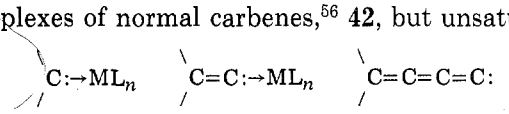
interest for they possess added functionality and hence the potential for further synthetic manipulation. A great deal of work has been done on transition-metal complexes of normal carbenes,[56] **42**, but unsaturated



42 **43** **44**

carbene–transition-metal complexes,[57] **43**, need further attention.

(56) D. J. Cardin, B. Cetinkayser, and M. F. Lappert, *Chem. Rev.*, **72**, 545 (1972); E. O. Fischer, *Adv. Organomet. Chem.*, **14**, 1 (1976).
(57) A. N. Nesmeyanov et al., *J. Organomet. Chem.*, **110**, C36 (1976); **137**, 55, 69 (1977); R. B. King, *Ann. N.Y. Acad. Sci.*, **239**, 171 (1974); O. S. Mills and A. D. Redhouse, *J. Chem. Soc. A*, 1282 (1968).

More work will be forthcoming in the area of synthetic applications of alkylidenecarbenes, particularly in the preparation of strained small-ring compounds. Extended carbenes, such as **44** and others, and their chemistry will be explored.

Vinyl triflates will no doubt play a continuing role in some of these developments. Perhaps the greatest unrealized potential of vinyl triflates is their full development as useful synthetic reagents.

I wish to acknowledge the experimental skill of my co-workers, as cited, whose effort and dedication made this work possible. Financial support by the Petroleum Research Fund, administered by the American Chemical Society, the Research Corporation, and NIH Research Grant No. 1-RO-1-CA16903-01 and -02 are gratefully acknowledged.

Structural Basis of the Activation and Action of Trypsin[1]

ROBERT HUBER* and WOLFRAM BODE

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, Germany

Received March 23, 1977

A hundred years ago, W. Kühne first introduced the word "enzym" to describe the pancreatic protease trypsin.[2,3] Trypsin turned out to be an enzyme of utmost importance. It is a prominent member of a whole family of functionally and structurally related digestive enzymes. Among these are chymotrypsin and elastase, which have also been analyzed as to their three-dimensional structures.[4,5]

Trypsin forms the functional principle of some large and highly specific proteases involved in blood clotting[6] and complement binding.[7] Amino acid sequence studies of some of these proteases show the presence of a trypsin-like core with large segments attached to it which modify the specificity and are responsible for interaction with other macromolecules of the system.[8]

Most of the individual steps in the cascade reactions leading to blood clotting or complement binding are specific proteolytic cleavages liberating and activating yet another protease. Also, in this respect trypsin is a prototype as it is biosynthesized as an inactive precursor, trypsinogen, which is activated by limited proteolysis.[9] Limited proteolysis is also involved in phage maturation.[10] Limited proteolysis is indeed a most important regulatory phenomenon.

A recent fascinating finding was that some of the naturally occuring protease inhibitors are liberated from inactive precursors (pro-inhibitors) by limited proteolysis.[11] This indicates the existence of a regulatory hypercycle involving proteases and their inhibitors: proteases activate, but also destroy (by proteolytic action on temporary inhibitors[12]) their inhibitors.

A common effector in protease action is calcium, which possibly acts by stabilizing the three-dimensional structure.[13] Trypsin has a well-defined calcium binding site,[14] and calcium influences its functional and structural properties.

Because trypsin is in so many ways a prototype, a detailed understanding of its structure and function is

Robert Huber received the degree Dr. rer. nat. from the Technical University of Munich, and then joined the staff at the Max-Planck-Institut für Eiweiss- und Lederforschung. In 1972, he was appointed Director of the Department of Structure Research at the Max-Planck-Institut für Biochemie in Martinsried. His research has focused on studies of the structure and function of proteins, enzymes, and antibodies, and on methods in protein crystallography.

Wolfram Bode studied chemistry at the Universities of Göttingen, Tübingen, and München. After he received the degree Dr. rer. nat. at München, he joined J. Engel's research group at the Max-Planck-Institut für Eiweiss- und Lederforschung to study bacterial flagella. Dr. Bode moved to the Max-Planck-Institut für Biochemie in 1972. His research concerns the structure and function of proteolytic enzymes.

(1) See IUPAC–IUB Commission on Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970), for abbreviations. Other abbreviations used: PTI, pancreatic trypsin inhibitor (Kunitz); STI, soybean trypsin inhibitor; e, electrons; DIP, diisopropylphosphoryl; rms, root mean square; K_S, dissociation constant of the enzyme–substrate complex; conformational angles are defined according to IUPAC–IUB;[1] residues of the inhibitor are indicated by (I) or underlined.
(2) W. Kühne, *Verhl. Naturhist.-Med. Ver. Heidelberg*, 1, 194–198 (1877).
(3) H. Gutfreund, *FEBS Lett.*, **62**, (Supplement) E1–E12 (1976).
(4) B. W. Matthews, P. B. Sigler, R. Henderson, and D. M. Blow, *Nature (London)*, **214**, 652–656 (1967).
(5) D. M. Shotton and H. C. Watson, *Nature (London)*, **225**, 811–816 (1970).
(6) E. W. Davie and K. Fujikawa, *Annu. Rev. Biochem.*, **44**, 799–829 (1975).
(7) H. J. Müller-Eberhard, *Annu. Rev. Biochem.*, **44**, 697–724 (1975).
(8) S. Magnusson, L. Sottrup-Jensen, T. E. Petersen, G. Dudek-Wojciechowska, and H. Claeys, in "Miami Winter Symposia", Vol. 11 "Proteolysis and Physiological Regulation", D. W. Ribbons and K. Brew, Ed., Academic Press, New York, N.Y., 1976, pp 203–239.
(9) H. Neurath and E. Davie, *J. Biol. Chem.*, **212**, 515–529 (1955); P. Desnuelle and C. Fabre, *Biochim. Biophys. Acta*, **18**, 49–57 (1955).
(10) T. Isobe, L. W. Black, and A. Tsugita, *J. Mol. Biol.*, **110**, 165–177 (1977).
(11) G. Bretzel and K. Hochstrasser, *Hoppe-Seyler's Z. Physiol. Chem.*, **357**, 487–489 (1976).
(12) M. Laskowski, Sr., and F. C. Wu, *J. Biol. Chem.*, **204**, 797–805 (1953).
(13) R. H. Kretsinger, *Annu. Rev. Biochem.*, **45**, 239–266 (1976).
(14) W. Bode and P. Schwager, *J. Mol. Biol.*, **98**, 693–717 (1975).

0001-4842/78/0111-0114$01.00/0

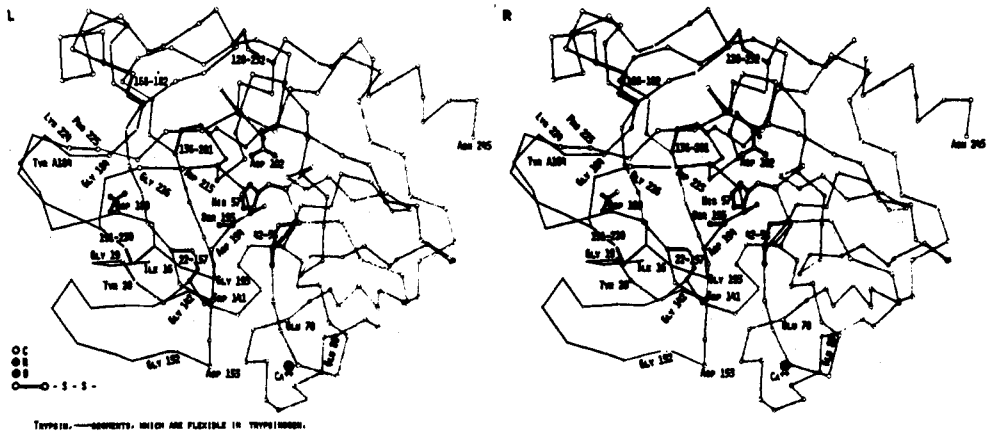


Figure 1. Stereo drawing of the C^{α} carbon positions in trypsin. Residues linked by single lines are flexible in trypsinogen. Residues linked by double lines are fixed. The catalytic residues as well as the hinge residues are indicated.

of general relevance. The structure and mechanism of the closely related protease chymotrypsin have recently been lucidly described in an Account by Blow,[15] who has also studied the complex formed by STI trypsin inhibitor and trypsin.[16] We focus in this Account on the structural basis of the activation mechanism of trypsin and on different views of the catalytic mechanism. We note that the crystal structures of DIP-inhibited trypsin and trypsinogen have been determined independently.[17]

We have determined various structures of the bovine trypsin family and refined them at the highest resolution allowed by the crystalline order ranging from 1.9 to 1.5 Å: trypsin (in its benzamidine-inhibited and free forms at pH 8 and pH 5[14,18]), trypsin inhibitor PTI,[19] trypsinogen,[20] trypsin inhibitor–trypsin complex,[18,21] trypsin inhibitor–anhydrotrypsin complex,[22] trypsin inhibitor–trypsinogen complex,[20,23] trypsin inhibitor–trypsinogen complex + the dipeptide Ile-Val.[20,23] These four different crystal structures and their various isomorphous variants provide a detailed yet static view of trypsin activation and action. The accuracy of the refined models is estimated to about 0.1 Å for well-defined segments from crystallographic considerations and, most objectively, from a comparison of molecular models for different crystal structures.

Trypsin and Trypsinogen Structure

The amino acid residues of trypsin fold into a globular molecule. The secondary structure is characterized by predominance of the β structure, but little helix. The β structure is organized in two barrels. As in chymotrypsin,[4] the interior of the barrels is packed with hydrophobic amino acid side chains, as densely as observed in organic crystals.[24,25] There are some hollows in trypsin filled with water molecules. These water molecules are integral constituents of the molecular structure and have been found in free trypsin,[14] in trypsin inhibitor–trypsin complex,[18,21] in trypsinogen,[20] and also in closely similar positions in chymotrypsin.[26]

The pro-enzyme trypsinogen is converted to trypsin by cleaving the N-terminal activation hexapeptide.[9] Both molecules have identical structures for about 85% of the chain. The deviation of main chain atoms in this part is 0.2 Å, insignificant with respect to the error level of both analyses, but the rest of the molecule is entirely different, consisting of four segments tightly interdigitating in trypsin: the N-terminus to Gly-19, Gly-142 to Pro-152, Gly-A184 to Gly-193, and Gly-216 to Asn-223. We call these segments, indicated in Figure 1, the "activation domain".

There is no significant electron density for these segments in the Fourier map of trypsinogen because they are either flexibly wagging in the crystalline state or adopting (at least three) different conformations statistically. The latter deduction stems from significance considerations of the electron density map. Both situations would produce a similar effect on the appearance of the electron density, either smearing it out over a large volume or distributing it on several sites, but with low weight. A decision as to which situation holds might be provided by spectroscopic methods or low-temperature crystallographic analysis.

Flexibility starts rather abruptly in single residues. It is remarkable that in five of the seven hinges where chains become flexible a glycine residue is located. These glycines are conserved in serine proteases, suggesting similar structural transitions upon activation. Glycine, which has no side chain for other parts of the molecule to grab, is a preferred candidate to mediate flexibility. Three of the seven hinges in trypsinogen have an aromatic residue adjacent to glycine. These residues are fixed and might serve as anchors. The segments around 190 and 220 are part of the pocket binding the specificity side chain of the substrate. They are connected by the disulfide 191–220 which is also flexible and can be selectively reduced.[27]

(15) D. M. Blow, *Acc. Chem. Res.*, **9**, 145–152 (1976).
(16) R. M. Sweet, H. T. Wright, J. Janin, C. H. Chothia, and D. M. Blow, *Biochemistry*, **13**, 4212–4228 (1974).
(17) R. M. Stroud, L. M. Kay, and R. E. Dickerson, *J. Mol. Biol.*, **83**, 185–208 (1974); A. A. Kossiakoff, J. L. Chambers, L. M. Kay, and R. M. Stroud, *Biochemistry*, **16**, 654–666 (1977).
(18) W. Bode, P. Schwager, and R. Huber in ref 8, pp 43–76.
(19) J. Deisenhofer and W. Steigemann, *Acta Crystallogr., Sect. B*, **31**, 238–250 (1975).
(20) (a) W. Bode, H. Fehlhammer, and R. Huber, *J. Mol. Biol.*, **106**, 325–335 (1976); (b) H. Fehlhammer, W. Bode, and R. Huber, *ibid.*, **111**, 415–438 (1977). Figures 3, 4, 6, and 7 of ref 20b are reprinted with permission. (c) 1977 Academic Press.
(21) R. Huber, D. Kukla, W. Bode, P. Schwager, K. Bartels, J. Deisenhofer, and W. Steigemann, *J. Mol. Biol.*, **89**, 73–101 (1974); A. Rühlmann, D. Kukla, P. Schwager, K. Bartels, and R. Huber, *ibid.*, **77**, 417–436 (1973).
(22) R. Huber, W. Bode, K. Kukla, U. Kohl, and C. A. Ryan, *Biophys. Struct. Mech.*, **1**, 189 (1975).
(23) W. Bode and R. Huber, *FEBS Lett.*, **68**, 231–235 (1976).
(24) B. Lee and F. M. Richards, *J. Mol. Biol.*, **55**, 379–400 (1971).
(25) C. H. Chothia and J. Janin, *J. Mol. Biol.*, **100**, 197–212 (1976).
(26) J. J. Birktoft and D. M. Blow, *J. Mol. Biol.*, **68**, 187–240 (1972).
(27) J. R. Knights and A. Light, *J. Biol. Chem.*, **251**, 222–228 (1976).

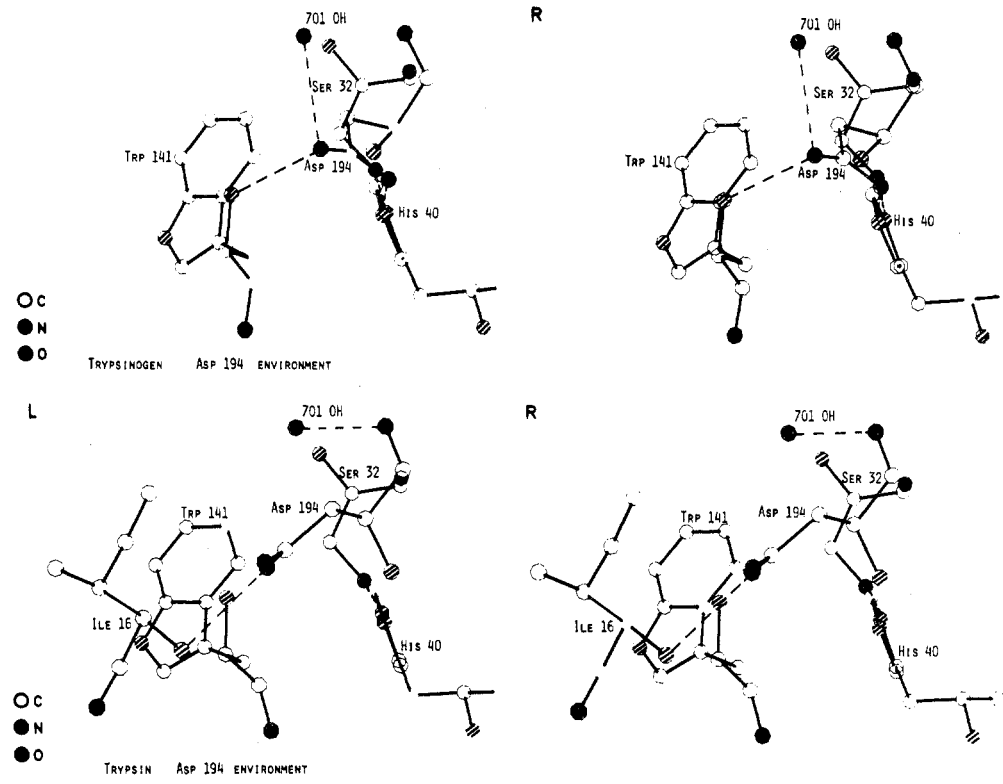


Figure 2. Stereo drawing of the Asp-194 conformation and environment in trypsinogen (a) and trypsin (b).

The triggering event that leads to the formation of the rigid, correctly designed specificity pocket is the conformational change of Asp-194 which forms a link to His-40 in trypsinogen and an internal salt bridge to Ile-16, the newly formed N-terminus, in trypsin (Figure 2a,b). This structural change was first found in chymotrypsinogen.[28]

The chain segments which are flexible in trypsinogen form a tightly interdigitating structural unit in trypsin, as shown in the hydrogen-bonding diagram (Figure 3). There are more than 20 hydrogen bonds cross-linking the segments of the activation domain in trypsin which are lost upon mobilization in trypsinogen. Particularly noteworthy is the hydrogen-bonding network of the Asp-194 carboxylate–Ile-16 ammonium ion pair, which appears to act as a clamp. The hydrogen bonding between the activation domain and the rest of the molecule is weak, consisting of only three linkages. The activation domain appears to be a rather separate unit.

The N-terminus Ile-16 occupies a pocket in trypsin forming a salt link and several hydrogen bonds to other residues of the activation domain and to structure water associated with it (Figure 4). Consequently, the buried Ile N-terminus has a high pK.[29,30] Titration of this group leads to inactivation and possibly to a species resembling trypsinogen. This suggests a conformational linkage between Ile-16 and the specificity pocket. Indeed, it is possible to demonstrate such a linkage rigorously by inducing the structural transition from both sides. Sufficiently strong binding of a specific inhibitor to trypsinogen rigidifies the specificity pocket *and* the Ile-16 pocket, although there is no Ile-16 N-terminus. Such a species is observed in the trypsinogen–PTI complex.

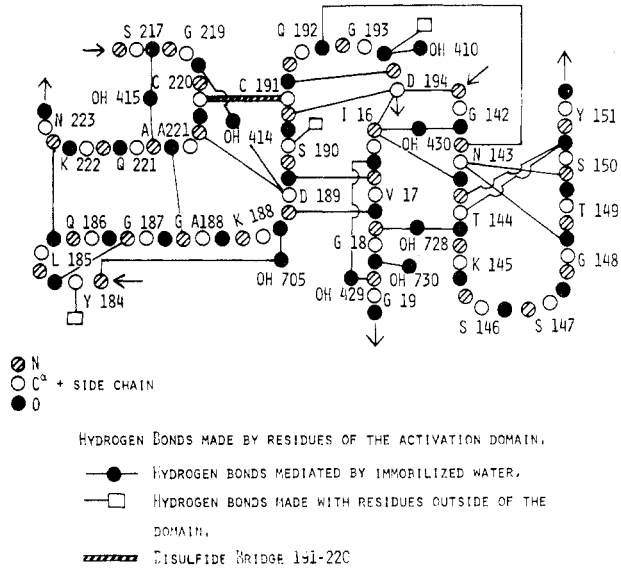


Figure 3. The activation domain in trypsin. (—) Hydrogen bonds between residues of the activation domain, (—□) hydrogen bonds to residues outside the domain.

Rigidification of the Ile-16 binding pocket with concomitant formation of the specificity pocket can also occur on reaction of trypsinogen with a peptide sequentially related to the trypsin N-terminus, i.e., Ile-Val, according to spectroscopic evidence.[23] This interaction is highly specific. Even the closely related Val-Val peptide is 30-fold less active. But the presence of an Ile-Val dipeptide alone is not sufficient. In addition, a strong, specific inhibitor, *p*-guanidobenzoate covalently bound to trypsinogen, is required to bring about the structural transition.[23] The ability of the "foreign" Ile-Val peptide to induce the trypsinogen–

(28) S. T. Freer, J. Kraut, J. D. Robertus, H. T. Wright, and Ng. H. Xuong, *Biochemistry*, **9**, 1997–2009 (1970).
(29) A. R. Fersht, *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 71–73 (1971).
(30) W. E. Spomer and J. F. Wootton, *Biochim. Biophys. Acta*, **235**, 164–171 (1971).

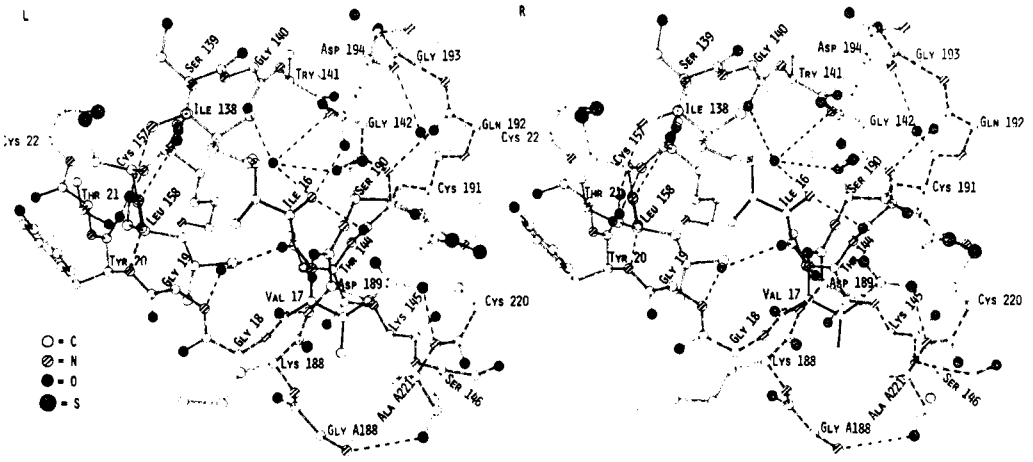


Figure 4. Stereo drawing of the Ile-16 pocket in trypsin. (- - -) Residues which are flexible in trypsinogen.

trypsin structural transition is therefore far inferior to that of the molecule's own Ile-Val N-terminus. This is due to a lower effective concentration (we compare a bimolecular with unimolecular reaction) and probably also to conformational constraints. Similar effects are discussed for rate and binding enhancement in intramolecular and enzymatic reactions.[31-33]

The trypsinogen–PTI complex has been analyzed crystallographically in detail. Here, the Ile-Val binding pocket is empty but binds an added Ile-Val dipeptide with about 8 kcal/mol.[34] The analysis of the Ile-Val addition compounds shows the Ile-Val peptide identical with the N-terminus in the trypsin–PTI complex.

In the trypsinogen–PTI complex the activation domain is structured, except the N-terminus, which remains mobile to Gly-19. Binding of the Ile-Val peptide leads to a stronger fixation of the 142–152 loop indicated by reduced temperature factors (rms vibration amplitudes of 0.58 and 0.46 Å for the complexes without and with dipeptide bound). The trypsinogen–PTI complex is very similar to the trypsin–PTI complex with a mean main chain deviation of 0.15 Å.

Trypsinogen has a fascinating facet as an allosteric enzyme. We may depict a thermodynamic cycle linking the various structures and their equilibria (Figure 5). Apart from minor structural variations in external amino acid side chain conformations due to lattice effects and the different degrees of structuring of the 142–152 loop there are only two different, interconvertible structures. The various species observed crystallographically are A, B, C, D, and E. Equilibrium constants are experimentally determined (large numbers)[23,34-37] or have been inferred (small numbers) from these measurements on the assumption that the equilibria are identical for species with the same structural features. Species F, G, H, and I are unstable, but have to be postulated as intermediates in the

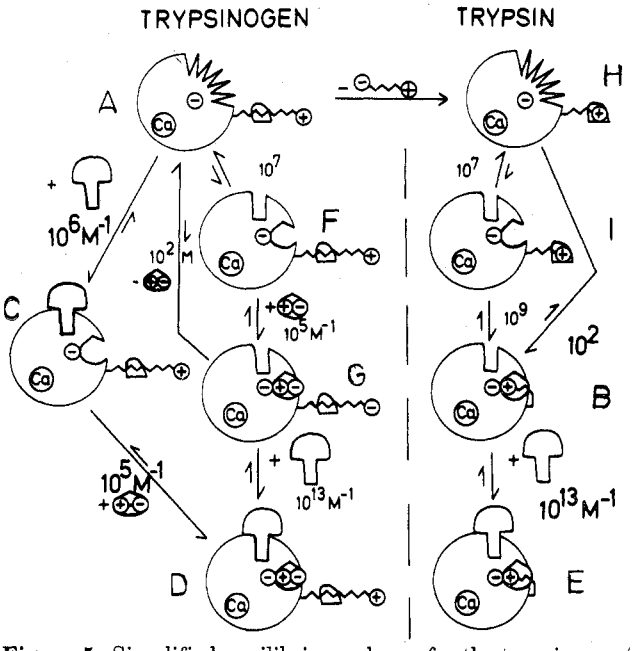


Figure 5. Simplified equilibrium scheme for the trypsinogen/PTI/Ile-Val system (left) and the trypsin/PTI system (right) (〰) flexible segments, (⊕⊖) Ile-Val dipeptide.

thermodynamic scheme. The allosteric properties are mediated by the segmental flexible–rigid transition of the activation domain discussed.

This type of signal transfer is different from what is seen in the oligomeric hemoglobin.[38] Here, and probably in other allosteric oligomeric proteins, the allosteric effects are mediated by a change in quaternary structure.

The inactive trypsinogen represents an incompletely folded trypsin species. As folding is believed to occur domain-wise, trypsinogen might be an intermediate on the folding pathway of trypsin. The final folding of the activation domain is possible only in the presence of the Ile-Val N-terminus.

It is unclear whether the allosteric induction of activity in trypsinogen and related zymogens without activation peptide cleavage is of physiological significance, but there are observations of the development of proteolytic activity prior to activation peptide cleavage in the plasminogen streptokinase complex.[39]

(31) M. I. Page and W. P. Jencks, *Proc. Natl. Acad. Sci. U.S.A.*, **68**, 1678–1683 (1971).
(32) D. E. Koshland, Jr., K. W. Carraway, G. A. Dafforn, J. D. Gass, and D. R. Storm, *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 13–20 (1971).
(33) T. C. Bruice, *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 21–27 (1971).
(34) W. Bode, in preparation.
(35) J. P. Vincent and M. Lazdunski, *Biochemistry*, **11**, 2967–2977 (1972).
(36) J. P. Vincent and M. Lazdunski, *FEBS Lett.*, **63**, 240–244 (1976); U. Quast, in preparation.
(37) B. S. Hartley and D. M. Shotton, *Enzymes, 3rd Ed.*, **3**, 323–373 (1971).
(38) M. F. Perutz, *Nature (London)*, **228**, 726–734 (1970).
(39) S. P. Bajaj and F. J. Castellino, *J. Biol. Chem.*, **252**, 492–498 (1977).

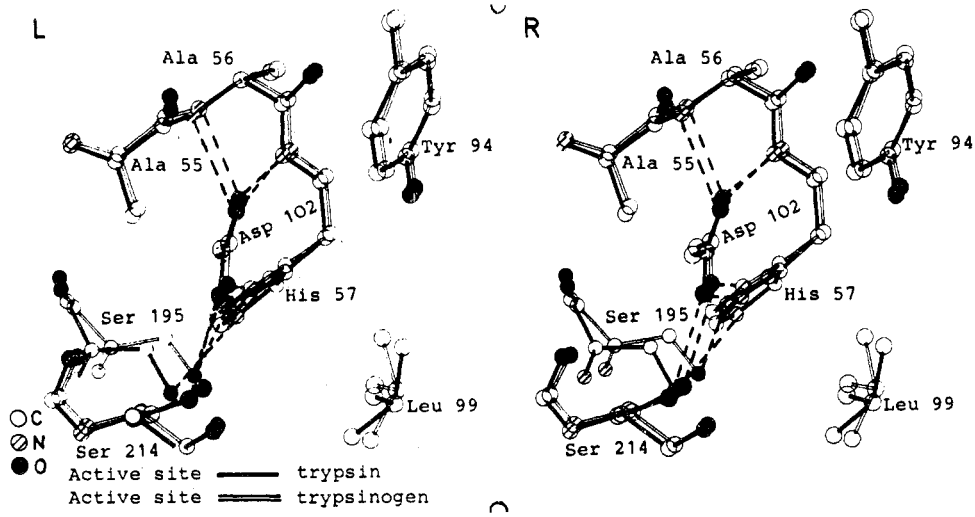


Figure 6. Stereo drawing of the catalytic residues in trypsinogen (═) and trypsin (━) overlaid.

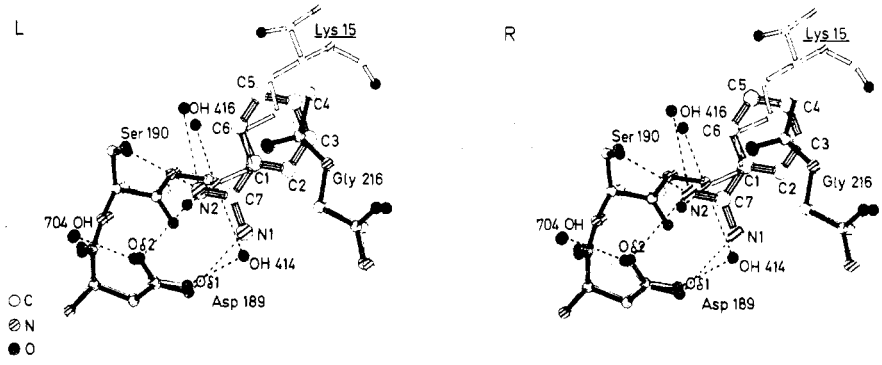


Figure 7. Stereo drawing of the Lys-15(I) binding in the specificity pocket of trypsin a seen in the PTI–trypsin complex (═) overlaid with benzamidine (≡) as seen in benzamidine-inhibited trypsin (━).

Similar phenomena appear to occur in the complement system.[7]

The activation domain in trypsinogen plays an essential role in specific binding of the substrate, as described below. This domain is flexible in trypsinogen and unable to mediate substrate binding. The Gibbs free energy for the conformational transition to trypsin is estimated to be 9 kcal/mol from the difference in binding energy of PTI to trypsin as compared to trypsinogen.[35,36] As K_S for good, specific substrates is around 6 kcal/mol,[40] there should be no binding to trypsinogen. The catalytic residues described below have nearly identical conformations in trypsin and trypsinogen (Figure 6). Such structural features agree with the observation of some (very low) activity of trypsinogen against specific substrates, but high activity against nonspecific substrates.[41]

Inhibitor Structure and Action

The basic pancreatic trypsin inhibitor (PTI) is a small, pear-shaped protein molecule.[19] It is an extremely potent inhibitor of trypsin and other proteases. It binds to trypsin with an association constant of 10^{13} M^{-1}, the highest value known for protein–protein interactions.[35]

It was of fundamental importance to find, with other natural inhibitors[42] and recently also with PTI,[43] that interaction with proteases involves formation of a modified inhibitor species which has the active site peptide bond cleaved. This tells us, in general, that inhibitor–protease interaction involves catalytic action and that the complex is an intermediate in catalytic peptide bond cleavage.

The kinetic reaction scheme and the thermodynamic data of PTI–chymotrypsin and –trypsin interaction have been thoroughly studied.[44-46] There is kinetic evidence for a minimal mechanism

$$E + I \rightleftharpoons L \rightleftharpoons C \rightleftharpoons L^* \rightleftharpoons E + I^*$$

where E is enzyme, I, inhibitor, L, loose complex, C, stable complex, and * identifies the inhibitor species with the active site peptide cleaved. As there is little doubt that peptide bond hydrolysis involves an acyl-enzyme species,[47,48] the equation to the right of C should be expanded to include the acyl intermediate, although there is no definitive kinetic evidence for it. Formation

(40) U. Quast, J. Engel, H. Heumann, G. Krause, and E. Steffen, *Biochemistry*, **13**, 2512–2520 (1974).
(41) M. A. Kerr, K. A. Walsh, and H. Neurath, *Biochemistry*, 14, 5088–5094 (1975).
(42) W. R. Finkenstadt and M. Laskowski, Jr., *J. Biol. Chem.*, **242**, 771 (1967).
(43) H. Jering and H. Tschesche, *Angew. Chem., Int. Ed. Engl.*, **13**, 661–663 (1974).
(44) J. Engel, U. Quast, H. Heumann, G. Krause, and E. Steffen, in "Bayer Symposium V Proteinase Inhibitors", H. Fritz, H. Tschesche, L. J. Greene, and E. Truscheit, Ed., Springer-Verlag, Berlin, Heidelberg, New York, 412–419 (1974).
(45) W. R. Finkenstadt, M. A. Hamid, J. A. Mattis, J. Schrode, R. W. Sealock, D. Wang, and M. Laskowski, Jr., in: Bayer Symposium V "Proteinase Inhibitors", H. Fritz, H. Tschesche, L. J. Greene, E. Truscheit, Eds., Springer Verlag, Berlin, Heidelberg, New York, 1974, pp 389–411.
(46) M. Lazdunski, J. P. Vincent, H. Schwertz, M. Peron-Renner, and J. Pudles, in ref 45, pp 420–431.
(47) G. P. Hess, *Enzymes*, *3rd Ed.*, **3**, 213–248 (1971).
(48) R. Henderson, C. S. Wright, G. P. Hess, and D. M. Blow, *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 63–70 (1972).

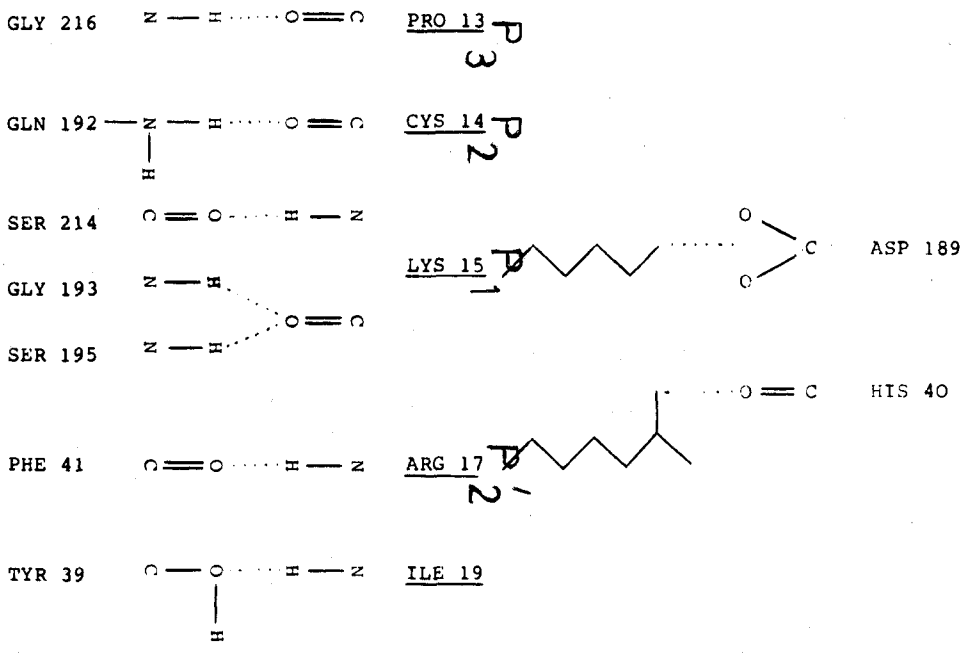


Figure 8. Schematic hydrogen-bonding interactions formed between PTI and trypsin as seen in the complex.

of L is a very fast, entropy-driven reaction, while the conversion of L to C is slow and has also a favorable enthalpy contribution.[44,45] Dissociation rates C → L (and also C → L*, as measured for soybean trypsin inhibitor[45]) are extremely slow at neutral pH. Quite generally, PTI and other natural inhibitors have the characteristics of excellent substrates in the association step. They inhibit by virtue of their slow dissociation.

The Complex

Only a small proportion of both molecules is in contact in the complex: 14 amino acid residues out of 58 of the inhibitor, and 24 amino acid residues out of 224 of trypsin. The contact is characterized by a complicated network of hydrogen bonds and a large number of van der Waals contacts. It is tightly packed.[25]

A dominant interaction is made by the specificity side chain of PTI (Lys-15(I)) inserted into the specificity pocket (Figure 7). The specific interaction occurs between the positively charged Lys-15(I) ammonium group of the inhibitor and the Asp-189 carboxylate of the enzyme. This carboxylate is responsible for the primary specificity of trypsin for positively charged side chains. N^{ζ} of Lys uses fully its hydrogen-bonding capabilities, donating three hydrogen bonds to the carbonyl oxygen of Ser-190 and two water molecules, 416 and 414, one of which is bonded to Asp-189. The binding of a small inhibitor molecule benzamidine to trypsin is overlaid in Figure 7. Benzamidine is a small synthetic inhibitor simulating arginine residues. It forms hydrogen bonds directly to the Asp-189 carboxylate, to Ser-190 O^{γ}, to a water molecule, 416, and to Gly-219 O.

There is a small structural rearrangement of the specificity pocket to fulfill the different hydrogen-bonding requirement of an ammonium group (lysine) and a guanidinium group (benzamidine, arginine). Water-414, which is also present in free trypsin, is expelled when benzamidine binds, and Ser-190 O is attracted by lysine and the peptide plane is slightly rotated.

The specificity side chain held in the specificity pocket is one anchor to fix the inhibitor or substrate at the enzyme surface. A second important binding interaction is between the substrate polypeptide chain and the enzyme. These hydrogen bonding interactions are schematically represented in Figure 8. The substrate main chain from P_3 to P_2' is linked via six hydrogen bonds to the enzyme. The enzymatic split occurs between P_1 and P_1'. An essential interaction appears to be the bifurcated hydrogen bond of P_1 CO with the NH groups of Gly-193 and Ser-195, which form the oxyanion binding hole.

This particular bond, together with the side-chain interaction in the specificity pocket, precisely orients the scissile peptide group with respect to the catalytic residues. It should be recalled that neither the specificity pocket nor the oxyanion binding hole is formed in trypsinogen due to mobility of the segments involved.

The arrangement of the catalytic triad Asp-102, His-57, and Ser-195 was first observed in chymotrypsin, and their essence has been lucidly described in a recent Account.[15] We therefore concentrate on those points where our structural studies might help in clarification.

Trypsin activity as well as the rate of inhibitor association depends critically on pH (pK around 7).[44-46] The group responsible was believed to be His-57, but both NMR studies and chemical modification experiments are controversial in assigning a pK around 7 to His-57 or Asp-102.[49-53] Our structural studies show that free trypsin at pH 5 and 8 is extremely similar. Figure 9 shows the catalytic site residues overlaid. There is virtually no structural variation, except a slight Ser-195 O^{γ} rotation which improves the His N^{ϵ}–O^{γ} hydrogen bond in the low pH form. (This hydrogen bond is 3.26

(49) G. Robillard and M. G. Shulman, *J. Mol. Biol.*, **86**, 519–540 (1974); G. Robillard and M. G. Shulman, *ibid.*, **86**, 541–558 (1974).
(50) M. W. Hunkapiller, S. H. Smallcombe, D. R. Whitaker, and J. H. Richards, *Biochemistry*, **12**, 4732–4742 (1973).
(51) J. L. Markley and M. A. Porubcan, *J. Mol. Biol.*, **102**, 487–509 (1976); J. L. Markley and I. B. Ibanez, *Biochemistry*, submitted for publication.
(52) R. E. Köppe and R. M. Stroud, *Biochemistry*, **15**, 3450–3458 (1976).
(53) A. R. Fersht and J. Sperling, *J. Mol. Biol.*, **74**, 137–149 (1973).

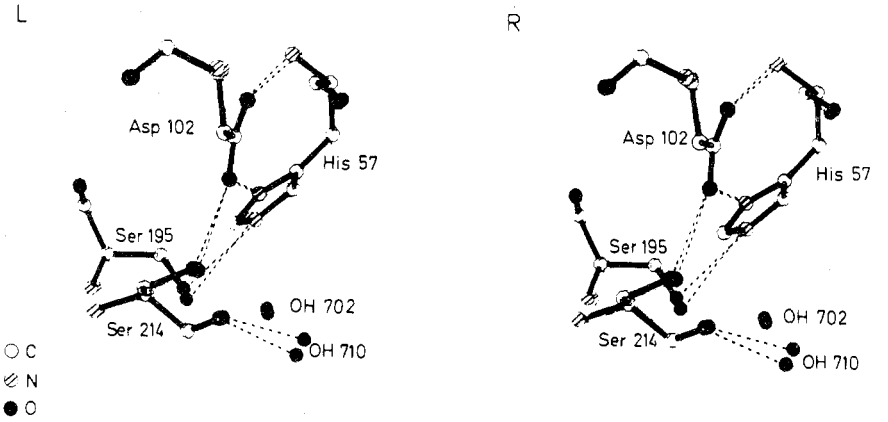


Figure 9. Stereo drawing of the catalytic residues as seen in trypsin at pH 5 (—) and pH 8 (=) overlaid.

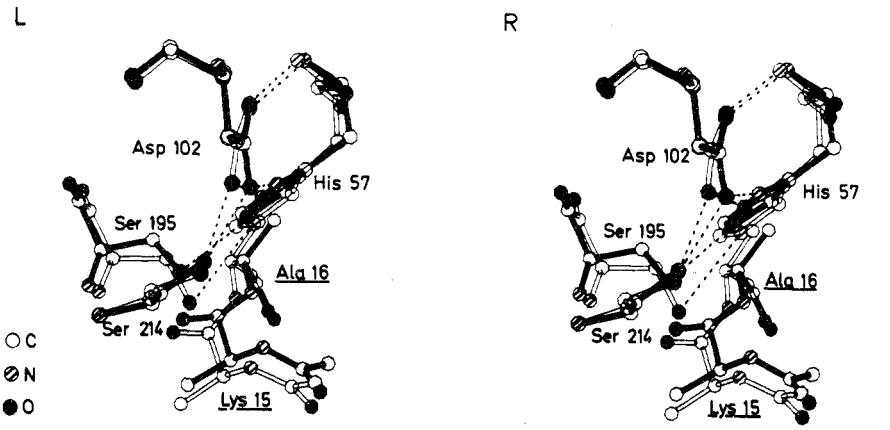


Figure 10. Stereodrawing of the catalytic residues as seen in the complex (—) and in free trypsin (=) overlaid. Ser-195 is pushed out of the way and His-57 seems to follow, improving the hydrogen bond. The inhibitor segment deduced from the free inhibitor conformation is also shown (=). This conformation would be sterically impossible.

Å and rather bent at pH 8, while it is 2.90 Å and more linear at pH 5.) A second change is in the active-site-associated solvent 702, which is characterized by a broad, smeared-out density. The integrated electron density corresponding to this solvent is 30e at pH 5 and 18e at pH 8. Its position is 3.75 Å from Ser-195 O^{γ} at pH 5 and 3.08 Å at pH 8. This is compatible with the presence of a disordered sulfate ion at low pH, which is lost to a large extent and/or replaced by water at high pH. Tulinsky and Wright presented evidence for this sulfate in chymotrypsin crystals at low pH.[54]

This observation indicates a change in charge of the catalytic groups and prefers a positively charged His-57 at low pH. Whichever residue is protonated at low pH, its effect is felt outside the molecule. It is conceivable that protonation of the catalytic system reduces the nucleophilicity of Ser-195 O^{γ} considerably.

Further complication is brought into the puzzle through the observation that the pH dependence of the anhydrotrypsin–PTI association rate is similar to the trypsin–PTI association rate except for a pK change of 0.6 unit.[55] Anhydrotrypsin has no Ser-195 O^{γ} and nucleophilicity of Ser-195 O^{γ} is excluded as the source of the pH dependence here. We believe that sites far from the catalytic residues should also be considered.

A second aspect, where we prefer a different view from that of Blow,[15] concerns the Ser-195 O^{γ} position in the free enzyme. Trypsin at pH 5 and pH 8 is characterized by a Ser-195 O^{γ} dihedral angle (χ^1) of −95 and −60°, respectively. As this angle is −83° in the PTI–trypsin complex (tetrahedral adduct)[21] or −64° in the acyl-enzyme,[26] little conformational change is required to proceed in the catalytic reaction steps.[18] In contrast, in α-chymotrypsin crystals at pH 4.5 a χ^1 angle of 93° has been observed.[26] Such a conformation requires a major change to reach the tetrahedral adduct.

We will now focus on the relative arrangement of the scissile peptide group and the catalytic residues in the complex; see Figure 10. The nucleophilic Ser-195 O^{γ} is approximately perpendicular to the Lys-15(I)–Ala-16(I) peptide plane above the carbonyl carbon. This is close to the minimum energy pathway for a nucleophilic addition reaction as shown by Bürgi, Dunitz, and Shefter.[56] The best line of attack would be along the tetrahedral bonding direction of the carbonyl carbon.

The peptide carbonyl carbon is tetrahedrally distorted with the carbon half way between a trigonal and a tetrahedral conformation. The distance between Ser-195 O^{γ} and Lys-15 C is 2.6 Å, longer than expected for a covalent bond but shorter than a van der Waals distance. We regard this as an intermediate state of the nucleophilic addition reaction frozen by constraints imposed by enzyme and inhibitor. Similar phenomena had been observed in small-molecule crytal structures.[56] Theoretical calculations of the reaction coordinate of

(54) A. Tulinsky and L. H. Wright, *J. Mol. Biol.*, 81, 47–56 (1973).
(55) J. P. Vincent, M. Peron-Renner, J. Pudles, and M. Lazdunski, *Biochemistry*, 13, 4205–4211 (1974).

(56) H. B. Bürgi, J. D. Dunitz, and E. Shefter, *J. Am. Chem. Soc.*, **95**, 5065–5067 (1973).

nucleophile–electrophile addition show that the energy of the system changes little in the range of 3- to 2-Å distance.[57]

The contribution of the Ser-195 O^{γ}–Lys-15(I) carbonyl carbon interaction to the association energy is small. This is experimentally established by investigations of anhydrotrypsin.[58] Anhydrotrypsin lacks the Ser-195 O^{γ}, but binds the inhibitor nearly as strongly as native trypsin.[55] In anhydrotrypsin Ser-195 has been converted to a dehydroalanine by a chemical reaction which inactivates the enzyme completely. Structure analysis of the anhydrotrypsin–PTI complex[22] indicated that the pyramidalization of the Lys-15(I) carbonyl carbon is identical with that of the native complex. This distortion is obviously brought about by the interaction of the Lys-15(I) oxygen with the oxyanion binding hole. Such distortion favors formation of the tetrahedral adduct sterically and electronically.

In general, the arrangement of the scissile peptide and the catalytic groups allows the minimum energy Ser-195 O^{γ}--C approach, i.e., close to the best line of nucleophilic attack and preserving the His-57–Ser-195 O^{γ} hydrogen bond.[18]

Complex formation is accompanied by few distortions in the inhibitor binding segment other than pyramidalization of the carbonyl carbon of Lys-15(I). The inhibitor adapts to the enzyme. It would not fit to the enzyme in its native conformation (Figure 10). There is also some rearrangement of the Ser-195 main chain of the enzyme upon complex formation. The inhibitor appears to push Ser-195 slightly out of the way. The hydrogen bond between Ser-195 O^{γ} and His-57 N^{ϵ} improves considerably in the complex where it is perfectly linear with a length of 2.7 Å, compared to the long and bent bond in free trypsin at pH 8. This has functional relevance as it enhances the nucleophilicity of O^{γ} in the complex.

Speculating about the structural events related to the kinetic steps, the formation of L might be primarily a desolvation process of enzyme and inhibitor contact surfaces. This would involve not only the few rigidly bound solvent molecules associated with the active-site residues but in particular the numerous, mobile water molecules loosely bound to polar surface groups. This process could provide the entropy contribution observed. In L, both molecules might still not be properly aligned with considerable relative mobility. Formation of C involves the slight conformational changes described, allowing optimal fitting of both molecules accompanied by a favorable enthalpy contribution.

The geometry observed in the complex (scheme, Figure 11) uniquely defines the minimum energy pathway leading to the "true" tetrahedral intermediate with a covalent Ser-195 O^{γ}–C bond. This requires slight further "down" rotation of O^{γ} and an "up" movement and further pyramidalization of C. C–O and C–N bonds are lengthened. Vibration of C along C–N leads to the acyl species with a planar C^{α}, C, O, O^{γ} configuration and the C–N bond broken (Figure 12). The hydrogen is transferred from Ser-195 O^{γ} to His-57 N^{ϵ} during C–O bond formation, in accord with the estimated pK of the groups involved.[59] Proton transfer occurs along an existing hydrogen bond and should be fast and kinetically insignificant.

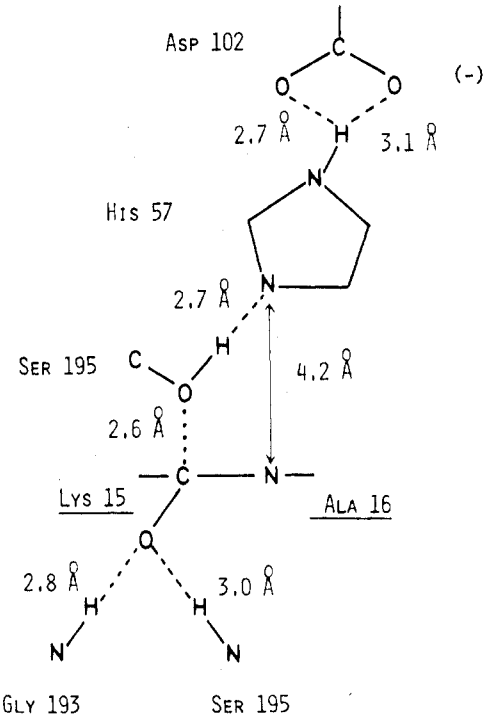


Figure 11. Scheme of the arrangement of the catalytic residues and the scissile peptide as seen in the complex.

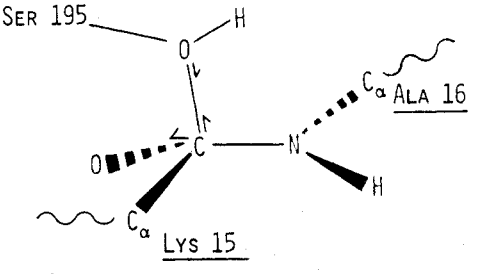


Figure 12. Vibrational modes of C of the scissile peptide leading to the O–C covalent bond and the acyl-enzyme.

Transfer of the proton from His-57 to N Ala-16 (I) of the leaving group involves a conformational change of the imidazole. It is more than 4 Å away in the complex, and it is unfavorably oriented for proton transfer (Figure 11). A conformational change of His-57 as the rate-limiting step in serine protease catalysis is postulated, in accord with experimental data[59] as well as theoretical calculations.[60] Further crystallographic evidence about the acyl-enzyme intermediate in the catalytic pathway has been described in Blow's Account.[15]

These structural features might provide an explanation for the extremely slow dissociation of the complex: Formation of I might be difficult because the His-57 conformation change discussed for acyl-enzyme formation is hindered. In the PTI–trypsin complex the imidazole is indeed tightly packed and shielded from water, in particular by the chain segment Gly-36(I) to Cys-38(I) antiparallel to the Lys-15(I) – Ala-16(I) chain.

(57) H. B. Bürgi, J. M. Lehn, and G. Wipff, *J. Am. Chem. Soc.*, **96**, 1956–1957 (1974).
(58) H. Ako, R. J. Foster, and C. A. Ryan, *Biochemistry*, **13**, 132–139 (1974).
(59) A. C. Satterthwait and W. P. Jencks, *J. Am. Chem. Soc.*, **96**, 7018–7031 (1974).
(60) S. Scheiner and W. N. Lipscomb, *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 432–436 (1976).

This chain is absent in STI, but here the bulky Tyr-62(I) and Ile-64(I) (equivalent to Cys-14(I) and Ala-16(I) in PTI) shield His-57. Dissociation to I might be difficult because the interactions in the complex have to be broken cooperatively. The rigid partner molecules allow no conformational changes required for a stepwise (zipper) breakage of the interactions.

The inhibitor has a structure almost perfectly complementary to the enzyme with minimal adaptation required. This is in contrast to a flexible substrate chain with many degrees of freedom to be frozen to make interaction with the enzyme possible.[31-33]

Resonance Raman Labels: A Submolecular Probe for Interactions in Biochemical and Biological Systems

PAUL R. CAREY* and HENRY SCHNEIDER

Division of Biological Sciences, National Research Council, Ottawa, Canada K1A 0R6

Received May 20, 1977

Key aspects of many biochemical interactions occur when a relatively small molecule is in contact with a highly specialized, local environment. Elucidation in chemical terms of these ligand–active-site interactions represents some of the crucial problems of molecular biochemistry. Resonance Raman labels[2] provide a precise indicator or reporter technique to study the nature of events produced by biological active sites on the bound ligand. The basis of the method is the use of resonance Raman labels which, while yielding vibrational and electronic spectral data, are at the same time biologically active molecules. The technique will be introduced using relatively simple protein–ligand systems where time-dependent effects are not observed, such as drug–receptor[3] and antibody–hapten[4] interactions. In these the drug and hapten are simultaneously resonance Raman reporter groups and key biological components. Extension to the time domain is illustrated by enzyme–substrate reactions.[5-7] In these studies, through the use of substrates which are resonance Raman labels the conformation and electronic structure of the substrate in the enzyme's active site can be monitored as a function of time. Extension to other important areas such as membrane processes and nucleic acid–protein interactions will be presented in outline.

A resonance Raman label provides a detailed vibrational spectrum from a specially designed chromophore when it is bound to a biochemically active site. Specificity is achieved by synthesizing a chromophore which mimics as closely as possible a true biological component. Selectivity, the ability "to see" the label to the essentially complete exclusion of everything else present, is achieved by utilizing the resonance Raman effect. Resonance Raman spectra are obtained by illuminating a sample with laser light whose wavelength lies in an absorption band of a chromophore in the sample. Pronounced intensity enhancement of certain vibrations of the chromophore results thereby. The enhancement may be 10^3- to 10^6-fold compared to normal Raman spectra where the excitation wavelength is far from absorption bands. In practice, the resonance Raman effect allows spectra to be obtained from chromophores at concentrations of 10^{-4} M or less. At these concentrations in a complex biological system the resonance Raman spectrum from a given chromophore often dominates the recorded spectrum to the extent that it alone is observed. An example of such selectivity is shown by Figure 1a. The spectrum in Figure 1a was given by a chromophoric sulfonamide drug bound to the active site of the enzyme carbonic anhydrase in an aqueous buffer. Only the sulfonamide contributes to the absorption band near 450 nm (Figure 1, inset) as the other components absorb below 300 nm. By using either of the excitation wavelengths shown in the inset, the spectrum is that of the sulfonamide bound in the enzyme's active site (Figure 1a), unobscured by enzyme of solvent features.

The information content of resonance Raman spectra follows from two principles. The peak positions are a property of the ground electronic state, while peak intensities are strongly dependent on the excited electronic states as well as on the ground state. Thus, in a manner analogous to IR spectra, resonance Raman spectra contain features which represent vibrational modes of the ground electronic state. The region covered is approximately from 200 to 2000 cm^{-1}.

Paul Carey was educated at the University of Sussex in England. After receiving his Ph.D. in 1969 he joined the National Research Council of Canada as a postdoctoral fellow. His graduate work centered on the application of NMR to chemical problems. However, soon after arriving in Ottawa he became interested in using resonance Raman spectroscopy in biology. Although his interests in molecular biophysics are fairly catholic, he spends most of his time developing the kind of work described in this Account.

Henry Schneider was born in Montreal and received a B.Sc. from Sir George Williams University, a M.Sc. from the University of Western Ontario, and a Ph.D. from McGill. After 2 years of postdoctoral study at Cornell and some time in industry, he joined the National Research Council. His interests center about biochemical mechanisms, in particular those concerning membrane process.

Both authors are Associate Research Officers.

(1) Issued as NRCC No. 16417.

(2) P. R. Carey, H. Schneider, and H. J. Bernstein, *Biochem. Biophys. Res. Commun.*, **47**, 588 (1972).

(3) (a) K. Kumar, R. W. King, and P. R. Carey, *FEBS Lett.*, **48**, 283 (1974); (b) *Biochemistry*, **15**, 2195 (1976).

(4) P. R. Carey, A. Froese, and H. Schneider, *Biochemistry*, **12**, 2198 (1973).

(5) P. R. Carey and H. Schneider, *Biochem. Biophys. Res. Commun.*, **57**, 831 (1974).

(6) P. R. Carey and H. Schneider, *J. Mol. Biol.*, **102**, 679 (1976).

(7) P. R. Carey, R. G. Carriere, K. R. Lynn, and H. Schneider, *Biochemistry*, **15**, 2387 (1976).

0001-4842/78/0111-0122$01.00/0