophan,[37] (b) the complex of bovine trypsin with bovine pancreatic trypsin inhibitor,[24] (c) the complex of pork trypsin with soybean trypsin inhibitor;[25] (d–f) show the same atoms from a viewpoint 90° away. The orientation and structures have been arranged to give, as nearly as possible, an exactly comparable view in each case. These diagrams are derived from three completely independent crystallographic structure determinations of three different pairs of molecules. The largest difference is in the position of the histidine ring, confirming the idea that this side chain is able to move after the substrate is bound. With this exception, all homologous atoms are similarly placed within less than 1 Å. I thank Dr. R. Huber for providing the unpublished coordinates used to generate (b) and (e).

ation,[20] and the tetrahedral complex has only a transient existence. These differences of stabilization energy can be reasonably explained by the different contact areas in the complexes which the enzyme forms with an inhibitor or a small substrate.[50]

Proposed Catalysis in Subtilisin

Subtilisin has a similar charge-relay system, and there is a similar arrangement of reacting groups around the substrate.[51] It has been proposed that the subtilisin substrate moves by about 1 Å in the transi-

(50) C. H. Chothia and J. Janin, *J. Mol. Biol.*, **100**, 197–212 (1976).

(51) J. D. Robertus, R. A. Alden, J. J. Birktoft, J. Kraut, J. C. Powers, and P. E. Wilcox, *Biochemistry*, **11**, 2439–2449 (1972).

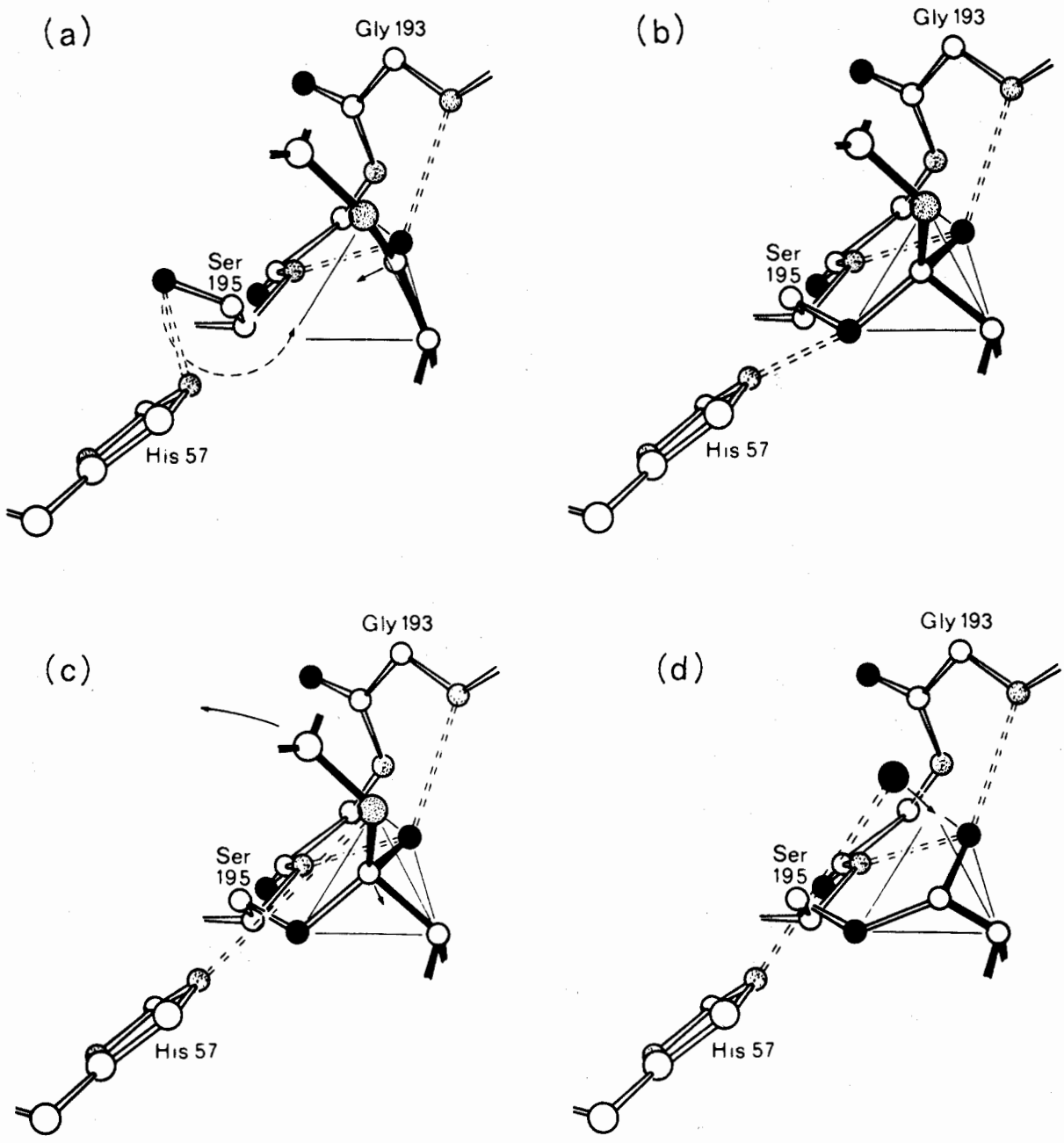


Figure 6. A stereochemically realistic series of drawings,[54] indicating the proposed steps in formation of the acyl-enzyme, following the scheme proposed previously.[38,53] Bonds between atoms in the substrate are indicated in black. (a) Rotation of O^{γ} (Ser-195) to attack the carbonyl carbon of the substrate, which moves into a tetrahedral conformation. This must be accompanied by some displacement of His-57, since otherwise O^{γ}-195 would pass too close to $C^{\epsilon 1}$-57.[55] (b) The tetrahedral intermediate, as observed in the inhibitor complexes. (c) The hydrogen bond from the displaced His-57 is transferred from O^{γ}-195 to the leaving group X. After the proton has been transferred from His-57 to X, the leaving group is expelled. (d) The tetrahedral carbon moves into the plane defining the ester group of the acyl enzyme. The leaving group is replaced by a water molecule hydrogen bonded to His-57, which can move in to attack the ester group. Further steps of the process would be indicated by (c), (b), (a), where this water molecule (or an OH group) takes the place of the leaving group.

tion from the Michaelis complex to the acyl-enzyme. In this process, the distance from the substrate NH to CO-125 (the analogue of chymotrypsin's CO; see Figure 2) shortens from about 4 Å to a hydrogen-bond distance. It has been suggested that the tighter binding of the substrate which results from this movement provides a driving force for catalysis.[52] By analogy, it is argued that a similar process might operate in trypsin and chymotrypsin.

Such a movement appears unlikely in chymotrypsin, since the substrate-binding pocket greatly restricts possible movements of the substrate. Moreover, the chemical changes in the various steps of the reaction provide no obvious mechanism which would explain a change in the length of the hydrogen bond to CO-214.

The structural results for the two trypsin–inhibitor complexes confirm that the substrate movement proposed for subtilisin does not occur in trypsin. In Figure 5 we show comparable views of the complex of chymotrypsin with the pseudosubstrate formyl-L-tryptophan (which we take as a model of a Michaelis complex)[37] and of the two trypsin inhibitor complexes.[24,25] The C^{α} atom in the mouth of the primary

(52) J. D. Robertus, J. Kraut, R. A. Alden, and J. J. Birktoft, *Biochemistry*, 11, 4293–4303 (1972).

binding site is in the same position (relative to mean enzyme coordinates) within experimental error. The NH···CO-214 hydrogen bond is shorter in both inhibitor complexes (3.3[24] and 3.2 Å[25]) than in the pseudo-substrate complex (at least 3.5 Å), but this difference is probably at the level of experimental error.

Stereochemistry of an Enzyme-Catalyzed Reaction[38,53,54]

When the scissile bond lies in the orientation dictated by the primary binding site and by the hydrogen bonds made from NH-193 and NH-195 to the carbonyl oxygen O′, the carbonyl carbon C′ is near C^{β}-195, and the amide nitrogen N′ is close to $N^{\epsilon 2}$-57 and O^{γ}-195. The leaving group X cannot be brought into the position observed in the trypsin–inhibitor complexes, because these contacts would become too short.

Development of the Tetrahedral Intermediate (Figure 6a). In the first step of the reaction, the serine O^{γ} rotates about its C^{α}–C^{β} bond. This rotation must involve a small displacement of His-57, since otherwise O^{γ} would pass too close to $C^{\epsilon 1}$ 57.[55] When O^{γ} begins to come in contact with C′, it is moving roughly perpendicular to the plane defined by the three ligands of C′. On closer approach to C′, bonding will develop, and as a consequence C′ will be displaced toward O^{γ}, out of this plane and into a tetrahedrally coordinated conformation.

In this conformation, observed in the trypsin inhibitor complexes, the $N^{\epsilon 2}$–O distance is still suitable for a hydrogen bond (Figure 6b). It is now possible for the full leaving group interactions to be made, because of the closer approach of C′ to C^{β}-195 which follows from the C′–O^{γ}–C^{β} bonds. The relaxation of the planarity of the peptide group which results from formation of the additional bond may assist in developing the leaving group interactions.

At this stage we may assume the proton originating from the serine $O^{\gamma}H$ to be localized on $N^{\epsilon 2}$, leaving His-57 and Asp-102 electrically neutral, and with a negative charge developing on O′, as in Scheme II.

Collapse of the Tetrahedral Intermediate to the Acyl Enzyme (Figure 6c). The structures of the inhibitor complexes show that $N^{\epsilon 2}$ is too far from N′ for a hydrogen bond, and probably His-57 must swing back slightly toward the substrate, away from O^{γ}. During the process the hydrogen bond is transferred from O^{γ} to N′. The proton can now be transferred onto N′, initiating the collapse of the tetrahedral intermediate toward the acyl-enzyme. Protonation of N′ implies the repulsion of N′ from C′, leaving C′ three-coordinated, so that it will be pulled a few tenths of an ångström into the plane of O^{γ}, O′, and C^{α}. This plane also includes C^{β}-195, as is usually found in esters.

The repulsion of N′ from C′ must move it far enough to bring it to a nonbonded distance from C′, while accommodating the proton accepted from $N^{\epsilon 2}$. The movement of N′ is sufficient to dislodge the interactions made by the leaving group X, facilitating diffusion of the leaving group into the solvent. As soon as possible, N′ will pick up a further proton from the solvent, becoming NH_3^+. The Asp-His system again polarizes His-57 because $N^{\delta 1}$ is permanently protonated (above pH 7) and $N^{\epsilon 2}$ is now deprotonated. This polarization causes a water molecule to be strongly hydrogen bonded to $N^{\epsilon 2}$ at site W (Figure 6d).

Deacylation. This water molecule also forms a hydrogen bond to CO-40, and in this position its lone-pair electrons point toward C′, from a direction roughly perpendicular to the plane of the ester group.[38] The attack of these electrons on C′ causes it to revert to four-coordination, with development of a negative charge on O′ and movement of C′ back to its "tetrahedral" position. The tetrahedral conformation can collapse by expulsion of O^{γ}. The proton is transferred from the water molecule first to $N^{\epsilon 2}$, then to O^{γ}, as His-57 and O^{γ} move back to their starting positions. C′ and its ligands now form a carboxylate group, and the remaining product of hydrolysis leaves the enzyme in this form.

Conclusion

The activity of chymotrypsin can be considered to derive from several different effects: (i) *chemical apparatus* (the charge relay system) which, by polarizing hydrogen bonds facilitates the proton transfer steps which occur in different steps of the reaction; (ii) a *binding site,* which immobilizes the scissile bond of a substrate at the active site; (iii) a precise *orientation* between the two, which permits each reaction step to involve only a small rotation about a single bond, and which aligns the attacking groups near to optimal directions; (iv) *lowering the activation energy* of the reaction by features which reduce the energy of the enzyme–substrate complex, at the transition state.

All of these effects are, however, different aspects of the structure of the active site, and cannot really be separated from one another. Together they increase the rate of the hydrolytic reaction and cause the enzyme to function as a catalyst.

The rather simple nature of the reaction, and the small number of atoms which need to move, allow us to define the process in detail.

(53) R. Henderson, C. S. Wright, G. P. Hess, and D. M. Blow, *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 63–70 (1971).

(54) D. M. Blow and J. M. Smith, *Philos. Trans. R. Soc. London, Ser. B*, **272**, 87–97 (1975).

(55) Huber et al.[24] have suggested that, in the active enzyme at neutral pH, O^{γ} is in a different orientation from that observed in crystals, for which this movement of His-57 would not be needed.