CATALYSIS BY TRYPSIN

Thesis by

Roger E. Koeppe II

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1976

(Submitted May 13, 1976)

ACKNOWLEDGMENTS

I thank my research advisor, Robert M. Stroud, for providing ideas, support, and direction throughout my graduate career. The educational experience which I have had in his laboratory has been most rewarding.

It is a pleasure to acknowledge the joint participation and advice of Monty Krieger and of John Chambers in some of the experiments. I thank the National Science Foundation, the National Institutes of Health, and the California Institute of Technology for financial aid during my graduate studies. I express my appreciation to my wife, Jessie, for the encouragement which she has given me.

ABSTRACT

Investigations are reported concerning the mechanism of catalysis by an enzyme of known structure, bovine trypsin. Evidence for proton transfers at the active site residues His 57 and Asp 102 during catalysis is provided by experimental techniques which specifically detect the ionization state of either the carboxyl group or the imidazole ring. The pK_a of Asp 102 in a chemically modified--but still active--form of trypsin is shown to be 6.8 by difference infrared spectroscopic titration. This assignment is facilitated by the use of inhibitory copper (Cu^{++}) ion which lowers the pK_{app} of Asp 102 by binding to trypsin between Asp 102 and His 57, as is demonstrated crystallographically.

Information concerning the solvent accessibilities and mobilities of the three imidazole side chains of bovine trypsin is provided by measurements of the pH dependences of the rates of exchange of the ring C-2 protons with tritium in labeled water. Kinetic studies of substrate hydrolysis do not detect the pK_a of His 57 anywhere in the range 3 - 8.5. Pre-incubation of trypsin at low or at high pH shows that the failure to detect the pK_a of His 57 kinetically is not due to a slow pH-dependent conformational change.

The implications of the assignment of a pK_a of 6.8 to Asp 102, and of the pK_a of His 57 being below 3, are discussed with regard to the catalytic mechanism of the serine proteases.

TABLE OF CONTENTS

CHAPTER		PAGE
	Acknowledgments	ii
	Abstract	iii
I	Introduction	1
II	On the Mechanism of Hydrolysis by Serine Proteases: Direct Determination of the pK_a's of Asp 102 and Asp 194 in Bovine Trypsin Using Difference Infrared Spectroscopy	10
III	A Crystallographic Determination of the Site at Which Inhibitory Cu^{++} Ion Binds to Trypsin	86
IV	What is the pK_a of His 57?	95
Appendix A	A Pulsed Diffusion Technique for the Growth of Protein Crystals for X-ray Diffraction	109
Appendix B	Structure-Function Relationships in the Serine Proteases	117
Appendix C	The Effect of Pre-Incubation on Trypsin Kinetics at Low pH	137
Appendix D	The pH Dependence of Tritium Exchange with the C-2 Protons of the Histidines in Bovine Trypsin	152

CHAPTER I: Introduction

Since the time when Buchner (1897) obtained the fermentation system from yeast in a cell-free extract, there has been considerable effort devoted to the study of enzymes in isolation from the living cells from which they derive. Enzymology is one of the fields which lies at an interface between the biological and the physical sciences. A study of the molecular structure and catalytic mechanism of action of a purified enzyme represents a challenging physico-chemical problem; while an understanding of the role of the enzyme in a living system, the consequences of its absence or deficiency, and the means of correcting diseases associated with such a deficiency represent aspects of the biological side of the picture.

Buchner's discovery put to rest Pasteur's argument that fermentation was inseparable from living cells. The "vital force" theory of life has long since given way to the idea that living processes are carried out by molecules composed of the ordinary chemical elements. Therefore nearly all approaches to the chemistry of life have involved studying a simplification of the living cellular system, often to the point of purifying and characterizing an individual protein, nucleic acid, lipid, or other molecule. In other cases, assemblies of molecules, such as membranes, chromosomes, or ribosomes have been studied, but here again the work is often

performed with isolated components rather than with living cells. Purified biological macromolecules can be studied in greater detail than more complex cellular systems, and represent interesting objects in their own right. In addition, the premise behind such studies is that the pieces will eventually be put together once they are individually understood, and that a greater understanding of the whole cell will result.

The goals of physical and chemical studies of a purified enzyme include a description of the three-dimensional structure of the molecule; an explanation of the structural reasons for the specificity of the enzyme toward its natural and artificial substrates and for the action of various natural and synthetic inhibitors, activators, or cofactors which affect the action of the enzyme; and an understanding of the catalytic efficiency of the enzyme, including a correlation of the molecular structure with the mechanism of action and ultimately a quantitative description of the factors responsible for the observed rate enhancement which the enzyme achieves for the reactions which it catalyzes. While no enzyme is completely understood to the point of quantitating its catalytic efficiency, many of the details of substrate specificity, inhibitor binding, and mechanism of action have been correlated with molecular structure for a small but growing number of enzymes. The subject of this thesis concerns

this general area of the correlation of enzyme structure with enzyme function, the development of new methods for conducting such experiments with enzymes, and a detailed study of a particular enzyme molecule--trypsin from bovine pancreas.

One necessity in understanding any enzyme is the three-dimensional structure of the molecule, determined to a resolution which locates all of the non-hydrogen atoms. This can be achieved by combining electron density information obtained from x-ray diffraction of crystals of the enzyme with an independent chemical determination of the linear sequence in which the various amino acids are linked to make up the protein molecule. The slowest step in these long and involved procedures is sometimes the growth of crystals of the enzyme which are suitable for x-ray diffraction. Appendix A is a reprint of a paper which describes a new pulsed diffusion method which was developed for growing crystals of glutamate-aspartate transaminase from porcine heart and of thymidylate synthetase from Lactobacillus casei. The method is useful for eliminating unwanted amorphous precipitated protein which may interfere with crystal growth, and for reducing the number of microcrystals which form and thus allowing each crystal to grow larger. The pulsed diffusion technique can also be used to crystallize proteins which are available in only limited quantities and so should prove

useful for the crystallization of newly purified enzymes.

Once the three-dimensional structure of an enzyme is known, the importance of various parts of the structure to the activity of the enzyme can be probed using chemical, kinetic, and spectroscopic methods. Such techniques can also be used in the absence of a knowledge of structure, but can be interpreted with greater confidence if the complete structure is known. For example, chemical blocking of Ser 195 of chymotrypsin by diisopropyl-fluorophosphate (Jansen et al., 1949) and of His 57 by N-tosyl-L-phenylalanyl-chloromethyl ketone (Schoellman and Shaw, 1963) demonstrated that these two residues were essential for catalytic activity. However, the proximity of Asp 102 to these two residues was unknown until the work on the three-dimensional structure (Blow et al., 1969). The structure of chymotrypsin also showed that Asp 194 was involved in stabilizing the conformation of the catalytic serine hydroxyl by means of a salt bridge to the terminal α-ammonium group (Matthews et al., 1967). The importance of the amino terminus had previously been demonstrated by chemical and spectroscopic methods (Oppenheimer et al., 1966).

Another example of the correlation of chemical and crystallographic results is the story of the substrate binding pocket of trypsin which is specific for a lysine or arginine side chain (Hofmann and

Bergmann, 1939). Early work on the sequence of trypsin assigned residue 189* as an asparagine (Walsh and Neurath, 1964). The distinction of Asp from Asn, or Glu from Gln, is one of the most difficult problems in sequencing, and in 1964 there was no reason to suspect that the side chain of 189 would have any special significance for the action of trypsin. However, the sequence was later corrected to Asp 189 (Walsh, 1970). Then Eyl and Inagami (1971) showed that Asp 189 could be chemically modified by water-soluble carbodiimide (a characteristic of Asp but not of Asn). Furthermore, the modification was inhibited by benzamidine, a competitive inhibitor of trypsin. The x-ray structure of trypsin (Stroud et al., 1971) shows that Asp 189 is located at the bottom of the substrate binding pocket of the trypsin molecule. Therefore, all of the evidence points to the fact that the binding pocket is specific for positively charged side chains (lysine and arginine among the natural amino acids) because of the presence of Asp 189. In addition, the binding of benzamidine in a salt bridge with Asp 189 in the pocket has been shown crystallographically (Krieger et al., 1974).

Interest in the catalytic mechanism of the serine proteases has centered around the active site residues Ser 195, His 57, and Asp102.

---

* The numbering system referred to for trypsin is that of the homologous sequence of chymotrypsin (Walsh and Neurath, 1964).

The experiments reported in this thesis were carried out in order to find out about proton transfers at the active site residues His 57 and Asp 102 of bovine trypsin during catalysis. The aspartate and imidazole side chains together act as a base which participates in the catalysis by accepting the hydroxyl proton of Ser 195 during the nucleophilic attack by the serine hydroxyl oxygen of the substrate (Bender and Kezdy, 1964). An identification of the individual roles of Asp 102 and His 57 in this process requires the use of techniques which can distinguish the two side chains and subdivide the coupled system. The methods used here were chosen to specifically detect the ionization state of either the carboxyl group or the imidazole ring. One such method--NMR spectroscopy of α-lytic protease containing a ^{13}C label at the C-2 position of His 57--has previously been used by Hunkapiller et al. (1973). Our probes for the system have been infrared spectroscopy to monitor the carboxyl group, and hydrogen-tritium exchange at the C-2 position of the imidazole ring.

Difference infrared spectroscopy was used to determine that the carboxyl side chain of Asp 102 titrates with a pK_a of 6.8 in a chemically modified--but still active--form of trypsin. The implications of this result for the catalytic mechanism of serine proteases will be discussed. A part of the procedure of assigning the infrared titrations involved the use of specific copper ion (Cu^{++}) binding to

Asp 102. The location of the bound inhibitory copper ion was determined crystallographically from a difference electron density map obtained from crystals of trypsin with and without copper.

His 57 was studied by measuring the pH dependence of the rate of exchange of tritium from labeled water with the C-2 proton of the imidazole ring. Although the pK_a of His 57 has not yet been determined, the tritium exchange data do provide information concerning the solvent accessibility and mobility of the side chain. Boundary conditions can be put on the assignment of a pK_a to His 57 based on data which will be presented regarding the kinetics of substrate hydrolysis by trypsin. Because of the pH dependence of the rate of hydrolysis and the assignment of a pK_a of 6.8 to Asp 102, the pK_a of His 57 cannot be anywhere in the range 3.0-8.5. A study of the effect of pre-incubation of trypsin at low or at high pH on the kinetics demonstrates that the failure to detect the pK_a of His 57 kinetically is not due to a slow pH-dependent conformational change. Therefore, His 57 in the serine proteases must have a remarkably low pK_a.

References

Bender, M. L. and Kezdy, F. J. (1964), J. Amer. Chem. Soc. 86, 3704-3714.

Blow, D. M., Birktoft, J. J., and Hartley, B. S. (1969), Nature 221, 337-340.

Buchner, E. (1897), Berichte der Deut. Chem. Gesellschaft 30, 117-124.

Eyl, A. W. Jr. and Inagami, T. (1971), J. Biol. Chem. 246, 738-746.

Hofmann, K. and Bergmann, M. (1939), J. Biol. Chem. 130, 81-86.

Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. E., and Richards, J. H. (1973), Biochemistry 12, 4732-4743.

Jansen, E. F., Nutting, M. D. F., and Balls, A. K. (1949), J. Biol. Chem. 179, 201-204.

Krieger, M., Kay, L. M., and Stroud, R. M. (1974), J. Mol. Biol. 83, 209-230.

Matthews, B. W., Sigler, P. B., Henderson, R., and Blow, D. M. (1967), Nature 214, 652-656.

Oppenheimer, H. L., Labouesse, B., and Hess, G. P. (1966), J. Biol. Chem. 241, 2720-2730.

Schoellman, G. and Shaw, E. (1963), Biochemistry 2, 252-255.

Stroud, R. M., Kay, L. M., and Dickerson, R. E. (1971), Cold Spring Harbor Symp. Quant. Biol. 36, 125-140.

Walsh, K. A. (1970), Methods in Enzymol. 19, 41-63.

Walsh, K. A. and Neurath, H. (1964), Proc. Nat. Acad. Sci. USA 52, 884-889.

CHAPTER II: On the Mechanism of Hydrolysis by Serine Proteases: Direct Determination of the pK_a's of Asp 102 and Asp 194 in Bovine Trypsin Using Difference Infrared Spectroscopy

Abstract

The pK_a of Asp 102 in trypsin is shown to be 6.8 ±0.2 by difference infrared titration. All but 2.5 of the carboxyls in bovine trypsin were first modified with semicarbazide. The modified enzyme still retains full activity toward nonspecific substrates. The remaining free carboxyls include one equivalent each of Asp 102 and Asp 194. The absorbances associated with the C$\cdots$O and C$=$O stretching modes at 1570 cm^{-1} and 1710 cm^{-1} were used to monitor the proportion of ionized or protonated carboxyl present in the enzyme as a function of pD. The pK_a of 6.8 was assigned to Asp 102 using copper ions which bind to trypsin between Asp 102 and His 57, so lowering the pK_{app} of Asp 102. The implication of this result for the ionization of the active site, and for the mechanism of serine proteases, is discussed. Asp 194 and the C-terminus are shown to titrate with an average pK_a of 2.9.

Introduction

The crystallographically-determined structures of all of the pancreatic serine proteases (chymotrypsin, elastase, trypsin) studied so far show that the catalytic site contains the hydroxymethyl group of Ser 195, the imidazole side chain of His 57, and the carboxyl of Asp 102.[1] Subtilisin of different evolutionary heritage contains the same arrangement of side chains at the catalytic site (Alden et al., 1970), thus emphasizing the seemingly universal efficacy of this particular arrangement of side chains in peptide hydrolysis by serine proteases.

Early studies showed that activity depended on a single group of pK_a between 6 and 7 for trypsin and chymotrypsin (Hammond and Gutfreund, 1955). More detailed results demonstrated that a pK_a near 7 controls both acylation and deacylation (Gutfreund and Sturtevant, 1956; Bender et al., 1962). This has usually been attributed to His 57 and, since reaction occurs best above pH 7, most mechanistic schemes have used His 57 first as a general base. One notable exception is the charge relay mechanism of Blow et al. (1969), which first gave functionality to Asp 102 as the base in the Asp 102-His 57 system. Jencks (1969), however, pointed out that the group with apparent pK_a of 7 could represent another group in the enzyme

---

[1] The numbering system referred to is that of chymotrypsin (Walsh and Neurath, 1964).

controlling conformation, or effecting a change in rate-determining step.

The assumption that His 57 was the catalytic group having a pK_a of 7 seemed to indicate a mainly structural role for the carboxyl group of Asp 102--perhaps in maintaining the orientation of the imidazole of His 57, or in stabilizing an imidazolium ion during catalysis. Although this perhaps seemed reasonable for the trypsin family, the finding that subtilisin, of completely different evolutionary origin, has exactly the same structure at the active site seems to emphasize the essential role of Asp 102 and raises the question of an even more specific catalytic function for Asp 102. In addition, it has been somewhat puzzling as to why the pK_a of 7, if it belonged to His 57, were not more perturbed from that of imidazole in aqueous solution by the negatively-charged group of Asp 102. For these reasons, we sought ways of determining the microscopic pK_a's of residues in the active center of serine proteases. It should be noted that, although many different studies show rate dependence on, or spectral change associated with a single group of pK_a near 7 (which is required in its depronated form for activity), most do not distinguish the microscopic pK_a of His 57 from that of Asp 102. Hunkapiller et al. (1973) presented strong NMR evidence that His 57 in α-lytic protease is not the group which titrates at pH 6.8, and assigned the titration to

a neighboring group which they presumed to be Asp 102.

Two problems are apparent in the kinds of experiments which might attempt this definition. First, the presence of a label on the enzyme near the active center, which could be used to sense the changes in environment there, perturbs the enzyme in an unknown fashion. Consequently the system might no longer be analogous to the native enzyme and ionizations could easily be perturbed. Second, spectroscopic methods which can define the ionization state of particular types of side chains often present a problem of assignment to the particular group in question. The method used here overcomes this latter difficulty through the use of metal ions which specifically bind to the active center and perturb the native enzyme ionization in a predictable fashion. Nevertheless, the metal label is used only for assignment and is not present during the primary observations.

Timasheff and Rupley (1972) demonstrated that infrared difference spectroscopy could be used to study the overall ionization of all of the carboxyl groups in lysozyme. Bovine trypsin contains eleven carboxyl groups (Titani et al., 1975), of which three are most easily identified as being specifically linked to substrate binding (Asp 189), conformation (Asp 194), and catalytic function (Asp 102). Trypsin (rather than any other serine protease) was chosen for this study which used infrared difference spectroscopy, because it contains

fewer carboxyl residues than many of the other extracellular serine proteases, because its structure has been well-characterized and refined at 1.76 Å resolution (J. L. Chambers and R. M. Stroud, unpublished), and because it is the most specific of the pancreatic enzymes. As a result, ionization of the active site in the many well-known trypsin-inhibitor complexes can also be characterized. Autolysis is a particular problem with trypsin, so initial studies were carried out on benzamidine-inhibited trypsin.

Materials and Methods

Materials: Bovine trypsin from Worthington was 2x crystallized, salt-free, lots 2JB and 33P685. D_2O was from Stohler. Semicarbazide hydrochloride (Matheson) was recrystallized from ethanol/water before use. 1-ethyl-3(3-dimethylaminopropyl) carbodiimide from Sigma and benzamidine hydrochloride (Aldrich) were used without further purification.

^{14}C-SCZ[2] was prepared according to the procedure of Fersht

---

[2] Abbreviations used: ^{14}C-SCZ, ^{14}C-semicarbazide hydrochloride; ^{14}C-Gly-Et, ^{14}C-glycine ethyl ester hydrochloride; DTT, dithiothreitol; BAEE, N-benzoyl-L-arginine ethyl ester; NPGB, p-nitrophenyl-p'-guanidino benzoate; Ac-Gly-NPE, N-acetyl glycine-p-nitrophenyl ester; PIPES, piperazine-N, N'-bis(2-ethane-sulfonic acid); trypsin-SCZ, trypsin in which accessible carboxyl groups have been blocked with semicarbazide by means of a carbodiimide coupling procedure; DIP-trypsin, diisopropylphosphoryl trypsin.

and Sperling (1973), starting with ^{14}C-potassium cyanate (ICN) and aqueous hydrazine at pH 7. The reaction yielded 96% (of starting OCN^-) semicarbazide based on the 1,2-naphthoquinone-4-sulfonic acid assay (Fersht and Requena, 1971a). 4.3% of the initial OCN^- was recovered as bis-urea, a side product which after recrystallization had a specific activity of 215 ±2.1 cpm/nmole. ^{14}C-Gly-Et (50μCi) from New England Nuclear was diluted with cold carrier to a specific activity of 285 ±6.8 cpm/nmole. All scintillation counting was done using "Aquasol" (New England Nuclear) in a Beckman Model LS-250 Liquid Scintillation Counter. Reference solutions of known concentrations of ^{14}C-Gly-Et and ^{14}C-bis-urea were used to correct for counting efficiency.

Schwarz-Mann ultra-pure urea was freshly dissolved to 10 M concentration and deionized on a Column of Dowex AG 501-X8D, then immediately diluted to 8 M in the reaction mixture. DTT, histamine, thyrotropin-releasing hormone (Calbiochem), ethyleneimine (Eastman), L-histidine, glycine, and glycyl-L-histidyl-glycine (Sigma) were used without further purification.

Normal solutions of "DCl", "DNO_3", and "NaOD" were prepared as described by Timasheff and Rupley (1972) by dissolving the appropriate reagent in D_2O.

Substrates used to assay trypsin activity were BAEE (Schwarz-

Mann), NPGB (Cyclo), and Ac-Gly-NPE (Cyclo, recrystallized before use). PIPES, used in buffers, was from Calbiochem. All other chemicals were of reagent grade.

Preparation of trypsin-SCZ. Although native trypsin has a relatively small number of carboxyl groups (eleven) for a protein of its size, it is useful to be able to reduce this number in order to observe the ionizations of individual groups. It can be difficult to determine individual pK_a values from a titration curve which is the sum of eleven separate overlapping curves. The problem was simplified by modifying accessible carboxyls with semicarbazide using the carbodiimide reaction of Hoare and Koshland (1967) and the procedure of Fersht and Sperling (1973). The reaction has been shown to modify surface carboxyl groups while not affecting those which are buried inside trypsin (Eyl and Inagami, 1971) or chymotrypsin (Carraway et al., 1969; Fersht and Sperling, 1973).

Special precautions were taken to stabilize the active conformation of the enzyme and to prevent autolysis during the semicarbazide labeling procedure: a) 12 mM benzamidine was present throughout; b) the time of reaction was three hours, somewhat shorter than used by Fersht and Sperling for chymotrypsin. 11.1 g of semicarbazide hydrochloride (1 M) and 190 mg benzamidine hydrochloride (12 mM)

were dissolved in 100 ml H_2O containing enough NaOH to bring the pH to 3.7. 600 mg trypsin was added. The reaction mixture was stirred for three hours at room temperature and 1.2 g portions of the carbodiimide reagent were added at time zero and again after one hour. At the end of the three-hour reaction period, the solution was extensively dialyzed against 12 mM benzamidine at 4°C and then lyophilized to yield 583 mg of white powder.

Infrared spectra. Absorbance of carboxyl groups in the infrared spectrum occurs at 1710 cm^{-1} for the C=O stretch in the protonated carboxyl groups, and at 1570 cm^{-1} for the C···O stretch in the carboxylate anion. These peaks overlap with a strong water bending mode at 1640 cm^{-1}; therefore, the water of hydration was exchanged for twenty hours with D_2O. Nearly all remaining solvent hydrogen will be present at HOD (Timasheff et al., 1973), and the bending absorbances for both D_2O (1250 cm^{-1}) and HOD (1450 cm^{-1}) are well outside the 1710 - 1570 cm^{-1} region.

The amide I band of peptide bonds also overlaps with the carboxylic acid peaks, so reference protein solutions were used to cancel out pD-independent IR absorption bands. Difference IR spectra were obtained using pairs of accurately balanced solutions of benzamidine trypsin (native and semicarbazide modified) of different pD.

Trypsin becomes rapidly less soluble above pD 8.0, as the isoelectric point is approached. Therefore, the concentration of enzyme was that of a saturated solution (native or modified enzyme) at pD 7.8 (~1.5 mM), samples of which were pH-adjusted and used throughout the entire pD range. The enzyme concentration was constant for all spectroscopy. Unfortunately, trypsin is not soluble enough to give good IR spectra above pD ~7.8.

The pD values referred to above are 0.4 unit higher than the ones actually read on a Beckman pH meter (Glasoe and Long, 1960). This effect is almost exactly compensated by the fact that the pK_a's of carboxyls are observed about 0.5 unit higher in D_2O than in H_2O (Glasoe and Long, 1960). Therefore, we are omitting both corrections in reporting our results; that is, we are using direct pH meter readings and unaltered pK_{app}'s. These values are too low for the enzyme in D_2O, but accurately reflect the situation in H_2O.

Trypsin (400 mg) or trypsin-SCZ (400 mg) was dissolved in 8 ml D_2O containing 12 mM benzamidine and left overnight at 4°C for solvent hydrogen exchange. The sample was then lyophilized, reconstituted in 8 ml D_2O containing 6 mM NaCl, or 2 mM $Cu(Cl_2)$ with 4 mM NaCl, as appropriate, adjusted to pD 7.4 (meter reading) with 1 M NaOD, and filtered through a Millipore membrane (0.45μ). The pD of the bulk solution was adjusted in stepwise fashion from 7.4 to 1.5

using 1 M DCl, and 400 μl aliquots were removed at each step change in pD. Protein concentration was determined using the extinction coefficient for trypsin at 280 nm (OD_{280} = 1.54 for 1 mg/ml, and molecular weight of 23,891 (Robinson et al., 1971)), and was always near 36 mg/ml (1.5 mM). The difference between this and 50 mg/ml expected from initial weighing alone is accounted for by the water of hydration and, to a lesser extent, by the small amount of precipitated protein which comes out of solution prior to the filtration. The 0.4 ml samples were used immediately or were frozen until used to prevent degradation.

Difference spectra were recorded on a Perkin-Elmer Model 225 Infrared Spectrophotometer, or on a Beckman Model 4240 Infrared Spectrophotometer interfaced to a Fabri-Tek Model 1062 computer, to permit averaging of repetitive scans. Samples were contained in matched sealed Perkin-Elmer cells of path length 0.125 or 0.150 mm, or in a variable path cell, all equipped with BaF_2 windows. Baseline controls were run using solutions of identical pD in the two cells. The temperature of the samples in the infrared beam was kept somewhat below room temperature by blowing chilled dry air onto the windows and by circulating ice water through copper coils which were attached to the cells.

Enzymatic Activity. The activity of trypsin-SCZ was determined for specific and nonspecific substrates, and compared to "native" (Worthington) enzyme under the same conditions. One specific assay used the difference in absorbance at 253 nm between N-benzoyl-L-arginine and its ester (Schwert and Takenaka, 1955). The enzymes were kept in fresh stock solutions in 1 mM HCl, 50 mM $CaCl_2$, pH 3.0, until aliquots were injected into the assay mixture. Assay conditions were: 4.2×10^{-8} M enzyme, 0.3 mM BAEE, 60 mM PIPES buffer, pH 7.0, 25.0°C, in a 1.0 cm cuvette. Change in A_{253} with time was recorded using a Gilford Model 240 Spectrophotometer equipped with a moving chart.

Another assay involved use of NPGB in a specific titration of active sites (Chase and Shaw, 1967). A typical assay used 2.4 ml of borate buffer (12.5 mM, pH 8.3), 0.5 ml of enzyme (5 mg/ml), and 0.1 ml of dimethylformamide containing 5 mg/ml substrate. The initial burst of p-nitrophenol release was monitored at 412 nm. The number of active enzyme molecules containing both specific binding pockets and active catalytic sites was determined by comparing the A_{412} value to a standard curve for p-nitrophenol concentration in the assay solution.

Nonspecific activity was determined by following the increase in absorbance at 412 nm due to hydrolysis of Ac-Gly-NPE. The

substrate was dissolved in ethanol, the enzyme in pH 3 HCl as above, and aliquots of each were injected into the PIPES buffer. Final conditions were: 2.1×10^{-6} M enzyme, 1 mM substrate, 60 mM PIPES, pH 7.0, 25.0°C. Correction was made for the spontaneous hydrolysis of substrate in the absence of enzyme. The effect of 0.01 M benzamidine on both native and modified enzyme activities was tested.

Identification of Free Carboxyls in Trypsin-SCZ. In order to identify those buried carboxyl groups which had not reacted in the coupling of semicarbazide to native trypsin, the protein was unfolded in 8 M urea and newly exposed carboxyls were labeled with ^{14}C-SCZ. To assure complete unfolding, the six disulfide bridges of trypsin were first reduced, and then aminoethylated. Reduction of disulfides was according to Konigsberg (1972). Trypsin-SCZ (70 mg, prepared as described earlier) was dissolved in 5 ml of 8 M urea containing 0.5 M tris-HCl, pH 8.1, and 2 mM EDTA. The sample was warmed to 50°C under N_2 and treated with a 50-fold molar excess of DTT (138 mg) for four hours. The sample was cooled to room temperature, and ethyleneimine was added in three portions at ten-minute intervals according to the procedure of Raftery and Cole (1963), each addition (140 µl) being a threefold excess over the amount of DTT used. The product was extensively dialyzed against H_2O to remove excess

reagents, and then lyophilized. Part of the sample was saved for a control thermolysin digestion (see below) and part was used for the radioactive labeling of previously unmodified carboxyls.

In order to assay for the amounts of each carboxyl modified, ^{14}C-SCZ or ^{14}C-Gly-Et was coupled to the remaining free carboxyls in reduced S-aminoethyl trypsin-SCZ by a procedure similar to those of Carraway et al. (1969) and Fersht and Sperling (1973). 23 mg of the protein was dissolved in 1.6 ml of 10 M urea. 0.3 ml of a solution of ^{14}C-SCZ (0.8 M, 215 cpm/nmole) or ^{14}C-Gly-Et (1 0 M, 285 cpm/nmole) was added and the pH adjusted to 4.2 for SCZ labeling or 4.75 for Gly labeling, with 0.5 M HCl. 38 mg of the carbodiimide reagent was added, and the pH maintained at 4.2 or at 4.75 with HCl. Additional 38 mg aliquots were added after one hour and again after two hours to ensure complete reaction with label. After four hours, the carbodiimide reaction was quenched by adding 5 ml of 1 M sodium acetate, pH 4.75, which reacts with remaining reagent. The solution was then dialyzed against H_2O at 4°C until the radioactivity in the dialysate had been reduced to the level of the background in pure water. An aliquot of labeled protein solution was saved for scintillation counting and A_{280} measurement. The remaining sample was lyophilized prior to digestion with thermolysin.

Three batches of modified S-aminoethyl trypsin-SCZ were then

available: a "cold" sample which still had free those carboxyl groups which are buried in native trypsin, and two radioactively-labeled samples, one containing ^{14}C-SCZ and the other containing ^{14}C-Gly-Et attached to the previously buried carboxyls. Ten mg of each protein sample was dissolved in 0.6 ml of 10 mM tris-acetate buffer, pH 8.0, containing 2 mM Ca^{++} and 50 μg thermolysin (Matsubara and Sasaki, 1968). Enzymatic digestion of the sample was carried out at 37°C for three hours, and an additional 0.1 ml aliquot of the thermolysin solution was added after 1.5 hours to ensure maximum digestion. The reactions were stopped by addition of one drop of 15% acetic acid, and the samples were lyophilized.

To determine which peptides contained the label, each sample of thermolytic peptides was reconstituted in 100 μl H_2O. Ten μl or 20 μl portions were spotted on 46 cm x 57 cm sheets of Whatman 3 MM paper for peptide maps at pH 6.5 (Bennett, 1967). Descending chromatography (butanol:pyridine:water:glacial acetic acid, 15:10:12:3) in one direction was followed by electrophoresis at pH 6.5 in pyridine acetate buffer in the other direction. At least three maps were run for each sample. One of the maps in each set was dipped in staining solution (300 ml of 100% ethanol, 40 ml collidine, 0.5 g ninhydrin, 100 ml glacial acetic acid). Autoradiography was used to locate spots with labeled peptides, on both stained and unstained maps from the

same sample. Radiograms were made using Kodak RP-54 medical x-ray film placed in contact with the maps for five to eighteen days.

In order to define the location of labeled residues in the trypsin sequence, peptides were eluted from the paper using 10% aqueous NH_4OH. After lyophilization each peptide sample was reconstituted in 1.00 ml H_2O containing an accurately known amount of norleucine to serve as an internal standard during amino acid analysis. 100 μl aliquots were taken for scintillation counting. The remaining portions were hydrolyzed in redistilled constant boiling HCl at 110° for twenty hours. Appropriate aliquots were used for quantitative amino acid analysis on a Beckman Model 120-C or a Durrum Model D-500 Amino Acid Analyzer.

Results

Infrared Spectra of Benzamidine Trypsin. The direct spectrum of trypsin dissolved in D_2O (pD = 4.6) is shown in Figure 1. The most notable features of the spectrum are the broad band centered at 1640 cm^{-1} due to absorbance by peptide backbone amide carbonyl groups, and the band at 1450 cm^{-1} which shows extensive deuteration of the peptide N—H groups. At the peak maximum at 1640 cm^{-1} the transmittance drops to 3.5%, or an optical density of 1.5. Thus, difference spectra will exhibit decreased sensitivity at 1640 cm^{-1} due to the

small amount of radiation passing through the reference and sample cells. Nevertheless, the situation is considerably improved at the frequencies which will be of interest in the difference spectra below, since the absorbance values are 0.6 at 1600 cm^{-1} and 0.8 at 1680 cm^{-1}. A shoulder due to side chain carboxylate absorbance can be seen at 1590 cm^{-1}, and the population of protonated carboxyl groups at pD 4.6 produces a barely detectable broadening at 1710 cm^{-1}.

Difference spectra were recorded in the range from 1800 cm^{-1} to 1500 cm^{-1} using pairs of samples with equivalent protein concentration from a series of protein solutions of varying pD. Examples of such spectra are shown in Figure 2. At 1710 cm^{-1} there is a peak representing the excess population of protonated carboxyl groups present in the low pH sample, while at 1570 cm^{-1} there appears a peak of opposite sense representing the corresponding excess of ionized carboxyls in the high pH sample. By plotting the heights of the peaks at either frequency from a series of spectra obtained using solutions of various different pD's, one can obtain a titration curve which is the sum of the titration curves of all of the carboxyl groups in the enzyme, as was done by Timasheff and Rupley (1972) for lysozyme. Because of variations in the baselines, the peak at 1710 cm^{-1} was easier to measure than the peak at 1570 cm^{-1}. Differences in baselines are attributed to small differences in protein concentrations and

in HOD content between the reference and sample cells. Measurements were made by converting the values of percent transmittance to absorbance at the baseline and at the maximum of each 1710 cm^{-1} peak (see arrows in Figure 2). The differences in absorbance were plotted on a graph in which the absorbance at pD 1.5 was arbitrarily set equal to zero (Figure 3). This graph is the sum of the titration curves of the carboxyl groups of bovine trypsin.

Rather than trying to extract and assign individual pK_a's based on the graph in Figure 3, a procedure which would reduce the number of titrating carboxyl groups in the enzyme was pursued. The characterization of the modified enzyme (trypsin-SCZ) will be presented prior to a consideration of the infrared spectra of trypsin-SCZ.

Catalytic Activity of Semicarbazide-Modified Trypsin. After blocking accessible carboxyl groups of trypsin with semicarbazide, the modified enzyme (trypsin-SCZ) was tested for activity toward specific and nonspecific substrates. The results in Table I show that 88-92% of the molecules have lost the ability to bind the specific substrates BAEE and NPGB, due to modification of Asp 189 in the binding pocket. The use of the active site titrant NPGB allows one to determine that the 8% residual activity is in fact due to molecules having normal binding pockets, and is not merely nonspecific hydrolysis by

blocked trypsin. The agreement between the NPGB (initial burst) and BAEE (overall rate) results indicate the the action of trypsin-SCZ toward specific substrates can be completely accounted for by the modification of Asp 189, and that any other carboxyl groups which have been modified have no effect on activity.

Enhanced catalytic activity is seen for trypsin-SCZ toward the nonspecific substrate Ac-Gly-NPE. This result is not surprising for trypsin, when a small molecule is bound in the specific side chain binding pocket. Inagami and York (1968) have observed similar effects upon binding of small side chain analogues such as methyl guanidine to trypsin. The activity of trypsin-SCZ toward Ac-Gly-NPE is not diminished by the presence of benzamidine (Table I), a fact which further confirms that the binding pocket has been blocked.

The blocking of Asp 189 in the trypsin binding pocket was also observed by Eyl and Inagami (1971) when they coupled glycinamide to accessible carboxyl groups using the carbodiimide procedure at pH 4.75. At that pH the modification of Asp 189 was shown to be inhibited by 0.4 M benzamidine. In our hands a much lower concentration of benzamidine (12 mM) did not prevent reaction of Asp 189 during carbodiimide coupling of semicarbazide at pH 3.9. These results are expected and consistent, since benzamidine binding has been shown to depend on a pK_a of 4.6 (East and Trowbridge, 1968). In our case both

the lower inhibitor concentration and the neutrality of Asp 189 act to reduce benzamidine binding. The size and geometry of the semicarbazide group in the binding pocket of trypsin were estimated by building the blocking group into a wire model of the active site region. Figure 4 shows the result of such model building using the trypsin coordinates of Krieger, Kay, and Stroud (1974). The minimum distance from the blocking group to the active site serine hydroxyl is 7.3 Å in the model. Therefore, no close contact with the catalytic groups or with nonspecific substrates is anticipated.

Identification of Remaining Free Carboxyl Groups in Trypsin-SCZ. A sample of trypsin-SCZ was denatured to expose free carboxyls which had escaped modification in the blocking reaction. The reduced S-aminoethylated trypsin-SCZ was dissolved in 8 M urea. The carbodiimide reaction was again used to couple ^{14}C-Gly-Et or ^{14}C-SCZ to the remaining free carboxyls. Each protein molecule incorporated 2.57 ±0.09 molecules of SCZ label, or 2.53 ±0.12 molecules of Gly label. Thus, the two labeling tests are in agreement and show that 2.5 carboxyl groups remained unblocked in trypsin-SCZ.

Thermolysin was used to digest S-aminoethylated trypsin-SCZ before and after incorporation of radioactive label. Two-dimensional peptide maps of unlabeled and ^{14}C-Gly-Et labeled samples are shown

in Figure 5. Two peptides in the unlabeled sample which run toward the positive electrode at pH 6.5 (Figure 5A) are absent from the map after ^{14}C-Gly-Et treatment (Figure 5B). Because of their amino acid composition (Table II) these peptides were identified as Leu$_{99}$-Asn-Asn-Asp$_{102}$ (major spot), and Leu$_{99}$-Asn-Asn-Asp$_{102}$-Ile-Met-Leu (minor spot), respectively. This provides a direct demonstration that Asp 102 in some of the molecules does not react with carbodiimide in native trypsin, but does react when the enzyme is denatured.

The Asp 102-containing peptide and other labeled peptides in the ^{14}C-Gly-Et map were identified by autoradiography. The radiogram in Figure 6 shows the locations of the labeled spots and the assignments made by amino acid analysis of eluted peptides (Table II). The peptide labeled "102" in Figure 6 includes the composition expected for the part of the sequence $\text{Leu}_{99}\text{-Asn-Asn-}\underset{\substack{|\\ *\text{Gly}}}{\text{Asp}}_{102}$, although contaminating peptides are also present in the same spot. With the bulky uncharged blocking group on the β-carboxyl of Asp 102 thermolysin cleavage is more efficient at the Asp$_{102}$-Ile bond; and the longer heptapeptide containing Asp 102 does not appear in the map. The "00" peptide is assigned to Ile$_{242}$-Ala-Ser-Asn$_{245}$-*Gly in the sequence. The group of peptides labeled "194" all have the same composition except that they vary in valine content. The composition is consistent

with the long peptide Leu-Glu_{186}-Gly-Gly-Lys-Asp_{189}-Ser-Cys-Gln-Gly-Asp_{194}-Ser-Gly-Gly-Pro-(Val)-(Val), which actually contains
*Gly
three acidic side chains which are potential candidates for carrying the ^{14}C-Gly-Et label. However, the studies of substrate specificity of trypsin-SCZ (see section above) exclude Asp 189 as a possibility. The three-dimensional structure of trypsin which shows Glu 186 exposed to solvent and Asp 194 buried, and the fact that Carraway et al. (1969) and Fersht and Sperling (1973) have shown that Asp 194 in chymotrypsin will not react with carbodiimide unless the enzyme is denatured, indicate that the assignment of Asp 194 as the radiolabeled group is correct. Specific activity measurements were used to exclude the possibility that both acids become fully labeled (see below).

To determine quantitatively the percentage of each carboxyl which remained unblocked in the trypsin-SCZ sample, it was necessary to use ^{14}C-SCZ for the second coupling reaction under denaturing conditions, so that labeled and unlabeled peptides would be chemically identical and therefore run to the same place on the peptide map. The specific activity of each peptide after thermolysin digestion was therefore a measure of the proportion of each particular carboxyl which remained unblocked in trypsin-SCZ. Figure 7 shows an autoradiogram and drawing of a peptide map of ^{14}C-SCZ labeled S-aminoethyl trypsin-SCZ. The C-terminus of Asn 245 did not come out in a

single peptide spot whose specific activity could be determined. Nevertheless, Asp 102 and Asp 194 were each labeled with one mole of ^{14}C per mole of protein. The compositions of the peptides, their sequence assignments and the results of the quantitative measurements are listed in Tables II, III, and IV, respectively. Asn 245 must be responsible for most of the remaining 0.5 mole of radioactivity. Thus, the modified enzyme trypsin-SCZ contains 2.5 free carboxyl groups, and these are identified as Asp 102, Asp 194, and about 50% of the molecules which have the C-terminus unblocked.

Infrared Spectra of Trypsin-SCZ. Difference spectra were recorded in the same way as for unmodified trypsin. Benzamidine was present in all solutions to prevent autolysis by the 10% of molecules which still had binding pockets containing unblocked Asp 189. Difference infrared peaks due to each of the two titrations in the pH range 1.7 - 7.5 are shown in the spectra in Figure 8. The COO^- peak in most of these spectra cannot be quantitated because these spectra were recorded without the use of a variable path length reference cell. Any small difference in path length will be reflected in the amide I band, centered at 1645 cm^{-1} (see Figure 1), and magnified by the number of total amides (~220) in the trypsin molecule. The spectra recorded in Figure 9A were made with a variable path

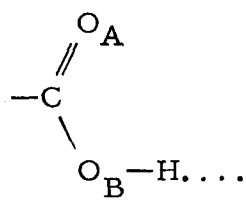
length cell where the balance is corrected experimentally.

In the low pH region, an absorption due to COOH appears at the usual frequency (1710 cm^{-1}) and disappears between pH 2 and pH 3 as the titration occurs (Figure 8, left). Although the reciprocal COO^- peak does not show up in Figure 8, it can be seen in Figure 9A at 1590 cm^{-1}. Between pH 3.5 and pH 6, the infrared spectrum of trypsin-SCZ contains no pH-dependent peaks in the region from 1800 cm^{-1} to 1500 cm^{-1} (Figure 8, center). Between pH 6 and pH 7.5, difference peaks appear at frequencies which are shifted by 30 cm^{-1} from the usual frequencies observed for carboxyl groups. A peak at 1680 cm^{-1} disappears as the pH is raised from 6 to 7.5. This peak can be seen by comparing the pairs of spectra (Figure 8, center and right) recorded using sample solutions of pD 6.6 and 7.05 with various reference solutions (pD 3.2, 3.8, 4.35, and 6.6). The spectrum which has $\Delta pD = 7.05 - 3.8$ contains an unusually small peak for two reasons: 1) part of the peak has been cut off by the uneven baseline due to amide I imbalance, and 2) the spectrum has been recorded at a relatively low position on the logarithmic percent transmittance scale.

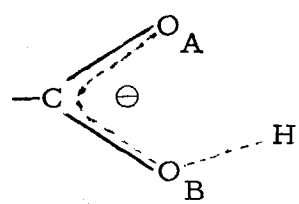
The corresponding peak at 1600 cm^{-1} can be seen along with the 1680 cm^{-1} peak in the $\Delta pD = 7.13 - 6.5$ spectrum of Figure 8, and in the $\Delta pD = 7.4 - 4.7$ spectrum of Figure 9. Since Figure 9 is on a

linear absorbance scale, the sizes of those peaks correspond to their relative intensities, a feature which is not true of the logarithmic percent transmittance scale. The COO^- absorption band is known to be several times as intense as the COOH band (Timasheff and Rupley, 1972).

Frequencies of 1680 cm^{-1} and 1600 cm^{-1} are reasonable for a carboxylic acid which is strongly hydrogen bonded. If the protonated acid were in a position to act as an unusually strong hydrogen bond donor,



then the H—O_B bond distance would be a little longer than normal, the C—O_B bond distance would be a little stronger than a single bond, and the C=O_A bond (whose stretching is normally observed at 1710 cm^{-1}) would be a little weaker than a normal double bond and would absorb radiation of a lower energy, namely 1680 cm^{-1}. By the same token, a carboxylate base which were hydrogen bonded,



would experience an uneven distribution of charge on the two oxygens and a nonequivalence of bonds C—O_A and C—O_B. In this case, the symmetric stretch would be expected at higher energy than the normal 1570–1590 cm^{-1} range, and the asymmetric stretch at lower energy than the normal 1410 cm^{-1}. Frequency shifts of 30 - 40 cm^{-1} from the usual positions are common for carbonyl groups which are hydrogen bonded (Susi, 1972).

The data from a number of difference spectra of solutions of trypsin-SCZ in D_2O at various pD's from 1.7 to 7.3 were combined to obtain the titration curve shown in Figure 10. The percent transmittance values at the maxima and the baselines of the 1710 cm^{-1} and 1680 cm^{-1} peaks were measured (arrows in Figure 8). After conversion to absorbance units, the data from the two peaks were plotted on the same scale. Two titrations can be seen, and both are steeper than would be expected for single or noninteracting groups. Nevertheless, the slope of the titration having its midpoint at pH 2.9 is 1.5 times the slope of the titration at high pH. The curve cannot be extended beyond pH 7.5 because of the decreased solubility of trypsin at high pH, so one cannot see the endpoint of the second titration. However, the ratio of the slopes of the two titrations, 1:1.5, and the total of 2.5 carboxyl groups in the trypsin-SCZ molecules strongly suggest that the low pH titration represents 1.5

carboxyl groups, while the titration at high pH is due to a single carboxyl group. In addition, the observation of an absorption at an unusual frequency (1680 cm^{-1}) can most reasonably be assigned to a single group in an unusually strong hydrogen-bonded environment. If the low pH titration is assumed to represent the titration of 1.5 groups, the expected endpoint (and therefore the midpoint) of the pH 6.8 titration can be estimated. The midpoint occurs at pH 6.8.

The average intensity of the absorption by one carboxyl group in trypsin can be determined from Figure 3 using the total of eleven carboxyl groups (ten of which absorb at 1710 cm^{-1}) in the entire molecule (Titani et al., 1975). When this average intensity is corrected for the difference in path length and compared with Figure 10, the low-pH limb of the titration curve is found to correspond to 2.0 carboxyl groups of "average" intensity. This value of 2.0 does not permit a choice between the alternatives of assigning 1.5 or 2.5 groups to the low-pH limb. In any case the extinction coefficient for a particular carboxyl group is expected to depend on its environment and should be quite variable (see control studies and Figure 12, below). Thus a comparison of extinction coefficients is not meaningful. The most straightforward assignment is that 1.5 carboxyl groups titrate with an average pK_a of 2.9, and one carboxyl titrates with a pK_a of 6.8.

Infrared Spectra of Trypsin-SCZ in the Presence of Copper (Cu^{++}) Ions. The use of copper ions as a specific probe permits one to assign the pK_a of 6.8 to Asp 102 rather than to either Asp 194 or Asn 245. In the presence of 2 mM or 6 mM $CuCl_2$ the difference infrared spectrum of 1.5 mM trypsin-SCZ no longer shows peaks between 6 and 7.5 (Figure 11, center), but instead the peaks at 1680 cm^{-1} and 1600 cm^{-1} appear between pH 4 and 5.5 (Figure 11, bottom, and Figure 9B). The titration of pK_a 6.8 is shifted to lower pH due to competition between Cu^{++} ions and protons. The observed pK_a depends on the binding constant for copper and the concentration of copper, as well as upon the intrinsic pK_a in the absence of copper. If the system is described by the following equilibria:

$$E^- + Cu^{++} = (E^-)(Cu^{++}), \quad pK_I = 4.0 \text{ (Berezin et al., 1967)}$$

$$E^- + H^+ = E - H, \quad pK_a = 6.8$$

then the observed pK_a will be the pH at which the ratio

$$\frac{E - H}{(E^-) + (E^-)(Cu^{++})}$$

is equal to 0.5 $\{K_a \text{ (apparent)} = K_a \text{ (true)}(1 + [\frac{Cu^{++}}{K_I}])\}$. This apparent pK_a will be 5.5 for the above ionization constants with 2 mM

Cu^{++}. This formulation assumes that Cu^{++} does not bind to the E-H form, since Cu^{++} binding has been shown to depend upon a group of pK_a of 6.8 (Martinek et al., 1969).

The location of the Cu^{++} binding site is derived from crystallographic studies of silver (Ag^{+}) ion binding to DIP-trypsin (Chambers et al., 1974) or to trypsinogen, and from the fact that Cu^{++} and Ag^{+} are inhibitors of trypsin and chymotrypsin which compete with each other (Martinek et al., 1969). An Ag^{+} ion binds between His 57 and Asp 102 in two crystalline forms of DIP-trypsin and in trypsinogen. Cu^{++} competes kinetically with Ag^{+} and therefore presumably binds in the same site. We are currently pursuing a crystallographic study of benzamidine trypsin containing bound Cu^{++} in order to unequivocally determine the location of copper binding.

Infrared Spectra of Control Compounds. Since 1680 cm^{-1} and 1600 cm^{-1} are not the most common frequencies for carboxyl group absorption, one may question whether an imidazole side chain could give rise to these bands in the infrared spectra. Being an aromatic ring, imidazole does indeed absorb between 1600 and 1700 cm^{-1}, although the extinction coefficient for this absorption is much lower than for carbonyl absorption. Therefore one would expect any imidazole bands to be difficult to detect in an enzyme solution which

must necessarily be dilute. Furthermore, it is not obvious that imidazole titration should shift the absorption and produce bands in a difference spectrum. Therefore, several compounds containing imidazole were used to test for the presence of any histidine-dependent titrations which would be seen in this region of the infrared spectrum.

Neither imidazole nor histamine at 20 mM concentration produce any peaks in a difference spectrum over the pH range 4.5-10.5 between the frequencies 1800 cm^{-1} and 1500 cm^{-1} (Figure 12-I and 12-II). This observation rules out the possibility that infrared absorption by a titrating imidazole could give rise to the peaks at 1680 cm^{-1} and 1600 cm^{-1} in difference spectra of trypsin-SCZ. Therefore, a carbonyl group of trypsin-SCZ is the only species which could give peaks between 1600 and 1700 cm^{-1} which are strong enough to be seen with the concentration of enzyme being used.

The question of whether a carbonyl absorption can be perturbed by titration of a neighboring group is a more interesting one. This phenomenon was studied using glycine and/or histidine in D_2O solutions (Figure 12-III and 12-IV). For the case of glycine, titration of the amino group does effect the position and intensity of the glycine carboxylate infrared absorption. This is presumably an electron-withdrawing effect of the ammonium group, on the carboxyl of

glycine. With histidine the same effect is seen when the amino group titrates. In addition, there is a further perturbation of the α-carboxyl due to titration of the imidazole, although this situation does not occur in the protein where there is no α-carboxyl.

Free histidine in solution is not a good model for histidine in a protein because the amino and carboxyl groups are not involved in amide bonds. The tripeptides Gly-His-Gly and pyro-Glu-His-Pro-amide (thyrotropin-releasing hormone) provide examples of histidine in peptides and enable one to test whether imidazole titration can affect neighboring peptide carbonyl groups. Gly-His-Gly contains two peptide carbonyl groups which absorb at 1655 cm^{-1} and the carboxyl terminal which absorbs at 1585 cm^{-1} at pD 4.5 in a direct spectrum with D_2O in the reference cell (Figure 12A-V). Difference spectra show that titration of the imidazole of Gly-His-Gly changes the intensity of the amide carbonyl absorption band. The ionization state of neither the amino terminal nor the imidazole ring has any effect on the infrared absorption of the carboxyl terminal of Gly-His-Gly (Figure 12A-V). The intensity changes which are seen in the control compounds such as glycine and Gly-His-Gly indicate that absolute extinction coefficients are meaningless for carboxyl or carbonyl groups in unusual environments such as proteins.

The infrared peaks due to the amide carbonyl groups of pyro-Glu-His-Pro-amide are also affected by imidazole titration, but only to a very small extent (Figure 9C and Figure 12-VI). This tripeptide has both amino and carboxyl termini blocked so that imidazole is the only ionizable species in the molecule. Figures 9C and 12-VI show that there are peaks in the difference infrared spectrum of this hormone between pH 4.5 and 10.5. The intensities of the peaks are quite low, and the spectrum is biphasic--unlike a simple difference spectrum containing one positive and one negative peak. The histidine must be influencing at least two of the four carbonyls in the molecule, and affecting them in different ways. The peaks are of such low intensity that they would almost not be seen at 1.5 m<u>M</u> concentration (Figure 12B). The two larger peaks in Figure 12A-VI have intensities of less than 5% and 22% of the peaks seen for trypsin-SCZ when normalized to equivalent concentration. This low intensity would seem to rule out a consideration of histidine perturbation of a neighboring carbonyl being responsible for the difference infrared spectrum of trypsin-SCZ, unless by coincidence the effects of the three histidines of trypsin happened to be additive. That such effects should be additive seems less likely than that they might tend to partially cancel each other, since a biphasic spectrum is observed for the effect of a single histidine on the several carbonyls of

thyrotropin-releasing hormone.

The most telling control experiment which shows that titrations of the three histidines of trypsin-SCZ could not be acting in concert to give the difference infrared peaks at 1680 cm^{-1} and 1600 cm^{-1} is the enzyme itself. In the presence of 2 mM Cu^{++} ions the peaks disappear from the pH 6-7 region and reappear at lower pH. Two of the histidines (His 40 and His 91) titrate with apparently normal pK_a's between 6 and 7 (Markley and Porubcan, 1975; Krieger et al., 1976). Yet no difference peaks are seen between pH 6 and 7 for 2 mM Cu^{++} with 1.5 mM trypsin-SCZ. The mole ratio of copper to enzyme is 1.3:1, so since Cu^{++} binds between Asp 102 and His 57 with $K_I = 10^{-4}$ M, there is no possibility of copper binding stoichiometrically to the other histidines, although some nonspecific copper binding has been observed (Berezin et al., 1967). Thus, the infrared difference peaks seen at 1680 and 1600 cm^{-1} for trypsin-SCZ must be due to a single group on the enzyme which binds copper. The intensities of the peaks require that they be due to a protonated and an ionized carboxylic acid, respectively, and the frequencies of the peaks are the ones expected for a strongly hydrogen-bonded carboxylic acid such as Asp 102.

Discussion

These experiments combine the use of difference infrared spectroscopy, chemical modification, and specific inhibitor binding to monitor the ionization state of a single carboxyl group at the active site of trypsin as a function of pH. That difference infrared spectroscopy could be used to follow the titrations of carboxyl groups in an enzyme has previously been shown by Timasheff and Rupley (1972). The usefulness of this spectroscopic technique in general is limited by the number of carboxyl groups in the enzyme, and by the effectiveness of curve-fitting methods that can be used to deduce individual pK_a's from an overall titration curve which represents the sum of all the carboxyl groups in the molecule. The method described in this paper overcomes these problems through chemical blocking of most of the carboxyl groups of trypsin, and uses a specific copper ion inhibitor to assign a particular pK_a, 6.8, to a particular carboxylic acid side chain, Asp 102.

The carbodiimide coupling of semicarbazide to accessible surface carboxyl groups reduced the complexity of the titration curve obtained from the difference spectra of trypsin. The trypsin-SCZ molecule had only 2.5 free carboxyl groups, 1.5 of which titrated with an average apparent pK_a of 2.9, and one which had a pK_a of 6.8.

The pK_a values of 2.9 and 6.8 were far enough apart to be easily resolved, and a two-step titration curve resulted. The pK_a's of the 1.5 groups (Asp 194 and the carboxyl-terminal of Asn 245) contributing to the low pH portion of the curve were not resolved and thus individual assignments cannot be made for these groups, although both titrate close to pH 2.9. The slopes of both limbs of the titration curve of trypsin-SCZ are steeper than would be expected for single or noninteracting groups. The reason for the steep slopes could be chemical, such as a pH-dependent conformation change of the enzyme, or cooperativity of more than one side chain ionizing over the same pH range. Alternatively, the explanation could involve difficulties in obtaining the intensities of the infrared peaks--such as the use of peak heights rather than areas; or the insensitivity of the instruments to the very small peaks which should be present between pH 3.5 and pH 6, but which were not observed.

The assignments of the two limbs of the titration curve to the carboxyls--Asp 102, and to Asp 194 and Asn 245--of trypsin-SCZ were made possible by the use of the spectra recorded with Cu^{++} present. The consequences of Asp 102 having a pK_a of 6.8 will be discussed below. That assignment leaves Asp 194 and the C-terminus with an average pK_a of 2.9, the individual values not resolved but probably neither one too far away from 2.9. A pK_a near 2.9 is

in agreement with previous suggestions of a pK_a of 3 for Asp 194 based on observed changes in optical rotation, UV-absorption, and fluorescence which accompany a pH-dependent conformation change in δ-chymotrypsin having a pK_a of 3 (Garel et al., 1974). The characteristics of the conformational change are similar to the ones observed for the alkaline pH transition controlled by the deprotonation of the amino terminal (Fersht and Requena, 1971b). Since Asp 194 is involved in an essential buried salt bridge with the amino terminal of the enzyme (Sigler et al., 1968), the two conformational changes might be expected to be similar. In addition, a lower than normal pK_a would be expected for a carboxyl group in a buried salt bridge due to the added stability derived from the overall neutrality of an ion pair. Thus, a pK_a for Asp 194 in the region near 2.9 is very reasonable. pH 2 - 3 is the normal range for a pK_a of an α-carboxyl group, such as the C-terminal of Asn 245, so again the assignment is reasonable. Furthermore, this carboxyl group exists in a solvent-accessible salt bridge with Lys 107 and Lys 87, which presumably also contributes to the diminished activity of this carboxyl to the blocking reagent.

The Ionization of the Catalytic Site. This experiment provides a direct determination of the pK_a of Asp 102 in trypsin-SCZ. The

modified enzyme is still fully active toward nonspecific substrates indicating that the active sites are affected to an insignificant extent by the labeling procedure. Furthermore, if either Asp 102 or Asp 194 had been labeled to any small extent, the labeled groups would not contribute anywhere to the infrared absorption spectra, which would respond only to the titration of the total number of free carboxyls in the enzyme solution. Therefore, the pK_a of Asp 102 in native (unmodified) trypsin must be about 6.8, and this pK_a, rather than that of His 57, is responsible for the low limb of the pH-activity profile. It should be noted that this assignment is totally consistent with the wealth of kinetic evidence on serine proteases which shows that an apparent pK_a of 6.8 controls both acylation and deacylation.

The enthalpy of ionization of the rate-controlling group of pK_a 6.8 has been determined and is 5.0 kcal/mole for acylation and 4.1 kcal/mole for deacylation of α-chymotrypsin (Rajender et al., 1971). These values are intermediate between the expected enthalpies of ionization in water solution for an imidazole ring (7 kcal/mole) and a carboxylic acid (1 kcal/mole; Greenstein and Winitz, 1961). The enthalpy of ionization of Asp 102 in trypsin is therefore 3-4 kcal/mole higher than that of a free carboxyl group in solution. A similar value of $\Delta H_{ioniz.} = 3.5$ kcal/mole has been reported for Asp 52 of lysozyme (Parsons and Raftery, 1972). Such unusual thermodynamic

parameters for carboxylic acids illustrate the ability of the folded protein to sequester specific groups away from aqueous solution in hydrophobic pockets.

The pK_a of His 57 in trypsin is still unknown, although there is evidence that the imidazole ring of His 57 in a bacterial serine protease remains neutral between pH 8 and pH 4. This derives from the NMR experiments of Hunkapiller et al. (1973), who incorporated histidine labeled with ^{13}C at the C_2 position into α-lytic protease (a bacterial homologue of the trypsin family). The C—H coupling constant in this enzyme resembled that of the neutral imidazole species down to about pH 4.0 even at 34°C. At pH 3.3, the C—H coupling constant and the line width behaved as though the imidazole were partially ionizing and partially ejected into solution at 34°C. This may represent partial or local denaturation of the protein. Therefore, the true pK_a of His 57 in the enzyme may in fact be lower than 3.3. In any case, their results provide strong evidence that His 57 in α-lytic protease, and by inference in the other serine proteases, has an abnormally low pK_a.

Other evidence for an abnormally low pK_a for His 57 comes from kinetic measurements which show that the pH rate profile can be satisfactorily explained by a single pK_a of 6.8 between pH 8 and pH 2-3. The rate of deacylation of acetyl-L-tryptophan-chymotrypsin

(Kezdy et al., 1964) and the rate of hydrolysis of acetyl-L-phenylalanine-p-nitrophenyl ester by δ-chymotrypsin (Fersht and Renard, 1974) both depend on a single pK_a of ~6.8 from pH 8 down to pH 2. The deacylation rate during trypsin-catalyzed hydrolysis has similarly been shown to depend only upon a pK_a of ~6.8 in the pH range 3-7 for the specific substrates N-benzoyl-L-arginine ethyl ester (Stewart and Dobson, 1965) and N-α-carbobenzoxy-L-lysine-p-nitrophenyl ester (Koeppe et al., 1976), and for the nonspecific substrate N-α-carbobenzoxy-L-tyrosine-p-nitrophenyl ester (Bender et al., 1964). If His 57 became protonated above pH 2 in chymotrypsin or pH 3 in trypsin, the slope of the line in a plot of log rate vs. pH (Dixon, 1953) should change from one to two with a small region of curvature at the pK_a value. Instead, data in all of the above experiments fall on single straight lines of unit slope over the pH ranges indicated. Therefore, His 57 must have a pK_a below 2 in chymotrypsin and below 3 in trypsin. The proportion of protonated imidazolium ion at His 57 is expected to be extremely low at neutral pH.

One puzzling feature of this system is that the pK_a of His 57 has to be less than 2.0. While there is precedent for a much raised pK_a of a carboxyl group in lysozyme (Parsons and Raftery, 1970; Timasheff and Rupley, 1972), there is as yet no other case of an imidazole in a protein having a pK_a as low as 2.0. The presence of

an adjacent carboxyl of high pK_a would be expected to stabilize the neutral form of imidazole somewhat, and so lower its pK_a whenever the carboxyl pK_a exceeds that of the imidazole. One can only conclude that it is exceedingly difficult to protonate the solvent-inaccessible δ-nitrogen of His 57, and that the high stability of the active center structure simply prevents this from occurring above pH 2.0.

Since the pK_a of Asp 102 in trypsin has been determined, the ionization of the active site residues at pH 6.8 is as described in Figure 13. It would previously also have been reasonable to assume that the pK_a of His 57 was about 7.0 while that of Asp 102 was fairly low. In either case the ionization at pH 6.8 leads to the same state for the active form of the enzyme above pH 6.8—namely a charged carboxyl and neutral imidazole. Below pH 6.8 the system as a whole takes up a single proton, and this is relayed via the imidazole to Asp 102. In the process the ε-nitrogen of His 57 receives a proton from solvent and the carboxyl of Asp 102 receives the proton previously on the δ-nitrogen of His 57. Below pH 6.8 the active center contains neutral aspartic acid and neutral imidazole (Hunkapiller et al., 1973), as opposed to an ion pair which would be present in the previously more accepted scheme. The proportion of imidazolium ion at pH 7.0 is probably less than 10^{-4}. Both the reverse separation of the pK_a's and the interaction between these groups in trypsin

above pH 7.0, which shows a symmetrical interaction between Asp 102 and His 57 (Krieger et al., 1974; Blow et al., 1974; Huber et al., 1974) are unlike the situation expected for solvated groups in aqueous solution, but demonstrate the importance of the protein folding in creating a special environment for its active site.

Robillard and Shulman (1974) have shown that a single proton can be seen in the PMR spectrum of chymotrypsin. This proton, easily identified by its downfield position (13-18 ppm downfield of H_2O), was ascribed to the proton on the δ-nitrogen of His 57. The change in chemical shift as a function of pH was ascribed to the titration of His 57 at about pH 7.0, and this conclusion is at variance with the result we report here. If indeed that is the proton in the active site, we would have to conclude that the change in chemical shift going down through pH 7.0 was caused by the proton moving over from histidine to the carboxyl of Asp 102. Yet, the chemical shift is apparently in the wrong direction for this interpretation. Another way out of this dilemma would occur if the proton seen in the low-field region could possibly be some other proton which is hydrogen bonded to Asp 102, and is in a particularly strong hydrogen bond. In any case, with our interpretation the titration seen in chemical shift presumably reflects the ionization state of Asp 102 rather than His 57.

The Mechanism of Catalysis by Serine Proteases. The ionization of the active center at pH 6.8 has important consequences for the mechanism of catalysis by serine proteases. Some of the implications of the assignment of the pK_a of 6.8 to Asp 102 have been presented in other discussions of the catalytic mechanism (Hunkapiller et al., 1973; and Stroud et al., 1975; Appendix B).

The new assignment scheme means that Asp 102 rather than His 57 is the base of pK_a ~7.0 responsible for the low limb of the pH activity profile found in all these enzymes, including subtilisin. During formation of the tetrahedral intermediate, a proton is removed from the hydroxyl group of serine 195 and, in the process, the hydroxyl proton may first be transferred to the ϵ-nitrogen of His 57, as first proposed by Cunningham (1957). The expectation then is that the proton in the hydrogen bond to Asp 102 and attached to the δ-nitrogen of His 57 will be in fast exchange and consequently will be immediately transferred to Asp 102. This gives obvious evolutionary significance to Asp 102 in its catalytic role as the base, but it does not immediately show why this situation should be mechanistically more favorable than the alternate system in which imidazole itself acts as a general base-general acid catalyst. Such a situation might not require an aspartate at all.

The most obvious hypothesis is that subsequent proton transfer

from His 57 to Asp 102 during catalysis would eliminate an unfavorable charge separation during the proton transfer (Jencks, 1971; Hunkapiller et al., 1973). It is clear that if the pK_a of His 57 were greater than that of Asp 102, as formerly seemed most reasonable, then the active site would contain a negatively-charged carboxyl group, a positively-charged imidazolium ion, and a negative charge on the substrate carbonyl oxygen in the tetrahedral intermediate. It now seems that as negative charge builds up on the carbonyl oxygen, the negative charge will be neutralized on Asp 102 by proton transfer from imidazole, and that imidazole itself will remain neutral.

This mechanism requires that two proton transfers be associated with the formation of the tetrahedral intermediate; namely the Ser 195 O—H proton moving to N—ϵ of His 57, and the His 57 N—δ proton moving to the carboxyl of Asp 102. These transfers could take place in either a stepwise or a concerted manner. A concerted scheme would seem to involve intermediates of lower energy, for then at no stage would there be a significant amount of positive charge anywhere in the system. The usual entropic disadvantage of concerted events would in this case be overcome by the precise alignment of the residues on the enzyme. Such concerted proton transfer in solution would require a simultaneous four-body encounter (substrate, Ser, His, and Asp), and therefore could not occur.

By bringing all of the reactants together in the proper orientation, the enzyme can provide access to an energetically favorable reaction pathway which would otherwise be impossible due to an enormous entropy of activation.

A small but apparently primary deuterium isotope effect, k_H/k_D in the range of 2-3, has been observed for both the acylation and deacylation steps during hydrolysis by chymotrypsin (Bender and Hamilton, 1962; Pollock et al., 1973). Based on a linear fall-off in rate recorded with an increasing atom fraction of deuterium in mixtures of water and deuterium oxide, Pollock et al. (1973) concluded that only one proton could be involved in the rate-determining step of the deacylation of α-chymotrypsin. However, Kresge (1973) pointed out that such data do not exclude multiple-proton transfer models, because the contributions of secondary isotope effects to the rates obtained using mixed isotopic solvents are unknown. Therefore, the isotope effects so far reported for hydrolysis by serine proteases are consistent with either stepwise or concerted transfers of the two protons.

As was pointed out by Hunkapiller et al. (1973), the scheme outlined above gives another role to Asp 102 besides the purely structural one of maintaining the orientation of His 57. Namely, the negatively-charged carboxyl rather than imidazole is the base of pK_a

6.8 seen in the catalysis. The imidazole ring shields Asp 102 from solvent, and thus raises the pK_a of the carboxyl above that of imidazole itself. It also facilitates proton transfer along the hydrogen bond, and so smoothes the transition state energy barriers between intermediates as well as reducing the internal energy of the tetrahedral intermediate states.

The binding studies of Johnson and Knowles (1966) show that the active form of chymotrypsin assists in repelling the final product from the enzyme surface. Their conclusion that a net of two negative charges in the product complex is in most part responsible for this effect is unaffected by the assignment of a pK_a of 6.8 to the carboxyl of Asp 102. One of the charges is due to the carboxylic acid of the second product of hydrolysis, and the other is due to the side chain of Asp 102 in the active form of trypsin or chymotrypsin (Figure 13).

A direct catalytic function for Asp 102 was first proposed by Blow et al. (1969). Our results are consistent with their suggestion that Asp 102 and His 57 could act as a proton relay and together increase the nucleophilicity of the serine 195 hydroxyl.

One might wonder whether the protonated or the deprotonated form of the serine hydroxyl were the attacking nucleophile. Unless the pK_a of Ser 195 were lowered by a large amount (for which there is as yet no evidence), the serine hydroxyl would be deprotonated for no

more than about 10^{-7} of the time (assuming a pK_a of 14 for serine) at neutral pH. Bender and Kezdy (1964) have estimated that the rate of acylation would in fact be several orders of magnitude higher than observed if the alkoxide ion were the attacking nucleophile. Furthermore, the nucleophile of deacylation has been replaced by isonitrosoacetone (Green and Nicholls, 1959), and by glycine- and phenylaceto-hydroxamic acids (Cohen and Erlanger, 1960). Whereas the pK_a of water (the normal nucleophile) is well above the range where chymotrypsin is active, the pK_a's of the other reagents are not. In these cases the pH-activity profile shows a dependence on the pK_a of the nucleophile, showing that it is required in its protonated form. By the principle of microscopic reversibility, the serine hydroxyl should likewise be required in its protonated form for activity (Bender and Kezdy, 1964).

The mechanism discussed here has interesting consequences for the activity of methylated chymotrypsin. Henderson (1971) has shown that chymotrypsin methylated at the ϵ-nitrogen of His 57 has about 1/5000 of the chymotrypsin activity, and that the pH profile shows again that activity depends on a group on the enzyme of pK_a ~7.0. As we have shown here that the pK_a of 6.8 normally belongs to Asp 102, one wonders what the group of pK_a 7.0 could be in the methyl enzyme where the ϵ-nitrogen is no longer available to

transfer the proton. Henderson (1971) proposed that residual activity could be retained if the imidazole were to swing out into the solvent, or even to reverse its orientation in the active site, leaving the δ-nitrogen free to participate in catalysis. Either way, the pK_a of Asp 102 can no longer be relayed to the active site via histidine because the hydrogen bonds are necessarily broken. However, the pK_a of 7.0 is not unreasonable for methyl imidazole swung out into aqueous solution. Thus, if the methyl imidazole of His 57 acts alone as the base in catalysis, this gives a measure of the upper limit to the rate enhancement due to Asp 102 in the normal catalytic mechanism. The actual effect of Asp 102 in contributing to the reaction rate may in fact be much less, since orientation effects must also contribute adversely to the activity of methyl chymotrypsin. Thus it seems that Asp 102 cannot contribute more than a factor of 5000 to the catalytic rate—maybe even much less. While methyl chymotrypsin is still a good catalyst, 5000 is nonetheless a large factor when one realizes that it would take 100 days to digest a meal with the methyl enzyme.

Acknowledgements

We wish to thank Michael A. Raftery and John A. Rupley for helpful discussions and advice, Harry B. Gray and John H. Richards for allowing us to use their infrared spectrophotometers, and Leroy E. Hood for the use of his facilities for peptide mapping and amino acid analysis.

References

Alden, R.A., Wright, C.S., and Kraut, J. (1970), Phil. Trans. Roy. Soc. Lond. B257, 119-124.

Bender, M. L. and Hamilton, G.A. (1962), J. Amer. Chem. Soc. 84, 2570-2576.

Bender, M. L. and Kezdy, F.J. (1964), J. Amer. Chem. Soc. 86, 3704-3714.

Bender, M. L., Killheffer, J.V. Jr., and Kezdy, F.J. (1964), J. Amer. Chem. Soc. 86, 5330-5331.

Bender, M. L., Schonbaum, G.R., and Zerner, B. (1962), J. Amer. Chem. Soc. 84, 2562-2570.

Bennett, J.C. (1967), Methods in Enzymol. 11, 330-339.

Berezin, I.V., Vill', Kh., Martinek, K., and Yatsimirski, A.K. (1967), Mol. Bio. (USSR) 1, 719-728.

Blow, D.M., Birktoft, J.J., and Hartley, B.S. (1969), Nature 221, 337-340.

Blow, D M., Janin, J., and Sweet, R.M. (1974), Nature 249, 54-57.

Carraway, K.L., Spoerl, P., and Koshland, D.E. Jr. (1969), J. Mol. Biol. 42, 133-137.

Chambers, J.L., Christoph, G.G., Krieger, M., Kay, L., and Stroud, R.M. (1974), Bioch. Biophys. Res. Commun. 59, 70-74.

Chase, T. Jr., and Shaw, E. (1967), Bioch. Biophys. Res. Commun. 29, 508-514.

Cohen, W. and Erlanger, B.F. (1960), J. Amer. Chem. Soc. 82, 3928-3934.

Cunningham, L. W. (1957), Science 125, 1145-1146.

Dixon, M. (1953), Biochem. J. 55, 161-170.

East, E. J. and Trowbridge, C. G. (1968), Arch. Biochem. Biophys. 125, 334-343.

Eyl, A. W. Jr. and Inagami, T. (1971), J. Biol. Chem. 246, 738-746.

Fersht, A. R. and Renard, M. (1974), Biochemistry 13, 1416-1426.

Fersht, A. R. and Requena, Y. (1971a), J. Amer. Chem. Soc. 93, 3499-3504.

Fersht, A. R. and Requena, Y. (1971b), J. Mol. Biol. 60, 279-290.

Fersht, A. R. and Sperling, J. (1973), J. Mol. Biol. 74, 137-149.

Garel, J. R., Epely, S., and Labouesse, B. (1974), Biochemistry 13, 3117-3123.

Glasoe, P. K. and Long, F. A. (1960), J. Phys. Chem. 64, 188-190.

Green, A. L. and Nicholls, J. D. (1959), Biochem. J. 72, 70-75.

Greenstein, J. P. and Winitz, M., Chemistry of the Amino Acids, Vol. 1, pp. 454, 492. John Wiley and Sons, New York, 1961.

Gutfreund, H. and Sturtevant, J. M. (1956), Biochem. J. 63, 656-661.

Hammond, B.R. and Gutfreund, H. (1955), Biochem. J. 61, 187-189.

Henderson, R. (1971), Biochem. J. 124, 13-18.

Hoare, D. G. and Koshland, D. E. Jr. (1967), J. Biol. Chem. 242, 2447-2453.

Huber, R., Kukla, D., Steigemann, W., Deisenhofer, J., and Jones, A. In Bayer Symposium V: Proteinase Inhibitors (Eds., H. Fritz, H. Tschesche, L. Greene, and E. Truscheit), pp. 497-512. Springer-Verlag, Heidelberg, 1974.

Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. E., and Richards, J. H. (1973), Biochemistry 12, 4732-4743.

Inagami, T. and York, S. S. (1968), Biochemistry 7, 4045-4052.

Jencks, W. P. (1969), Catalysis in Chemistry and Enzymology, McGraw-Hill, New York, p. 218.

Jencks, W. P. (1971), Cold Spring Harbor Symposium on Quantitative Biology 36, 1-11.

Johnson, C. H. and Knowles, J. R. (1966), Biochem. J. 101, 56-62.

Johnson, C. K. (1965), ORTEP-ORNAL-3794, Oak Ridge National Laboratories, Nashville, Tennessee.

Kezdy, F. J., Clement, G. E., and Bender, M. L. (1964), J. Amer. Chem. Soc. 86, 3690-3696.

Koeppe, R. E. II, Krieger, M., and Stroud, R. M. (1976), to be submitted to Biochem. Biophys. Res. Commun.

Konigsberg, W. (1972), Methods in Enzymol. 25, 185-188.

Kresge, A. J. (1973), J. Amer. Chem. Soc. 95, 3065-3067.

Krieger, M., Kay, L. M., and Stroud, R. M. (1974), J. Mol. Biol. 83, 209-230.

Krieger, M., Koeppe, R. E. II, and Stroud, R. M. (1976), Biochemistry--in press.

Markley, J. L. and Porubcan, M. A. (1975), submitted to J. Mol. Biol.

Martinek, K., Vill', Kh., Strel'tsova, Z. A., and Berezin, I. V. (1969), Mol. Bio. (USSR) 3, 554-565.

Matsubara, H. and Sasaki, R. M. (1968), J. Biol. Chem. 243, 1732-1757.

Parsons, S. M. and Raftery, M. A. (1970), Biochem. Biophys. Res. Commun. 41, 45-49.

Parsons, S. M. and Raftery, M. A. (1972), Biochemistry 11, 1630-1633.

Pollock, E., Hogg, J. L., and Schowen, R. L. (1973), J. Amer. Chem. Soc. 95, 968-969.

Raftery, M. A. and Cole, R. D. (1963), Biochem. Biophys. Res. Commun. 10, 467-472.

Rajender, S., Lumry, R., and Han, M. (1971), J. Phys. Chem. 75, 1375-1386.

Robillard, G. and Shulman, R.G. (1974), J. Mol. Biol. 86, 519-540.

Robinson, N.C., Tye, R.W., Neurath, H., and Walsh, K.A. (1971), Biochemistry 10, 2743-2747.

Schwert, G.W. and Takenaka, Y. (1955), Biochim. Biophys. Acta 16, 570-575.

Sigler, P.B., Blow, D.M., Matthews, B.W., and Henderson, R. (1968), J. Mol. Biol. 35, 143-164.

Stewart, J.A. and Dobson, J.E. (1965), Biochemistry 4, 1086-1091.

Stroud, R.M., Krieger, M., Koeppe, R.E. II, Kossiakoff, A.A. and Chambers, J.L. (1975). In Proteases and Biological Control (Eds. E. Reich, D.B. Rifkin, and E. Shaw), Cold Spring Harbor Laboratory, pp. 13-32.

Susi, H. (1972), Methods in Enzymol. 26, 381-391.

Timasheff, S.N. and Rupley, J.A. (1972), Arch. Biochem. Biophys. 150, 318-323.

Timasheff, S.N., Susi, H., and Rupley, J.A. (1973), Methods In Enzymol. 27, 548-557.

Titani, K., Ericsson, L.H., Neurath, H., and Walsh, K.A. (1975), Biochemistry 14, 1358-1366.

Walsh, K.A. and Neurath, H. (1964), Proc. Nat. Acad. Sci. USA 52, 884-889.

Table I

Activity of trypsin-SCZ toward specific and nonspecific substrates. The data show retention of the catalytic site but loss of the binding pocket carboxyl (Asp 189) in 90% of the molecules. Assay conditions are given in the Methods Section.

Substrate	Enzyme	Activity of pH 7.0, 25°C with 10^{-2} M BA*	
Specific BAEE	E	100%	
	E-SCZ	8% ±0.2%	
Nonspecific Ac-Gly-NPE	E	100%	16% ± 4%
	E-SCZ	146% ±8%	(92% ± 5% of E-SCZ)

Substrate	Enzyme	Moles Active Site per Mole A_{280} Units
Specific NPGB	E, from bottle	0.7 ± 0.1
	E, after G-25 column at pH 3	0.95 ± 0.07
	E-SCZ, after dialysis	0.11 ± 0.03

*BA refers to benzamidine hydrochloride.

Table II. Relative Molar Ratios of Amino Acids in Thermolytic Peptides of Trypsin-SCZ†

Amino Acid	Peptides C-1	C-2	G-102	G-194	G-245	Z-102	Z-194
Asp	3.00(3)	2.30(3)	3.00(3)	2.00(2)	1.00(1)	3.66(3)	2.00(2)
Thr	.05	.02			.14		
Ser	.07	.11	1.01	2.25(2)	.89(1)		2.27(2)
Glu	.09	.11	2.11	2.05(2)	.25	1.38	1.82(2)
Pro			0.79	1.01(1)	.26		0.86(1)
Gly	.07	.11	2.02(1)	5.96(6)	.93(1)	0.95	5.35(5)
Ala			0.63		.99(1)		
Val			0.69	1.61(1) or (2)	.35	0.34	1.39(1) or (2)
Met		1.14(1)					
Ile		1.00(1)	0.57		.34(1*)	0.18	
Leu	.26(1*)	1.23(2*)	0.91(1*)	0.56(1*)	.38	1.00(1)	0.79(1)
Tyr			0.61		.14		
Phe					.12		
His					.22		
Lys				1.03(1)	.14		1.17(1)
Arg					.21		

Table II (continued)

†The codes for the peptides refer to: C - peptides of trypsin-SCZ having a net negative charge from a pH 6.5 map stained with ninhydrin; G - radioactive peptides from trypsin-SCZ which had been denatured and labeled with ^{14}C-Gly-Et, pH 6.5 map stained with ninhydrin; Z - trypsin-SCZ, denatured and labeled with ^{14}C-SCZ, radioactive peptides from an unstained pH 6.5 map. The numbers in parentheses are the expected values for amino acid composition based on the sequence assignments for the peptide fragments as listed in Table III. * - N-terminal residue, partially destroyed by ninhydrin.

Table III

Peptide assignments based on the data in Table II and the sequence of Titani et al. (1975).

	Peptide	Contaminant
C-1	Leu_{99}-Asn-Asn-Asp_{102}	None
C-2	Leu_{99}-Asn-Asn-Asp_{102}-Ile-Met-Leu	None
G-102	Leu_{99}-Asn-Asn-Asp_{102} (*Gly-Et)	Yes
Z-102	Leu_{99}-Asn-Asn-Asp_{102} (*SCZ)	Val_{76}-Glu-Gly-Asn-Glu_{80} (SCZ, SCZ)
G-194 or Z-194	Leu_{185}-Glu-Gly-Gly-Lys-Asp_{189}-Ser-Cys-Gln-Gly-Asp_{194}-Ser-Gly-Gly-Pro-Val-(Val) (*(Gly-Et or SCZ))	None
G-245	Ile_{242}-Ala-Ser-Asn_{245}-*Gly-Et	Yes

Table IV

Radioactive Peptides Found in [^{14}C]-SCZ-treated Trypsin-SCZ

Sequence	Specific Activity	Unblocked Residue
Leu-Asn-Asn-Asp_{102}	1.02 ± 0.1	Asp 102
Leu_{185} to Val_{199}	0.97 ± 0.1	Asp 194
Ile_{242} to Asn_{245}	N. D. *	Asn 245
Entire protein	2.55 ± 0.12	

*Not determined. (Estimate 0.3 - 0.5 from ^{14}C-Gly peptide map.)

Figure Legends

FIGURE 1. Direct infrared spectrum of trypsin in D_2O, recorded on a Perkin-Elmer Model 225 Spectrophotometer. Path length, 0.125 mm. 1.5 mM trypsin, 12 mM benzamidine, 6 mM NaCl, pD 4.6. D_2O in reference cell.

FIGURE 2. Difference infrared spectra of trypsin recorded on a Perkin Elmer Model 225 spectrophotometer. Path length, 0.125 mm. 1.5 mM trypsin, 12 mM benzamidine, 6 mM NaCl. The pD's of the solutions in the sample and reference cells are listed next to each spectrum. Except in the two spectra where $\Delta pD = 0$ (where arrows indicate the expected carboxyl absorption frequencies), the arrows show the peak heights used for the titration curve in Figure 3.

FIGURE 3. Titration curve for the carboxyl groups of bovine trypsin. Data are taken from the heights of the peaks in the series of difference infrared spectra shown in Figure 2.

FIGURE 4. ORTEP drawing of the active site of trypsin with a model of semicarbazide built into the binding pocket. Coordinates

for the drawing are from Krieger et al. (1974). The alpha-carbon atoms are numbered. The SCZ group does not come closer than 7.3 Å from the hydroxyl of Ser 195. This computer plot was drawn using a modified version of ORTEP (Johnson, 1965).

FIGURE 5. Ninhydrin staining patterns of pH 6.5 peptide maps of thermolytic digests of S-aminoethyl trypsin-SCZ with and without carbodiimide coupling of glycine ethyl ester under denaturing conditions. The origin of each map is near the bottom edge and just to the right of center. Descending chromatography in the vertical direction was followed by electrophoresis in which the positive pole was on the right and the negative pole on the left. The map of trypsin-SCZ (A) contains two peptides of net negative charge. Both have been shown to be peptides containing Asp 102 (see text). After blocking the acid of Asp 102, these peptides run in the neutral region of the map at pH 6.5 (B). The negatively charged spot in map B is phenol red, which was used as a marker dye.

FIGURE 6. Autoradiogram and drawing of peptide map of ^{14}C-Gly-Et labeled S-aminoethyl trypsin-SCZ. The origin is in the lower right-hand corner of the map. Descending chromatography in the vertical direction was followed by electrophoresis in the horizontal

direction. The three major radioactive peptides contain Asp 102, Asp 194, and the α-carboxyl terminus of the protein at Asn 245 (00).

FIGURE 7. Autoradiogram and drawing of peptide map of ^{14}C-SCZ labeled S-aminoethyl trypsin-SCZ. The origin is in the lower right-hand corner of the map. Descending chromatography in the vertical direction was followed by electrophoresis in the horizontal direction. The two major radioactive peptides contain 1 mole of ^{14}C-label per mole of Asp 102 and Asp 194, respectively. In this case the C-terminus did not appear in a single cleanly digested peptide, but was scattered over the map.

FIGURE 8. Difference infrared spectra of trypsin-SCZ recorded on a Perkin Elmer Model 225 spectrophotometer. Path length, 0.150 mm. 1.5 mM trypsin-SCZ, 12 mM benzamidine, 6 mM NaCl. The pD's of the solutions in the sample and reference cells are listed next to each spectrum. The carboxyl groups which have been modified with semicarbazide do not show up in these spectra. The spectra indicate that one or more groups which absorb at 1705 cm^{-1} titrate at low pH, and one or more groups which absorb at 1680 cm^{-1} titrate around pH 6.8.

FIGURE 9. Difference infrared spectra recorded on a Beckman Model 4240 spectrophotometer interfaced to a Fabri-Tek Model 1062 computer. A) 1.5 mM trypsin-SCZ, 6 mM NaCl, 12 mM benazmidine. B) 1.5 mM trypsin-SCZ, 2 mM $CuCl_2$, 4 mM NaCl, 12 mM benzamidine. C) 20 mM pyro-Glu-His-Pro-NH_2. The vertical scale is absorbance (rather than transmittance as used elsewhere). All spectra are computer sums of 5 repeated scans; and all are plotted on the same scale, except that in Column B, the spectrum having $\Delta pD = 3.6 - 5.6$ is plotted on a scale expanded by a factor of 2 over the others. Baselines were obtained using protein solutions of identical pD in both cells and were subtracted to give the spectra shown in columns A and B.

FIGURE 10. Titration curve for the carboxyl groups of trypsin-SCZ. Data are taken from the peak heights in a series of difference infrared spectra including those shown in Figure 8. Of the total of 2.5 carboxyl groups per molecule, 1.5 titrate with an average pK_a of 2.9, while one carboxyl titrates with a pK_a of 6.8, as evidenced by comparing the slopes of the two titrations.

FIGURE 11. Difference infrared spectra of trypsin-SCZ in the presence of copper (Cu^{++}) ions recorded on a Perkin Elmer Model

225 spectrophotometer. 6 mM $CuCl_2$, 1.5 mM trypsin-SCZ, 12 mM benzamidine, path length = 0.150 mm.

FIGURE 12. Difference infrared spectra of control compounds.

III) 20 mM glycine; IV) 30 mM histidine; V) 20 mM Gly-His-Gly; and VI) 20 mM pyro-Glu-His-Pro-amide.

B - Schematic drawing to scale with each spectrum normalized to an equivalent concentration.

FIGURE 13. State of ionization of the active site residues above and below pH 6.8.

Figure 1

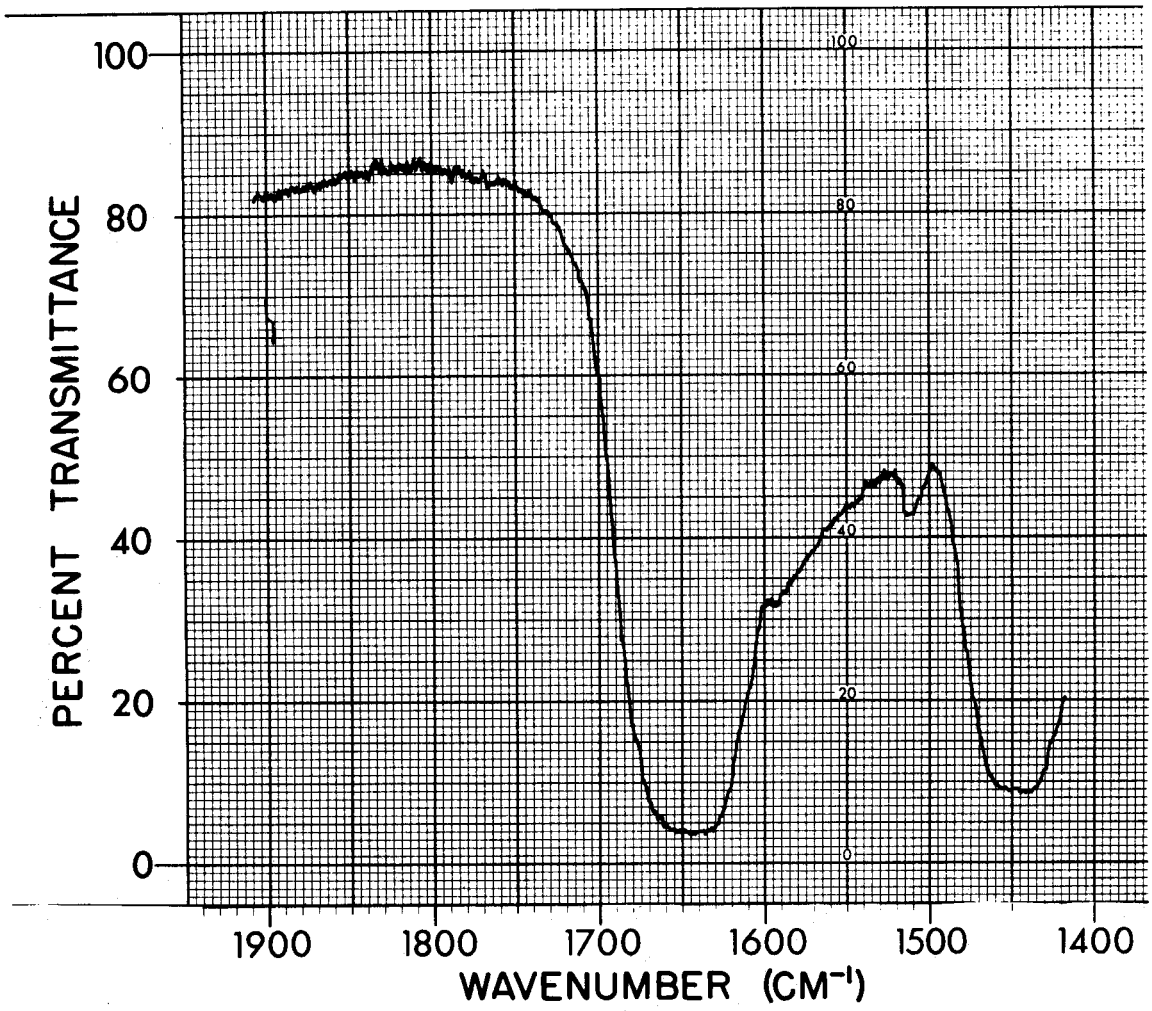

100
80
60
40
20
0
PERCENT TRANSMITTANCE
1900
1800
1700
1600
1500
1400
WAVENUMBER (CM⁻¹)

Figure 2

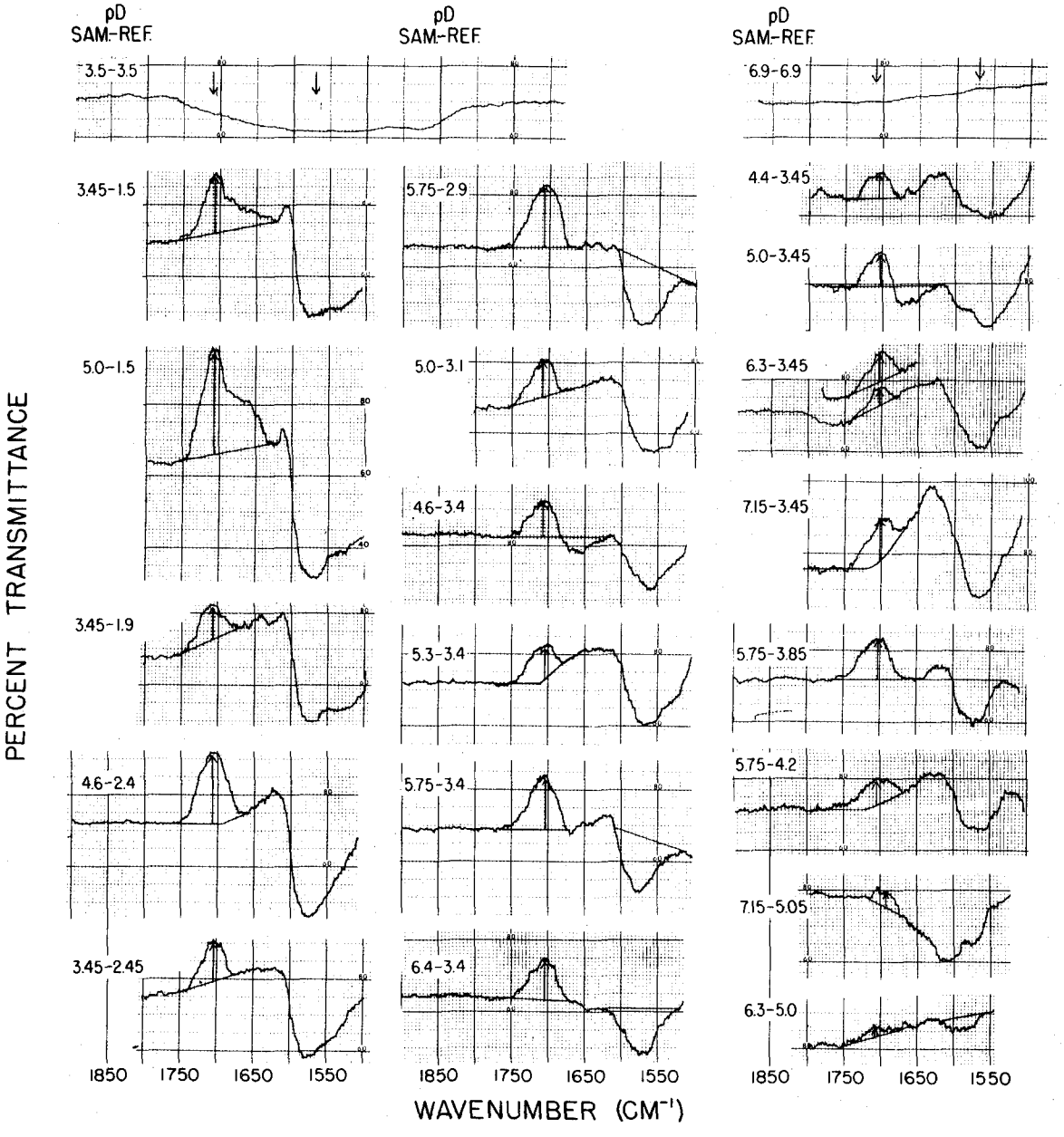

pD
SAM-REF.
3.5-3.5
3.45-1.5
5.0-1.5
3.45-1.9
4.6-2.4
3.45-2.45
pD
SAM-REF.
5.75-2.9
5.0-3.1
4.6-3.4
5.3-3.4
5.75-3.4
6.4-3.4
pD
SAM-REF.
6.9-6.9
4.4-3.45
5.0-3.45
6.3-3.45
7.15-3.45
5.75-3.85
5.75-4.2
7.15-5.05
6.3-5.0
PERCENT TRANSMITTANCE
1850
1750
1650
1550
WAVENUMBER (CM⁻¹)

Figure 3

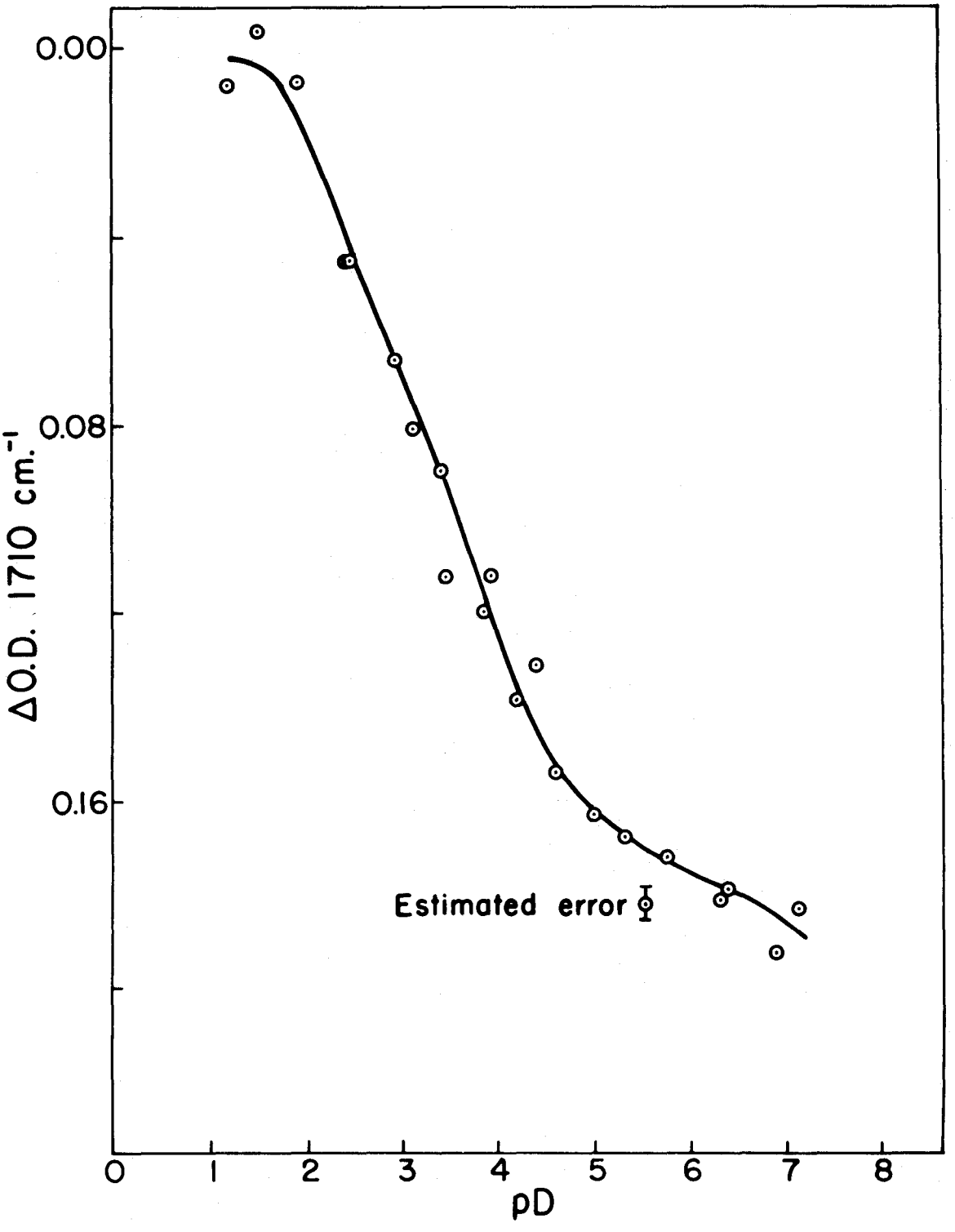


Figure 4

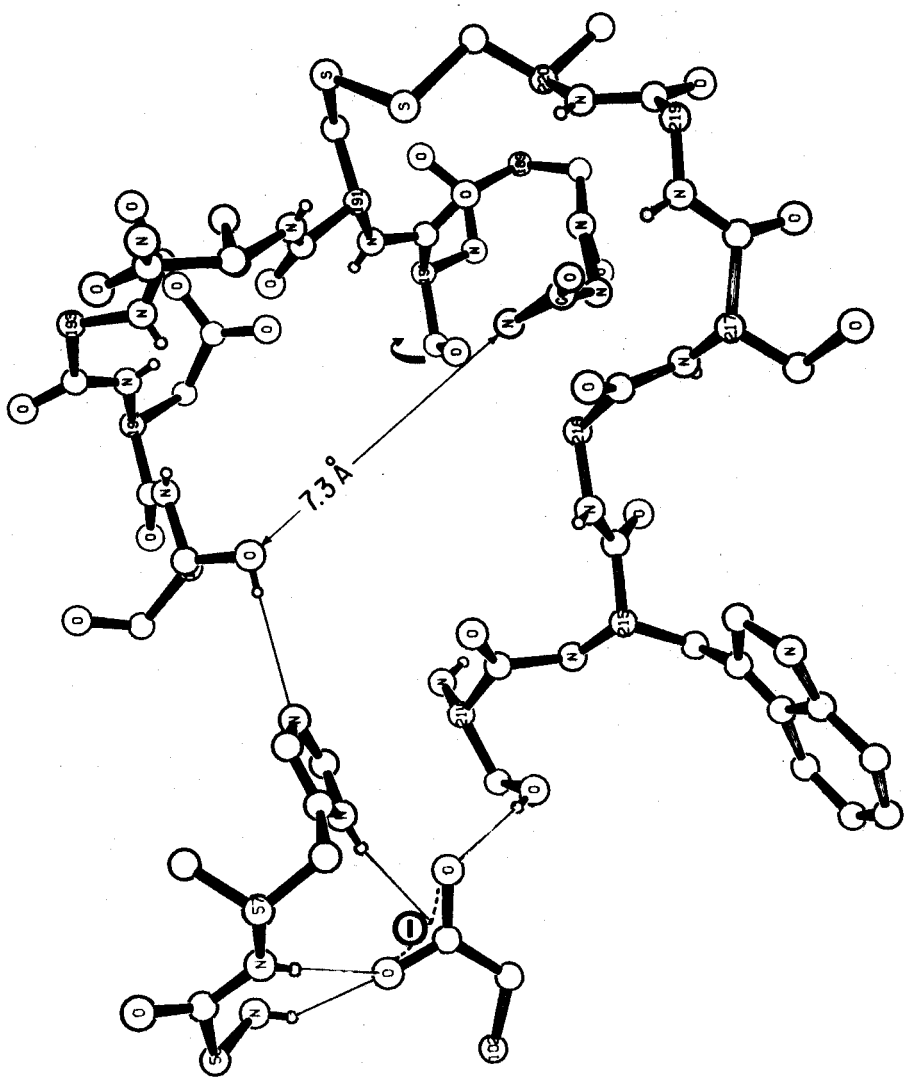
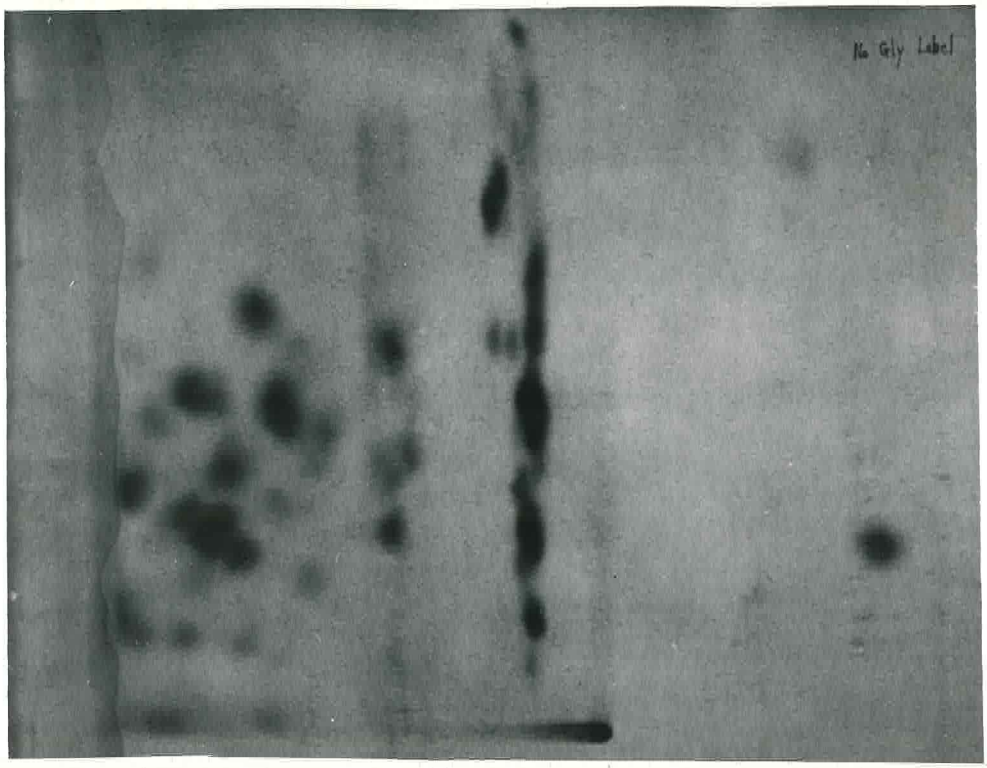


Figure 5

A



B

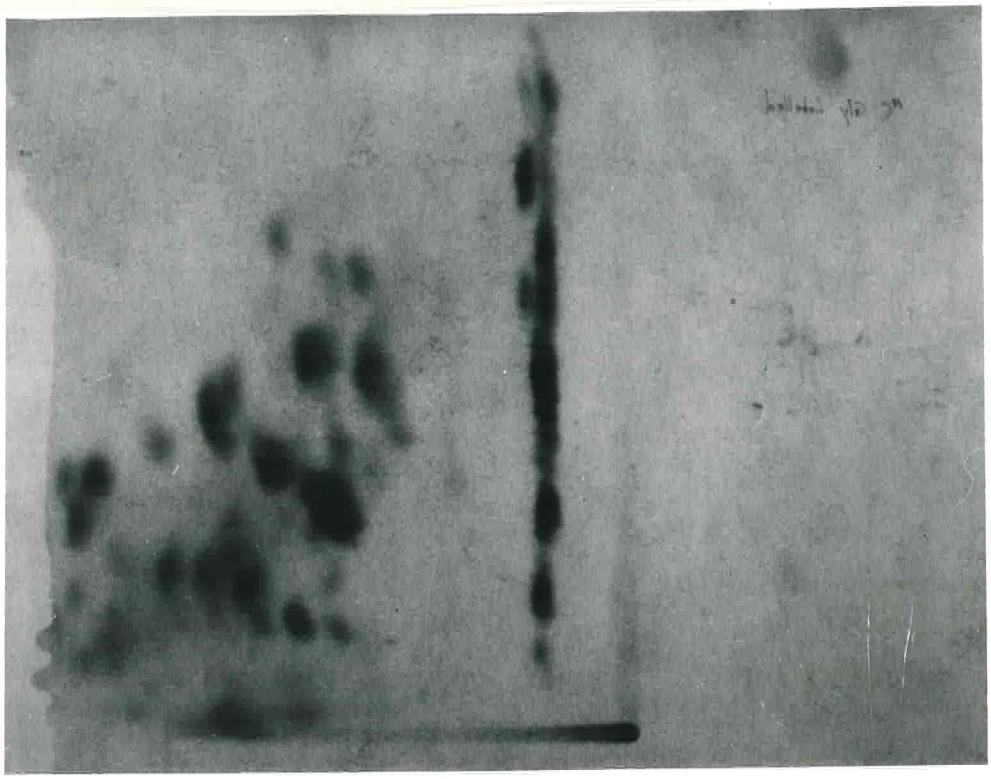


Figure 6

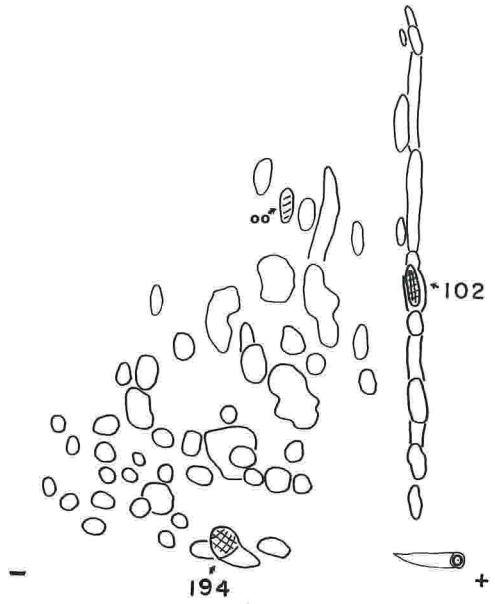


Figure 7

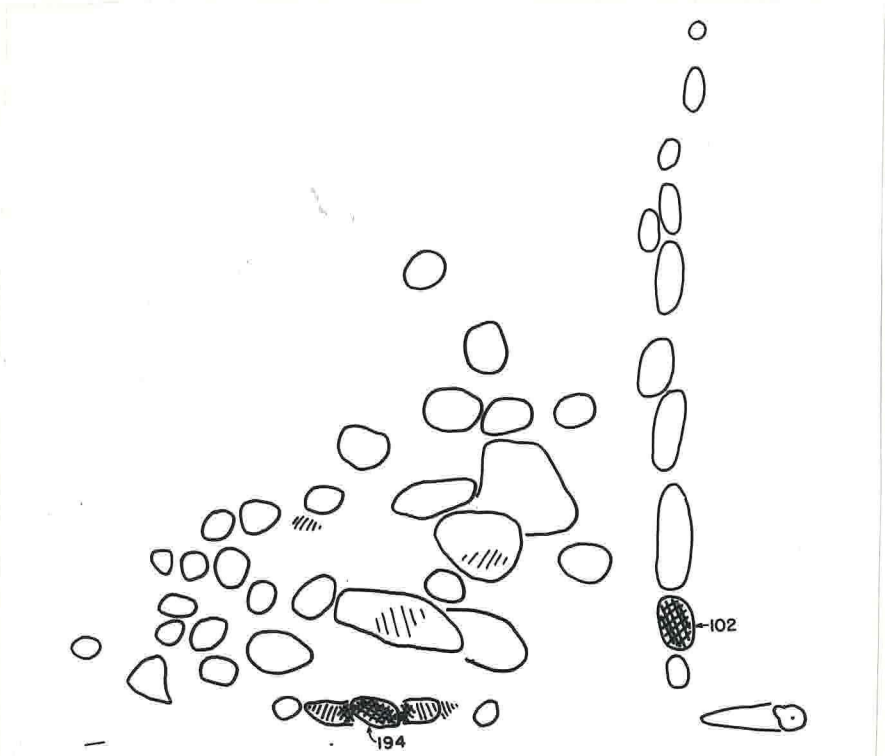


Figure 8

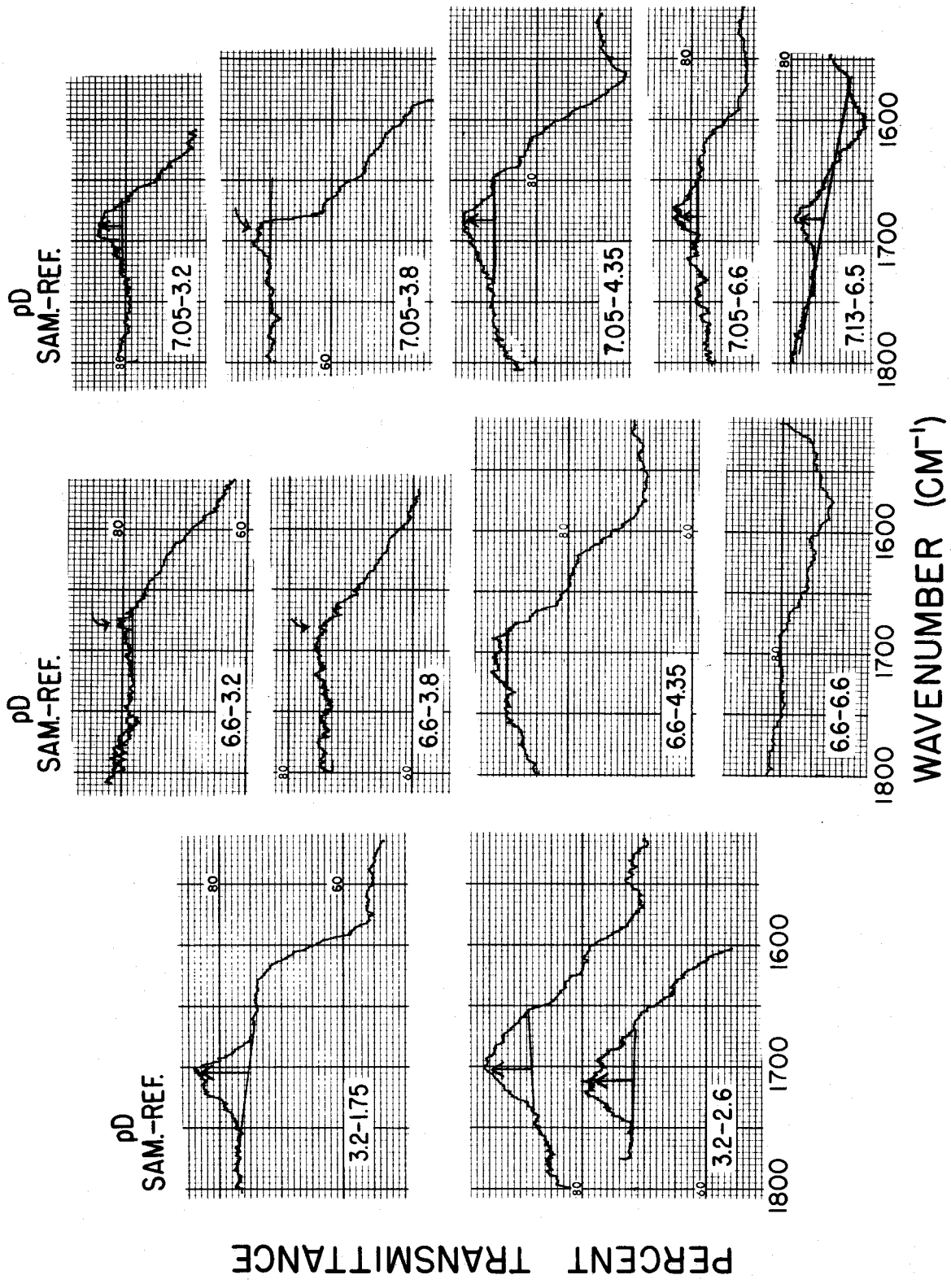

pD
SAM.-REF.
7.05-3.2
7.05-3.8
7.05-4.35
7.05-6.6
7.13-6.5
6.6-3.2
6.6-3.8
6.6-4.35
6.6-6.6
3.2-1.75
3.2-2.6
80
60
1800
1700
1600
WAVENUMBER (CM⁻¹)
PERCENT TRANSMITTANCE

Figure 9

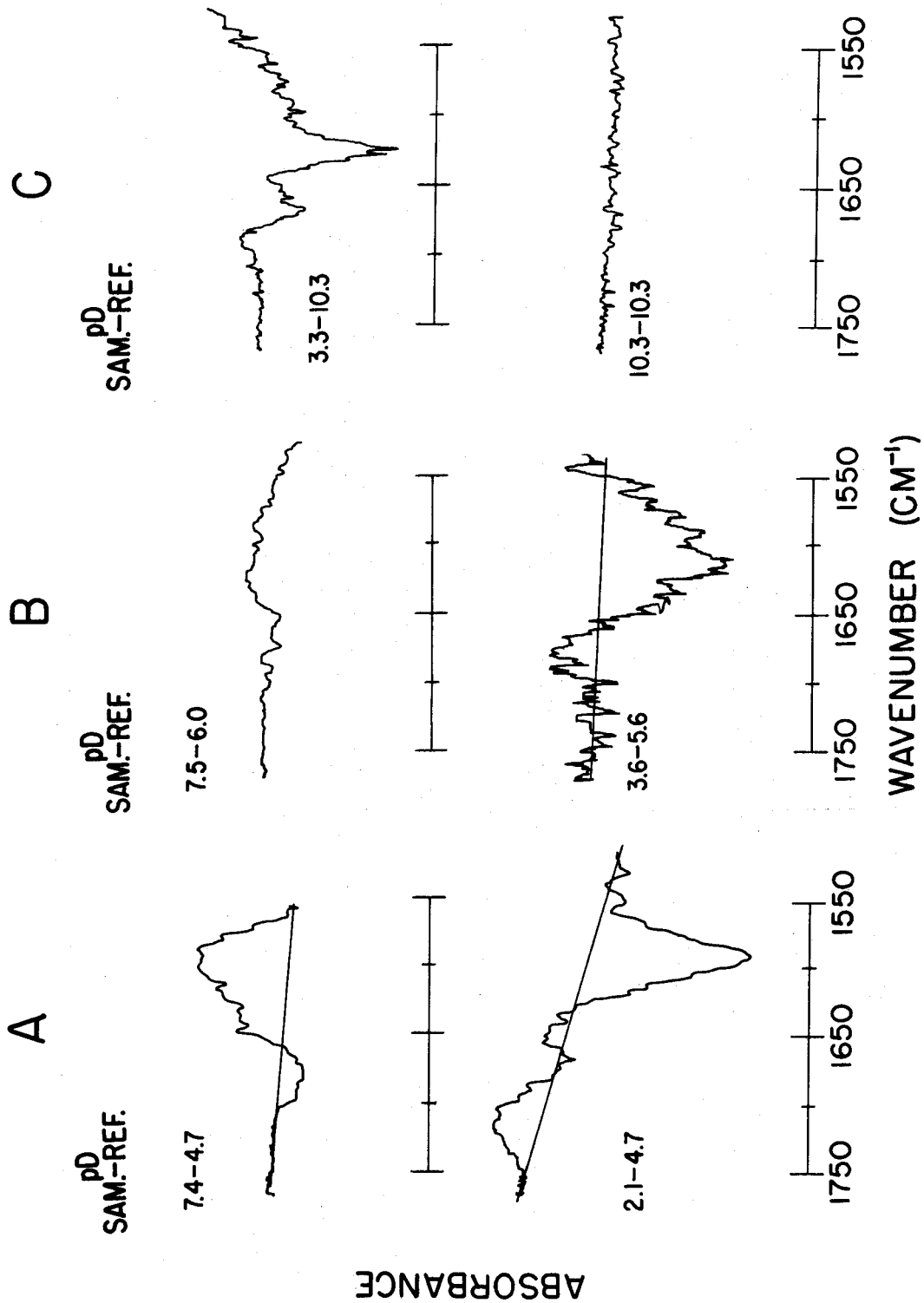

A
B
C
SAM.-REF.
pD
7.4-4.7
7.5-6.0
3.3-10.3
2.1-4.7
3.6-5.6
10.3-10.3
1750
1650
1550
WAVENUMBER (CM⁻¹)
ABSORBANCE

Figure 10

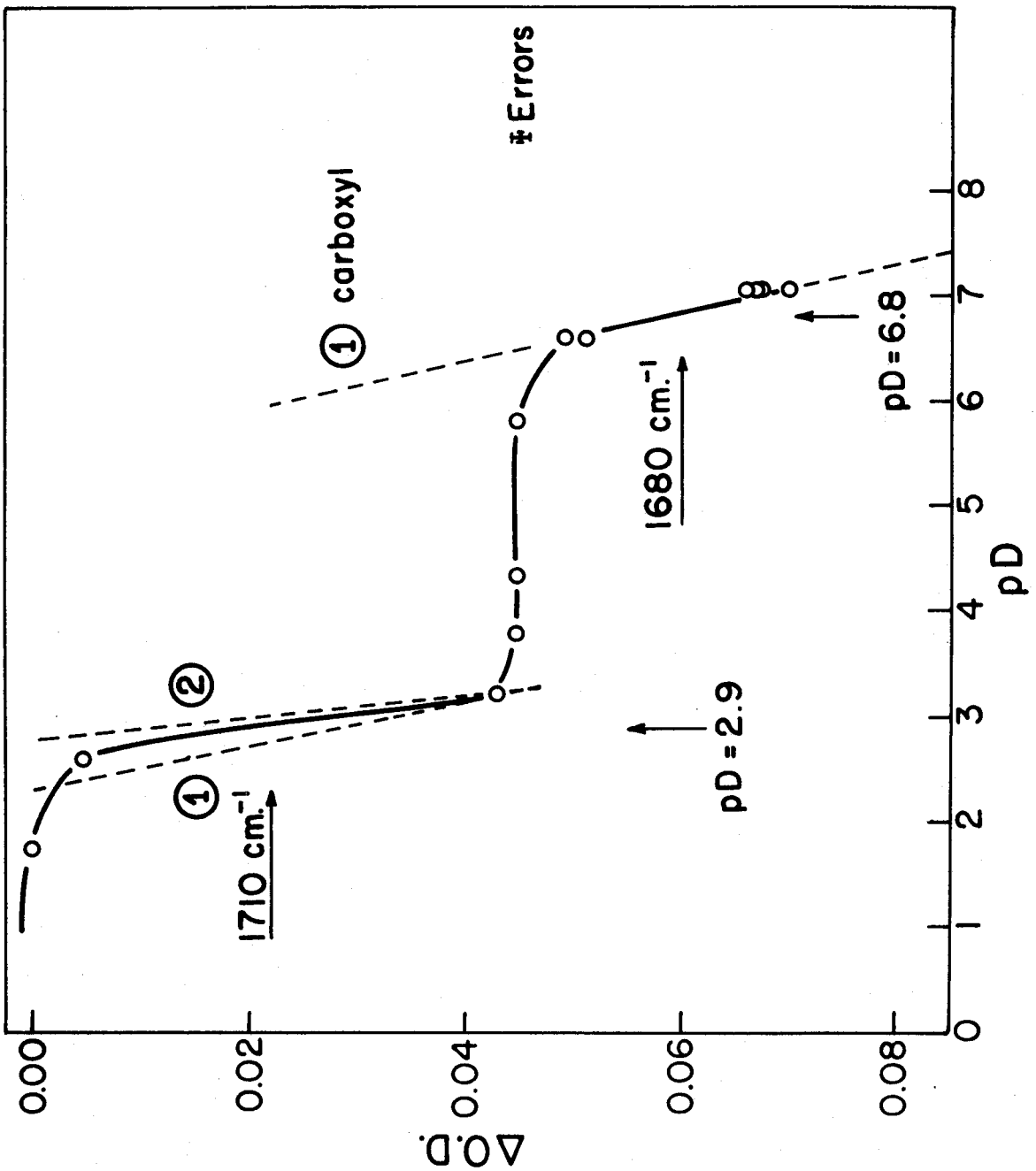

± Errors
① carboxyl
1680 cm.⁻¹
1710 cm.⁻¹
pD=6.8
pD=2.9
①
②
ΔO.D.
0.00
0.02
0.04
0.06
0.08
pD
0
1
2
3
4
5
6
7
8

Figure 11

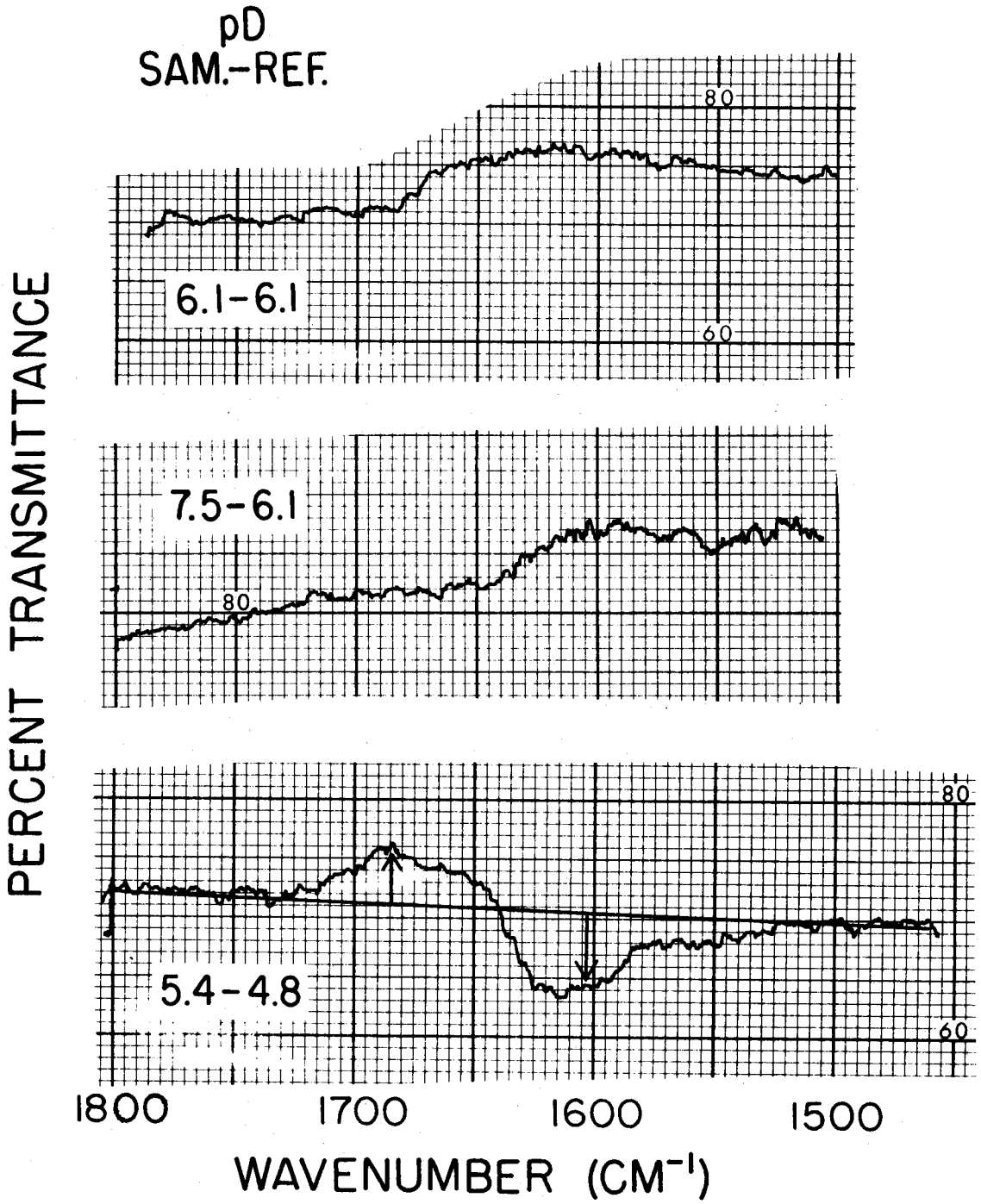

pD
SAM.-REF.
80
60
6.1-6.1
7.5-6.1
80
5.4-4.8
80
60
PERCENT TRANSMITTANCE
1800
1700
1600
1500
WAVENUMBER (CM⁻¹)

Figure 12A

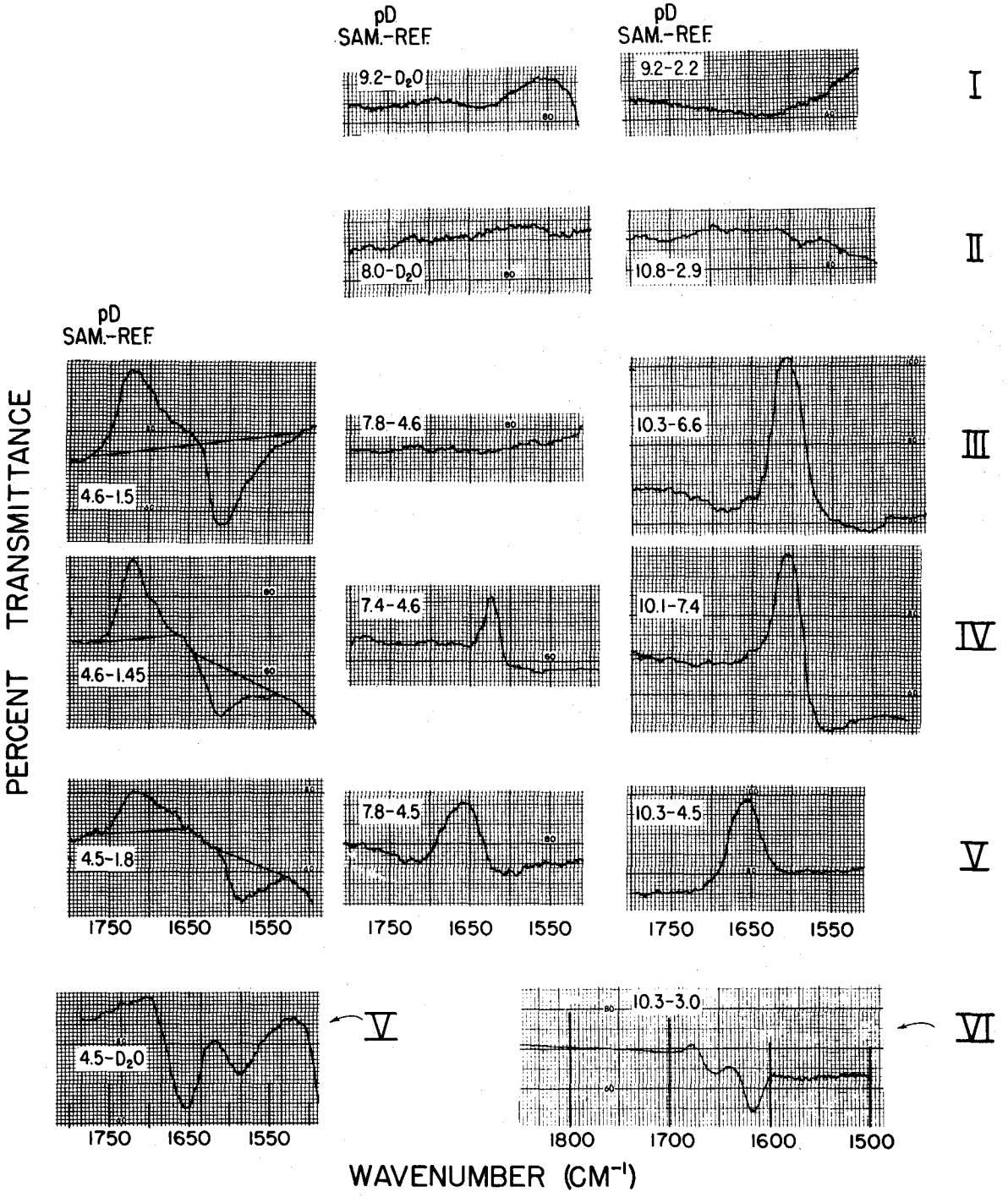

pD
SAM.-REF.
9.2-D2O
9.2-2.2
I
8.0-D2O
10.8-2.9
II
pD
SAM.-REF.
4.6-1.5
7.8-4.6
10.3-6.6
III
4.6-1.45
7.4-4.6
10.1-7.4
IV
4.5-1.8
7.8-4.5
10.3-4.5
V
1750 1650 1550
1750 1650 1550
1750 1650 1550
4.5-D2O
V
1750 1650 1550
10.3-3.0
VI
1800 1700 1600 1500
PERCENT TRANSMITTANCE
WAVENUMBER (CM-1)

Figure 12B

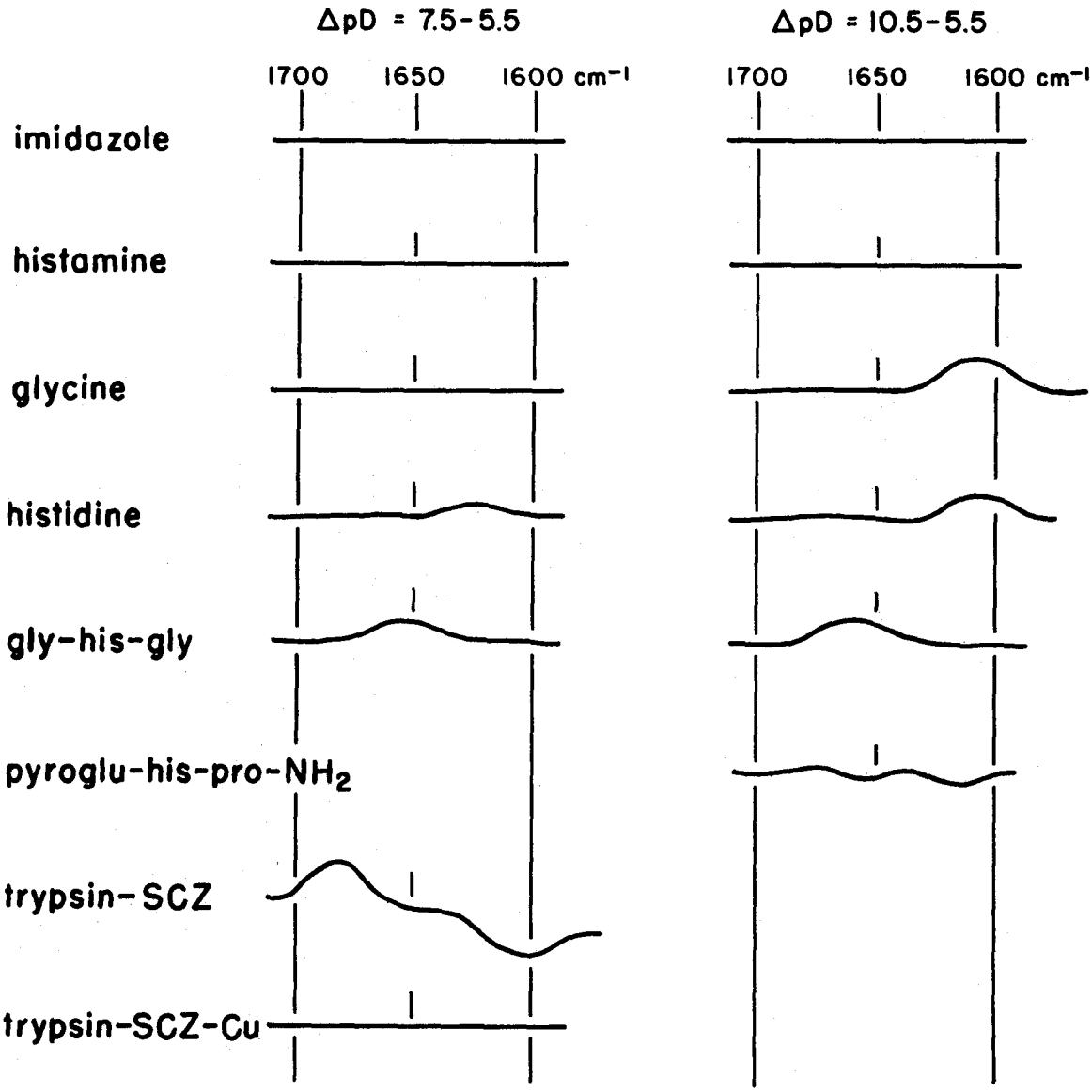
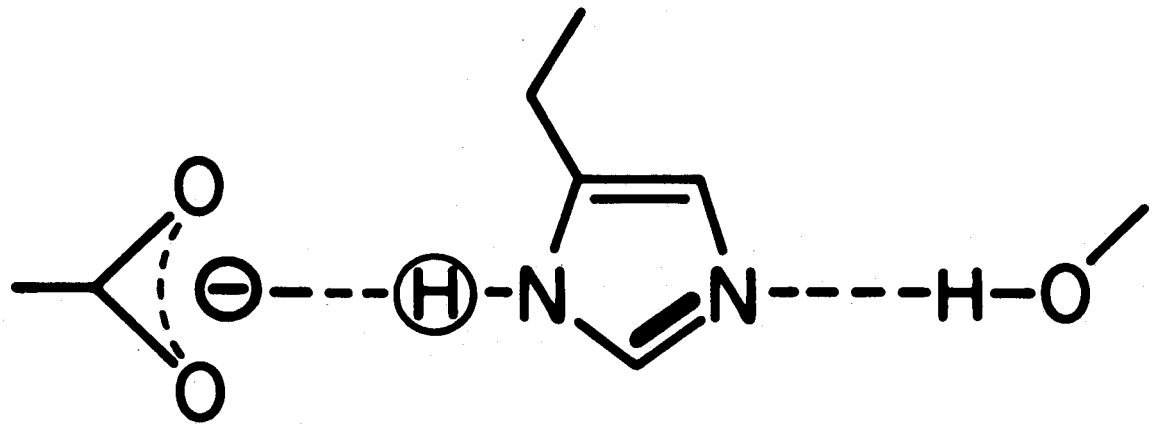
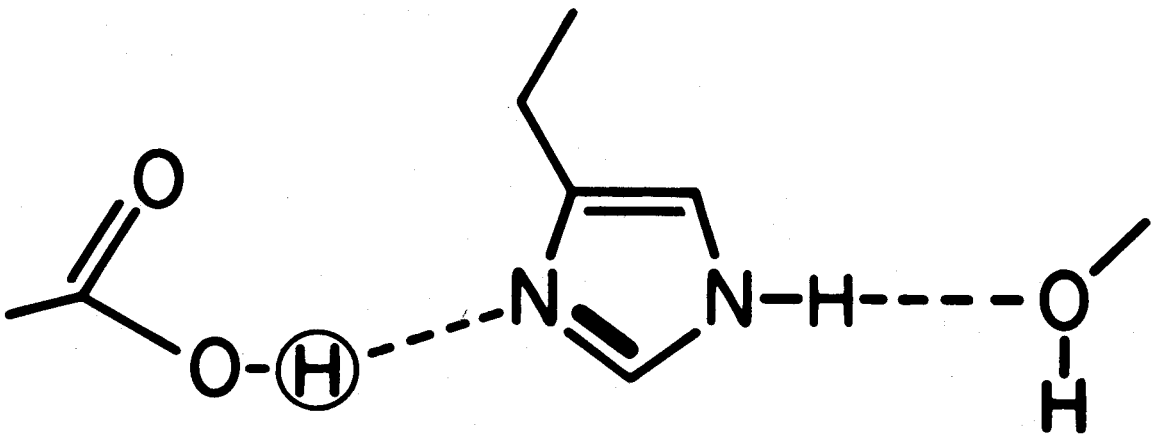

ΔpD = 7.5-5.5
ΔpD = 10.5-5.5
1700
1650
1600 cm⁻¹
imidazole
histamine
glycine
histidine
gly-his-gly
pyroglu-his-pro-NH2
trypsin-SCZ
trypsin-SCZ-Cu

Figure 13



Active form: above pH 6.7



Inactive below pH 6.7

CHAPTER III: A Crystallographic Determination of the Site at Which Inhibitory Cu^{++} Ion Binds to Trypsin

Introduction

The use of copper ions in Chapter II to assign the carboxyl infrared peaks of trypsin-semicarbazide requires a proof of the copper binding site. The assumption that Cu^{++} would perturb the titration of Asp 102 rather than Asp 194 or the C-terminus was based on two facts: 1) that Ag^{+} binds between His 57 and Asp 102 in DIP-trypsin (Chambers et al., 1974), and 2) that Cu^{++} and Ag^{+} compete with each other in inhibiting chymotrypsin (Martinek et al., 1969). Although this evidence is suggestive that Cu^{++} may bind between Asp 102 and His 57, the observation that silver and copper compete kinetically does not necessarily mean that they bind to the same site. A positive proof of the copper binding site can be provided by a crystallographic determination of the location of bound copper in crystals of trypsin. The experiment described in this chapter was carried out by John L. Chambers, Lois M. Kay, and the author.

Experimental

Crystals of DIP-trypsin were prepared as described by Stroud et al. (1974). Copper was diffused into the crystals by soaking them at 20°C, for six days in the dark, in a solution of 17 mM $Cu(NO_3)_2 \cdot 3H_2O$ (Merck reagent) containing 0.66 M $MgSO_4$, 11 mM benzamidine·HCl (Aldrich), and 0.05 M cacodylate buffer at pH 6.9.

X-ray diffraction data were collected using a Syntex $P\bar{1}$ automated four-circle diffractometer using the Wyckoff method of scanning through steps in ω (Wyckoff et al., 1967). (For a description of the four-circle geometry, including definitions of the 2θ, φ, ω, and χ angles, see Stout and Jensen (1968).) Only the tops of the peaks were scanned, not the entire profiles. Seven steps were used to cover the scan range of 0.1° in a time of 50 seconds. If the largest number of counts did not occur in the middle step of the seven, the diffractometer automatically took one or more extra steps--up to a maximum of six extra steps--until the highest step was flanked by at least three others on each side. In computing intensities, the five contiguous steps which gave the highest total number of counts were used. The scan rate was slowed to 0.08° per minute for the reflections which had values of 2θ between 29.5° and 34.0°, and to 0.05° per minute for 2θ between 34.0° and 35.5°.

The intensities of 7704 independent reflections were measured in the range of 2θ between 3.0° and 35.5°; 6905 of these reflections had intensities more than two standard deviations above background. For the monochromatized CuK_α radiation (λ = 1.54 Å) which was used, a value of 2θ = 35.5° corresponds to a resolution ($\frac{\lambda}{2\sin\theta}$) of 2.53 Å. In order to monitor damage to the crystal caused by the x-ray beam, five check reflections were measured repeatedly after every 120 reflections throughout the entire data collection process. The check reflections were well spaced from each other in the reciprocal lattice at values of 2θ ranging from 7.86° to 25.11°, and were used to calculate a two-parameter least squares fit for the decay of peak intensities as functions of time and of $\sin\theta$. A single crystal was used for the entire data collection, and the intensities of the check reflections diminished to values ranging from 85% to 97% of original during the 148-hour exposure to x-rays.

Rather than observing individual backgrounds next to each reflection, backgrounds were anisotropically interpolated from curves measured at the end of the data collection, according to the method of Krieger et al. (1974a). The backgrounds were subtracted before applying the time decay correction to each reflection.

The diffracted intensity measurements were further corrected for absorption (North et al., 1968), and for Lorentz and polarization

factors (Levy and Ellison, 1960) using standard techniques. The final corrected intensities (I) were converted to structure factors ($F \alpha \sqrt{I}$) and were scaled to the data for the isomorphous crystals of DIP-trypsin containing no copper (Stroud et al., 1974; Chambers et al., 1974). The scaling was performed according to the procedure of Krieger et al. (1974b) by determining the best two-parameter (k, B) fit to the equation

$$\ln \frac{\langle F \rangle_1}{\langle F \rangle_2} = \ln(k) - B \cdot s,$$

computed from 20 zones of $s = 4 \sin^2 \theta$; where $\langle F \rangle_1$ and $\langle F \rangle_2$ are the average values of F in the zone of thickness $\Delta s = 0.017$ for data sets 1 and 2 respectively.

Difference electron density maps were calculated between copper-DIP-trypsin and DIP-trypsin, using appropriately scaled structure factors. The difference maps were calculated using the Fourier transform expression:

$$\Delta\rho(\underline{r}) = \frac{1}{V} \sum_s m \Delta F_s \exp (2\pi i \, \underline{r} \cdot \underline{s} + \varphi)$$

where ρ is the electron density in real space at the point determined by the vector $\underline{r}$ from the origin and V is the volume of one unit cell of the crystal. φ and ΔF_s are respectively the phase and the difference

between the structure factors for the copper derivative and the native trypsin at the particular reflection in reciprocal space located by the vector s from the origin. The phases, φ, and associated figures of merit, m, were those determined by Stroud et al. (1974) using the method of multiple heavy atom isomorphous replacement for DIP-trypsin to a resolution of 2.7 Å.

The difference map was contoured at intervals of 0.1 electron/Å^3, beginning with the first contour at 0.2 electron/Å^3, using an approximate absolute scale factor (Stroud et al., 1974).

Results

The difference electron density map for Cu^{++}-DIP-trypsin minus DIP-trypsin contain only two significant peaks of positive electron density (Figure 1). These peaks are in the same positions found for silver-DIP-trypsin (Chambers et al., 1974), and correspond to the location of the bound Cu^{++} ion (large peak) and to the movement of the ring of His 57 (small peak). The binding of Ag^{+} or Cu^{++} forces the imidazole ring away from its native position and in the direction of the solvent.

This crystallographic determination of the site of copper binding in crystals of DIP-trypsin proves the assumption stated in

Chapter II that Cu^{++} and Ag^{+} inhibit trypsin at the same site. The result validates the use of Cu^{++} ion as part of the procedure of assigning the difference infrared peaks to the individual carboxyls of trypsin-SCZ. Cu^{++} perturbs the titration of the carboxyl group which has a pK_a of 6.8. Since Cu^{++} binds to Asp 102--and not to Asp 194 or to Asn 245--Asp 102 is the carboxyl group in trypsin which has a pK_a of 6.8.

References

Chambers, J.L., Christoph, G.G., Krieger, M., Kay, L., and Stroud, R.M. (1974), Biochem. Biophys. Res. Commun. 59, 70-74.

Krieger, M, Chambers, J.L., Christoph, G.G., Stroud, R.M., and Trus, B.L. (1974a), Acta Crystallogr. A30, 740-748.

Krieger, M., Kay, L.M., and Stroud, R.M. (1974b), J. Mol. Biol. 83, 209-230.

Levy, H.A. and Ellison, R.D. (1960), Acta Crystallogr. 13, 270-271.

Martinek, K., Vill', Kh., Strel'tsova, Z.A., and Berezin, I.V. (1969), Mol. Biol. (USSR) 3, 554-565.

North, A.T.C., Phillips, D.C., and Mathews, F.S. (1968), Acta Crystallogr. A24, 351-359.

Stout, G. H., and Jensen, L. H., X-ray Structure Determination (A Practical Guide), pp. 177-194, 217. Collier-Macmillan Ltd., London, 1968.

Stroud, R. M., Kay, L. M., and Dickerson, R. E. (1974), J. Mol. Biol. 83, 185-208.

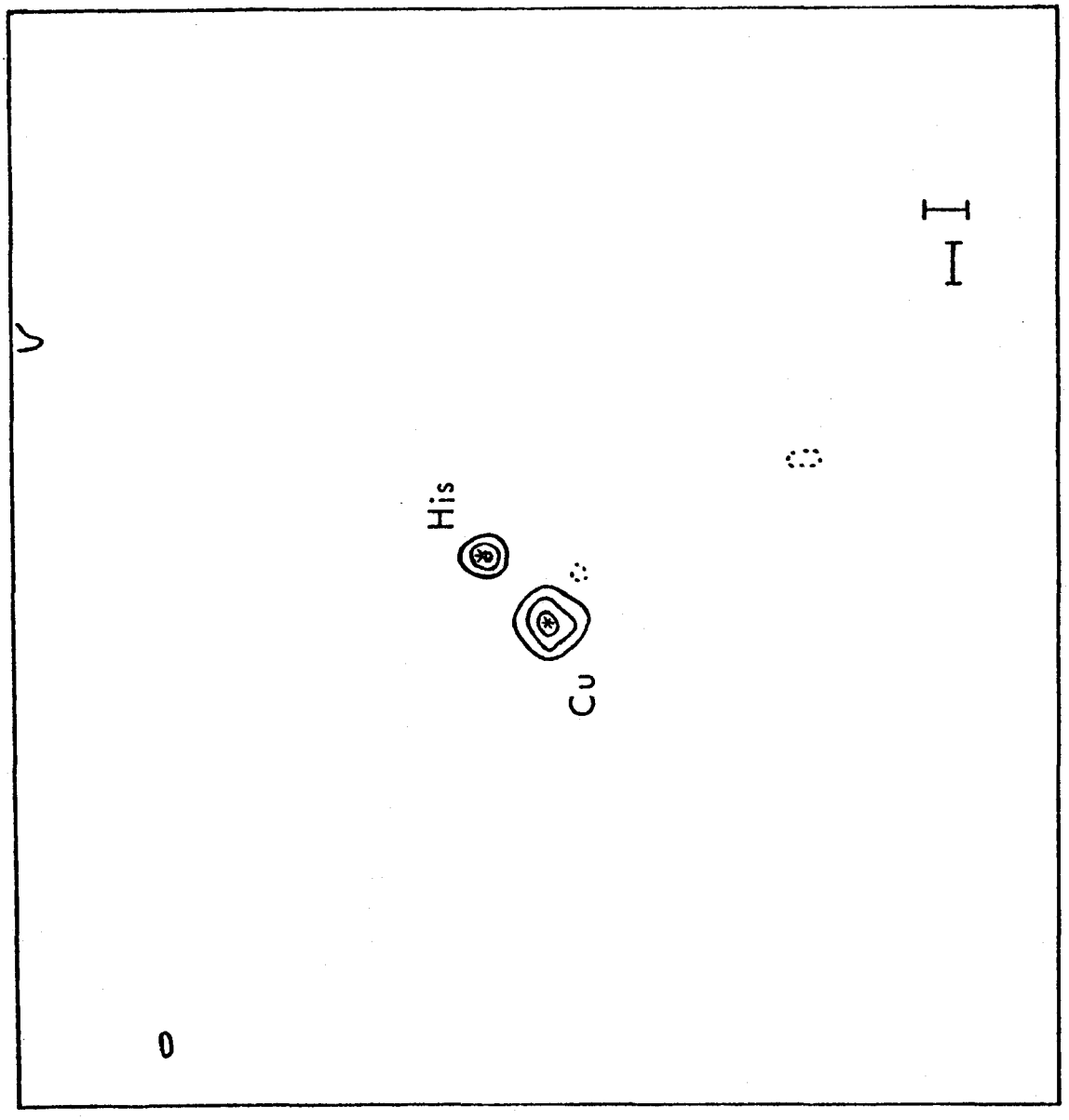
Wyckoff, H. W., Doscher, M., Tsernoglou, D., Inagami, T., Johnson, L. N., Hardman, K. D., Allewell, N. M., Kelly, D. M., and Richards, F. M. (1967), J. Mol. Biol. 27, 563-578.

Figure Legend

FIGURE 1. A view perpendicular to the crystallographic b axis of a difference Fourier map of DIP-trypsin using the coefficients $|F_{Cu^{++}}| - |F_{native}|$. The view shown is a projection through four sections and represents a total thickness of 4.7 Å. The scale markers in the lower right-hand corner represent 2 Å distances. Contours begin at 0.2 electrons per $Å^3$ and are drawn at successive intervals of 0.1 electrons per $Å^3$, using an approximate absolute scale (Stroud et al., 1974). Solid contours enclose regions of positive electron density; broken lines enclose regions of negative electron density. The remaining sections of the map along the b axis contained no peaks above the noise level of the map.

The two significant peaks correspond to the location of the copper ion binding site and to the movement of the imidazole ring of His 57 away from its native position in order to accommodate the ion. The imidazole movement peak occurs $2\frac{1}{2}$ sections (about 3 Å) above the copper peak, and the movement is in the direction of the solvent. The asterisks (*) mark the positions of the corresponding peaks in silver-DIP-trypsin (Chambers et al., 1974). Based on the approximate scale shown, the occupancy of the Cu^{++} site is about 50% in crystals of DIP-trypsin which have been soaked for six days in 17 mM Cu^{++} at pH 6.9.

Figure 1



CHAPTER IV: What is the pK_a of His 57?

Introduction

The assignment of the rate-controlling pK_a of 6.8 to Asp 102 in trypsin (Chapter II) leaves open the question of what the pK_a of His 57 might be. Attempts to measure this pK_a directly have been reported. Hunkapiller et al. (1973) used ^{13}C NMR spectroscopy to monitor the ionization state of the imidazole ring of His 57 in α-lytic protease, a bacterial homologue of the trypsin family. In that experiment the only histidine of α-lytic protease--His 57--was labeled with ^{13}C at the C-2 position by growing the bacteria on labeled histidine. Based on the ^{13}C—H coupling constant, the authors showed that His 57 was unprotonated above pH 4. At pH 3.3 and 34°C three NMR signals were seen and were assigned to: deprotonated imidazole in the native conformation, protonated imidazole in the native conformation, and protonated imidazole swung out into solution. The pK_a of His 57 in native α-lytic protease may be even lower than 3.3 because that enzyme--in contrast to trypsin or chymotrypsin--exhibits decreased solubility and stability at low pH (M. W. Hunkapiller, personal communication). The NMR signal which has characteristics similar to that of free histidine in solution may in fact be evidence of partial

denaturation at pH 3.3.

Support for the idea that the pK_a of His 57 is even lower than 3 comes from kinetic studies of trypsin and chymotrypsin. The pH-rate profiles can be explained by a single pK_a of 6.8 for both acylation and deacylation during hydrolysis of a number of substrates by trypsin and chymotrypsin in the pH range between 2 or 3 and 8 (Kezdy et al., 1964; Bender et al., 1964; Stewart and Dobson, 1965; Fersht and Renard, 1974). In all of these cases plots of log rate versus pH (Dixon, 1953) fall on straight lines of slope one in the pH range between 2 or 3 and 6. Between pH 6 and 7.5 the slope changes from one to zero, indicating the titration of the carboxyl of Asp 102.

Conflicting evidence regarding the pK_a of His 57 has been reported by Markley and Porubcan (1976), who used 250 MHz proton NMR spectroscopy to follow the resonances of the C-2 protons of the histidines of porcine trypsin as a function of pH. The chemical shift of the peak assigned to His 57 did not titrate in the pH range 6-9, but did undergo a titration with a midpoint of 5.0 which had half the normal amplitude. At pH 4.5 there was observed a discontinuity in the chemical shift of the peak assigned to His 57, apparently indicating a conformational transition between two states having lifetimes of approximately 18 msec. The data were interpreted in terms of either His 57 or Asp 102 titrating with a pK_a of 5.0 and the other

group having a pK_a of 4.5. One explanation advanced for the small amplitude of the chemical shift change for the titration at pH 5.0 was that perhaps there could be an equilibrium in the pH range 4.5-5.0 between two states: protonated His 57-negative Asp 102, and neutral His 57-neutral Asp 102; with the populations of the two states being roughly equal. Neither Asp 102 nor His 57 was assigned a pK_a of 6.8, for all observed titrations between pH 5 and pH 7 were assigned to the other three histidines of porcine trypsin. Nevertheless, all kinetic studies of the action of serine proteases show that both acylation and deacylation rates depend on a single group of pK_a 6.8.

The apparent conflict between the results of Markley and the kinetic data might be resolved if the ionization of His 57 were dependent on a slow pH-dependent conformational change which could not be detected in fast kinetic experiments. This question can be resolved by measuring the pH-dependence of the rate of substrate hydrolysis by trypsin which has been pre-incubated at various pH's.

This chapter is concerned with the possibility of a slow conformational change affecting the observed pK_a's of the active site residues of the serine proteases, as well as the general question of what the pK_a of His 57 is. The fact that only a single pK_a of 6.8 affects the rate of trypsin-catalyzed hydrolysis in the range of pH 3.0-7.0 will be reconfirmed; and the effect of pre-incubation of

trypsin at low or high pH on the rate of hydrolysis of a specific ester substrate will be presented. Finally will come a report of an attempt to directly measure the pK_a of His 57 by following the pH-dependence of the rates of exchange of the C-2 protons of the three histidines of bovine trypsin with tritium in tritiated water. These experiments were conducted jointly with Monty Krieger, and complete descriptions appear in Appendices C and D.

The Effect of Pre-Incubation on Trypsin Kinetics at Low pH

The rate of hydrolysis of N_α-carbobenzoxy-L-lysine-p-nitrophenyl ester (CLNE) by trypsin was measured over the pH range 2.0-5.0 using trypsin which had been pre-incubated for up to three hours at pH 6.9 or at pH 2.0. Complete details are given in Appendix C. Under the conditions used for this ester substrate, the deacylation of the N_α-carbobenzoxy-L-lysyl-trypsin was the rate-determining step of the reaction (Bender and Kezdy, 1965).

The deacylation rates were identical for trypsin pre-incubated at pH 2.0 or at pH 6.9 for up to three hours at 0°C. Since the mixing time before the reaction was a few seconds, this result excludes the possibility that the kinetics could be affected by any pH-dependent conformational change with a time constant between about fifteen

seconds and three hours. Since this time scale covers the range of times which are used for spectroscopic measurements of titrations, a conformational argument cannot be invoked to explain any discrepancy between kinetic and spectroscopic data. The methods should give consistent results.

The pH-dependence of the deacylation rate during hydrolysis of CLNE by trypsin between pH 3.0 and 5.0 is also available from the data of the pre-incubation experiment. A plot of log rate versus pH (Dixon, 1953) is linear between pH 3 and pH 5 and has a slope of approximately one, as would be expected if a single rate-controlling group had a pK_a anywhere above the pH 3.0-5.0 range. Since activity is known to depend upon a pK_a of 6.8, the Dixon plot referred to above provides evidence that no other group which affects the deacylation rate can have a pK_a between 3.0 and 6.8.

The pH dependence of the rate, the pre-incubation studies, and the assignment of a pK_a of 6.8 to Asp 102, taken together, indicate that the pK_a of His 57 in trypsin must be below 3. Any spectroscopic technique which measures a pK_a between 3.0 and 6.5 must be monitoring a group--or a fast conformational change--which has no effect on activity. The pH-rate profile for chymotrypsin further indicates that the pK_a of His 57 in chymotrypsin, and presumably in trypsin also, is below 2 (Kezdy et al., 1964; Fersht and Renard, 1974).

The pH Dependence of Tritium Exchange into the Histidine Side Chains of Trypsin

The Method

This section reports an attempt to determine the pK_a of His 57 in bovine trypsin. The technique used involves the rate of exchange of the C-2 proton of the imidazole ring in deuterated or tritiated water. A full account of the work appears in Appendix D.

The kinetics of deuteration of imidazole (Vaughan et al., 1970) and of N-acetyl-L-histidine (Matsuo et al., 1972) in aqueous solution have been studied as a function of pH. The pH dependence of the exchange rate can be explained by a mechanism which involves rate-determining abstraction of the C-2 proton by OH^- or by H_2O to form an ylide intermediate, followed by a fast protonation of the ylide (Vaughan et al., 1970). Evidence that proton abstraction is indeed involved in the rate-determining step is provided by the deuterium isotope effect of 7.5 which has been observed when comparing the rates of exchange of tritium into the tripeptide Gly-His-Gly with ^1H or ^2H in the C-2 position of the imidazole (Markley and Cheung, 1973). The pK_a of the imidazole ring can be determined by fitting the pH-exchange data to the rate equation derived from the ylide

mechanism.

The rate of exchange at the C-2 of imidazole is intermediate between the rates for fast-exchanging O—H and N—H protons and non-exchanging C—H protons. The uniqueness of this rate makes possible the specific labeling of the histidine rings of a protein in tritiated water (Matsuo et al., 1972). The tritium can be incorporated under mild (non-denaturing) conditions (37°C for two days); and the labile protons can be back-exchanged, thus leaving only histidines labeled. Studies of the pH dependence of tritium incorporation into the histidines of lysozyme (Matsuo et al., 1972) and ribonuclease (Ohe et al., 1974) have been used to determine the pK_a of each histidine in these proteins.

Our experiments with trypsin involved incubating the enzyme at 37°C in buffers of various pH ranging from 2 to 9, containing tritiated water. The incubation time was either 38 hours or 14 days; autolytic digestion during these long periods was prevented by including the inhibitor benzamidine in the solutions. Following the incubations, each labeled protein sample was separated from the bulk tritiated water by gel permeation chromatography and lyophilization. The three histidines of bovine trypsin were separated from each other by enzymatic digestion and subsequent isolation of pure peptides by means of chromatography and electrophoresis on paper (Bennett,

1967).

The peptides were isolated under conditions whereby the tritium attached to the C-2 carbon of histidine did not undergo back-exchange with the solvent. Nevertheless, all other tritions were eliminated from the samples by the back-exchange which took place during the purification procedures. That only the histidines remained labeled was shown by the absence of radioactivity in the peptides which did not contain histidine. The specificity of the incorporated label for the C-2 position in the ring was demonstrated by a separate study of the deuteration of L-histidine using NMR spectroscopy.

The specific activity of the histidine in each peptide was determined by scintillation counting and amino acid analysis. The time of incubation, and the specific activities of the incubation water and histidine samples, were used to calculate half-times and pseudo-first-order rate constants for the exchange reaction.

Results and Discussion

The observed rates of tritium-hydrogen exchange are extremely slow for all of the three histidines of bovine trypsin. In fact, the observed levels of tritium incorporation were so low that careful controls had to be performed to make sure that active trypsin--and not

just a small impurity of denatured trypsin--was actually exchanging at all. Two of the histidines, His 40 and His 91 in the sequence, exhibit half-times for exchange of greater than 1000 days at pH 9, a period which is indeed long when compared with the 1.5 day half-time for free histidine in solution (Matsuo et al., 1972). The very slow exchange rate for His 40 is consistent with the buried position of the imidazole ring in the tertiary structure of trypsin, which renders the ring totally inaccessible to solvent (Stroud et al., 1971; Chambers and Stroud, unpublished). In the case of His 91 the slow rate of exchange is not so easy to rationalize, for the side chain is quite near the surface of the protein, and the ring is not obviously hindered from interaction with solvent. An adequate explanation for the slowness of hydrogen-tritium exchange at His 91 has yet to be presented.

Although the C-2 proton of His 57, the active site histidine, exchanges more readily than those of His 40 or His 91, the rate of exchange is still very slow when compared to the rates for histidines in other proteins. The slowest rate reported for a protein other than trypsin is the case of His 48 of ribonuclease, which exchanges with a half-time of 58 days at 38°C and pH 9 (Ohe et al., 1974). The other histidines of ribonuclease and lysozyme exchange much more readily (Ohe et al., 1974; Matsuo et al., 1972). His 57 of trypsin, on the other hand, exhibits a half-time of 73 days for hydrogen-tritium

exchange at 38°C and pH 9. The slowness of this rate should be taken into account when interpreting the pH dependence of the rate.

The data for the pH dependence of tritium exchange into His 57 of trypsin can be fit to the ylide mechanism (Vaughan et al., 1970) with a pK_a of 6.55 for the imidazole ring. Although this mechanism is generally accepted, the fact that it fits the data does not prove that it is correct or that an alternative mechanism would not fit as well. Of particular interest are possible mechanisms in which the ionization of a neighboring group affects the rate of tritium exchange of the imidazole. There is at least one such alternate mechanism which would predict the same rate equation and the same pH-rate profile as the ylide mechanism, and would have its "primary" pH dependence due to the pK_a of Asp 102 (see Appendix D). However, the alternate mechanism predicts a different rate-determining step, which is not compatible with the deuterium isotope effect for tritium exchange into Gly-His-Gly (Markley and Cheung, 1973). Nevertheless, the situation in the enzyme coule possibly be different.

So far only His 57 in the active conformation of trypsin has been considered as the species undergoing tritium-hydrogen exchange at the C-2 carbon. However, the unusual slowness of the rate of exchange and the three-dimensional structure of trypsin (Stroud et al., 1974) indicate another possible explanation. The imidazole ring of

His 57 is half buried in trypsin, so that although the C-4 carbon and N-3 nitrogen are accessible to solvent, the C-2 carbon and N-1 nitrogen are sequestered away from solvent. The ring is sandwiched from above and below by other parts of the structure. The N-1 nitrogen is hydrogen-bonded to Asp 102 and the C-2 carbon points toward the center of the molecule. One explanation for the slow exchange rate is that the C-2 proton of His 57 simply does not exchange in the active conformation since it is inaccessible to solvent, but that the imidazole ring occasionally swings out from its normal position and exposes the C-2 proton to the solvent, at which point exchange with tritium can occur. If this "breathing" did occur, the pK_a of 6.55 would refer to the imidazole ring out in solution--not in its active conformation. The results could be satisfactorily explained by an "out" form of His 57 which has a pK_a of 6.55 and undergoes exchange at C-2 at a rate typical for histidine in solution or in other proteins, but which is populated only a small percentage of the time. His 57 in the active conformation might exchange at a rate comparable to those for His 40 and His 91, a possibility which seems reasonable from the structure.

The tritium exchange experiment takes such a long time that it cannot distinguish "in-out" swinging of the histidine from slow exchange of the "in" form. The method could trap a conformational

state in which a molecule existed only a very small amount of the time. Therefore, several interpretations are possible.

NMR and infrared techniques have provided evidence that Asp 102 is the group at the active site of the serine proteases which has a pK_a near 6.7 (Hunkapiller et al., 1973; Chapter II). The pH dependence of the kinetics of substrate hydrolysis by trypsin and chymotrypsin shows that only the single pK_a of 6.7 affects the rate of hydrolysis between pH 2 and pH 8 (Kezdy et al., 1964; Bender et al., 1964; Stewart and Dobson, 1965; Fersht and Renard, 1974; Appendix C). This would lead to the conclusion that His 57 of trypsin has a pK_a below 2 in the active conformation (the active conformation is not stable much below pH 2). The tritium exchange data would be consistent with this conclusion if histidine in the active form had a low pK_a and did not exchange, but had a pK_a of 6.55 when swung out into solution in the exchangeable form. Unlike His 91 or His 40, the imidazole ring of His 57 in a model of trypsin can be rotated out into solution without much disruption of the rest of the structure. This interpretation seems to be the most reasonable one which is consistent with all of the data because the "in" conformation of His 57 appears completely inaccessible to proton exchange at C-2.

References

Bender, M. L. and Kezdy, F. J. (1965), J. Amer. Chem. Soc. 87, 4954-4955.

Bender, M. L., Killheffer, J. V. Jr., and Kezdy, F. J. (1964), J. Amer. Chem. Soc. 86, 5330-5331.

Bennett, J. (1967), Methods in Enzymol. 11, 330-339.

Dixon, M. (1953), Biochem. J. 55, 161-170.

Fersht, A. R. and Renard, M. (1974), Biochemistry 13, 1416-1426.

Hunkapiller, M W., Smallcombe, S. H., Whitaker, D. R., and Richards, J. H. (1973), Biochemistry 12, 4732-4743.

Kezdy, F. J., Clement, G. E., and Bender, M. L. (1964), J. Amer. Chem. Soc. 86, 3690-3696.

Markley, J. L. and Cheung, S.-M., Proceedings of the International Conference on Stable Isotopes in Chemistry, Biology, and Medicine, Argonne, Illinois, May 1973. U. S. Atomic Energy Commission Conference, No. 730525, pp. 103-118.

Markley, J. L. and Porubcan, M. A. (1976), private communication of manuscript submitted to J. Mol. Biol.

Matsuo, H., Ohe, M., Sakiyama, F., and Narita, K. (1972), J. Biochem. 72, 1057-1060.

Ohe, M., Matsuo, H., Sakiyama, F., and Narita, K. (1974),

J. Biochem. 75, 1197-1200.

Stewart, J.A. and Dobson, J.E. (1965), Biochemistry 4, 1086-1091.

Stroud, R.M., Kay, L.M., and Dickerson, R.E. (1971), Cold Spring Harbor Symp. Quant. Biol. 36, 125-140.

Stroud, R.M., Kay, L.M., and Dickerson, R.E. (1974), J. Mol. Biol. 83, 185-208.

Vaughan, J.D., Mughrabi, Z., and Wu, E.C. (1970), J. Org. Chem. 35, 1141-1145.

APPENDIX A

A Pulsed Diffusion Technique for the Growth of Protein Crystals for X-ray Diffraction

R. E. KOEPPE II, R. M. STROUD

Norman W. Church Laboratory of Chemical Biology
California Institute of Technology
Pasadena, Calif. 91125, U.S.A.

V. A. PENA AND D. V. SANTI

Departments of Biochemistry and Biophysics
and of Pharmaceutical Chemistry
University of California
San Francisco, Calif. 94143, U.S.A.

Reprinted from *J. Mol. Biol.* (1975) **98**, 155–160

Crystals of proteins suitable for X-ray diffraction have been grown using a method that combines dialysis (Zeppezauer *et al.*, 1968) and relaxation (Salemme, 1972) techniques. This pulsed-diffusion method is useful for (a) eliminating unwanted amorphous precipitate, and (b) reducing the number of crystal nuclei which form and thus growing each individual crystal larger. The method is illustrated for glutamate–aspartate transaminase and thymidylate synthetase. Preliminary crystallographic data are reported for crystals of thymidylate synthetase.

1. Introduction

Prerequisite to the determination of the tertiary structure of an enzyme using X-ray diffraction is the growth of crystals of the enzyme. Unfortunately, many interesting and important enzymes cannot be studied because of the lack of suitable crystals. We report a technique that has proven useful for crystallizing enzymes that are available in only small quantities. Applications to crystallization of glutamate–aspartate transaminase from pig heart, and thymidylate synthetase from *Lactobacillus casei* are discussed.

Several methods for growing protein crystals have been described in the literature. In most cases such methods use a slow approach toward some predetermined precipitant concentration. Two micro-diffusion techniques have been described by Zeppezauer *et al.* (1968) and Salemme (1972). The Zeppezauer method involves holding the protein concentration constant inside a capillary closed by a dialysis membrane, and increasing the salt or organic solvent concentration outside the membrane until the protein becomes insoluble and crystallizes. Figure 1(a) illustrates that the precipitant concentration may be raised in one large increment or by a series of small steps, as first suggested by Zeppezauer (1968). In either case, supersaturation of the protein is approached in a slow exponential manner.

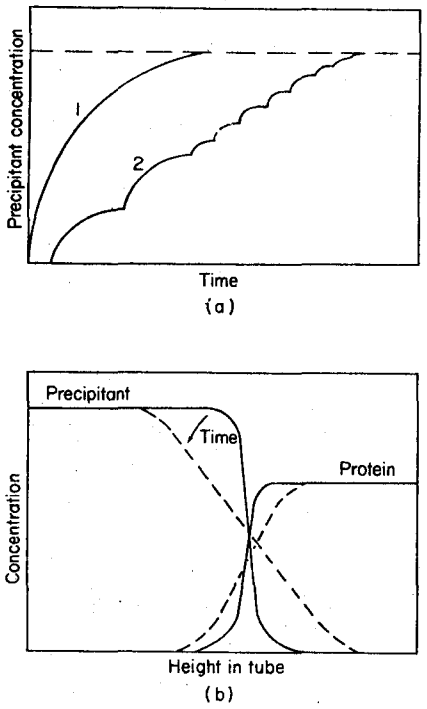


FIG. 1. (a) Schematic diagram illustrating the approach to protein saturation in a dialysis experiment described by Zeppezauer *et al.* (1968). (– – – –), Precipitant concentration at which the protein begins to crystallize. This concentration may be approached in a single step of dialysis (curve 1) or by gradually changing the precipitant concentration outside the dialysis capillary in a series of small steps (curve 2). (b) Schematic diagram showing the relaxation from transient supersaturation in a free diffusion experiment described by Salemme (1972). (————), The distributions of protein and precipitant in the tube near the beginning of the experiment. Nucleation occurs at the interface, where a high precipitant concentration meets a high protein concentration. (– – – –), The system at a later time, when the steep gradients of precipitant and protein have diminished and the condition of nucleation has changed into one favoring growth of existing nuclei.

Salemme's (1972) free diffusion method is fundamentally different in that a relaxation of the protein and precipitant concentrations from a transient condition of supersaturation is produced. Figure 1(b) shows the situation for protein solution of lower density layered above a salt solution of higher density. The protein at the interface experiences its highest salt concentration at the beginning of the experiment (solid curves in Fig. 1(b)), and the solution becomes nucleated with seed crystals. As time progresses (dashed curves) both protein and salt concentrations decrease so that small nuclei dissolve, but more protein comes into contact with salt and the large nuclei grow larger.

The method we have developed combines favorable features from both of the above methods, namely: (a) the maintenance of a constant and well-defined protein concentration; (b) the ability to carefully define conditions by dialysis; (c) a time-scale that can be readily varied and controlled (Zeppezauer *et al.*, 1968); and (d) a relaxation from supersaturation (Salemme, 1972). In addition, our method uses a cycling of precipitation conditions, a feature that has been exploited for small molecule crystal growth, particularly in growing crystals from melts (Samson, 1967).

Using a cell of the Zeppezauer design, a protein solution is dialyzed against a concentration of precipitant which is more than sufficient to bring the protein out of solution. A high gradient of precipitant across the dialysis membrane, and indeed throughout the entire capillary, serves to generate crystal nuclei. Before complete equilibration has taken place, the cell is transferred to a less concentrated precipitant solution so as to dissolve the precipitated protein. This method takes advantage of the observation that protein precipitate will generally redissolve more easily than protein crystals, and provides a means of forming seed crystals. The crystals may be grown larger by pulsed diffusion between alternately high and low concentrations of precipitant. Once crystal nuclei have formed, the protein will preferentially crystallize on them, rather than precipitate. Thus the method may be useful in reducing the number of crystals in one cell, and in growing the smaller number of crystals larger, while eliminating the amount of precipitated (non-crystalline) protein. Concentrations of precipitant used in this pulsed-diffusion method can be closely defined to regulate number and size of crystals and total precipitate. Two applications of the method are discussed.

2. Experimental

2-Methyl-2,4-pentanediol (Aldrich; 99%) was purified by stirring overnight with charcoal. Enzyme-grade ammonium sulfate was obtained from Schwarz–Mann. α-Ketoglutarate was obtained from Calbiochem. Cytoplasmic glutamate–aspartate transaminase from pig heart was the generous gift of Dr M. Martinez-Carrion, University of Notre Dame, Indiana.

Thymidylate synthetase was purified from amethopterin-resistant *L. casei* by two of us (Santi and Pena; see Santi *et al.*, 1974).

Crystallization cells closed by dialysis membrane were prepared as described by Zeppezauer *et al.* (1968), from thick-walled glass capillary tubes, 1 mm internal diam. and 20 mm long. The cells were placed horizontally on the bottoms of beakers during the dialysis. The rate of dialysis of ammonium sulfate into membrane-closed capillaries was determined by conductivity measurements. In general, about 16 h at room temperature were required for the average concentration inside the cell to reach half the concentration outside the dialysis cell ($t_{\frac{1}{2}}$). However, average concentrations do not reflect the steep gradient of ammonium sulfate inside the capillary itself, a gradient which probably contributes to the success of this method. Crystals have been observed near the membrane much earlier than in the other part of the capillary. Thus at $t_{\frac{1}{2}} = 16$ h, the concentration near the membrane was shown to be relatively close to that outside, while the concentration in the remote part of the capillary was still low.

3. Results

(a) *Glutamate–aspartate transaminase*

Initial trial experiments were made to determine (a) the precipitant concentration necessary to bring protein out of solution, (b) the precipitant concentration above which crystals were stable, and (c) the stability criteria for precipitated protein.

When this enzyme was dialyzed against 28% 2-methyl-2,4-pentanediol containing 10 mM-phosphate and 20 mM-α-ketoglutarate (pH 7·5, room temperature), a precipitate appeared after 1 to 2 weeks. Thus, 28% MPD† was identified as the precipitant concentration necessary to bring protein out of solutions (Fig. 2). Higher concentrations produced the precipitate faster and occasionally some tiny crystalline nuclei

† Abbreviation used: MPD, 2-methyl-2,4-pentanediol.

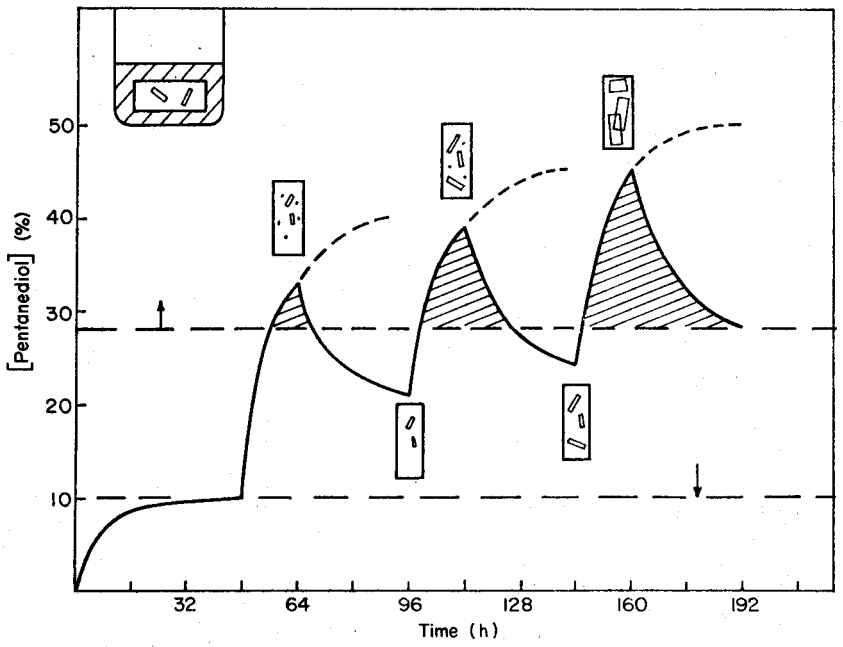


FIG. 2. Diagram of the pulsed-diffusion technique applied to glutamate–aspartate transaminase. Crystals were produced at room temperature using 10 mM-phosphate buffer (pH 7·5), 20 mM-α-ketoglutarate and varying concentrations of MPD. Above 28% MPD (arrow) the protein came out of solution either as crystals or as precipitate. The precipitate readily dissolved below 28% MPD, but the crystals were stable for several weeks at concentrations as low as 10% MPD (arrow).

could be seen along with the precipitate. Precipitated protein could be dissolved by lowering the concentration below 28%, although the crystal nuclei did not dissolve for several weeks at concentrations of MPD above 10%.

The above observations were used to design a pulsed-diffusion experiment that gave considerably improved results (Fig. 2). Enzyme solution (20 mg/ml) was put into Zeppezauer capillaries (1 mm internal diam. by 20 mm long), which were placed on their sides (horizontally) in a beaker for two days of dialysis against 10% MPD containing 10 mM-phosphate and 20 mM-α-ketoglutarate (pH 7·5, room temperature). A vertical orientation of the capillaries did not work as well, for the aggregates of protein tended to fall to one end of the tube. The initial time lapse permitted the pH and cofactor to equilibrate with the enzyme, and also served to decrease the steepness of the precipitant gradient in the next dialysis step. The capillary cell was then transferred to a solution containing 40% MPD (all other conditions unchanged), but was left there for only 16 hours, enough time to pass through the critical concentration (28% MPD) to initiate precipitation in the end of the cell closest to the membrane, but not long enough to actually reach 40%. The cell was then returned to 10% MPD to dissolve the precipitate that had formed. At this point, two or three small crystals (up to 0·05 mm long needles) were present in the capillary. Two more successive 16-hour pulses in 45% MPD produced larger crystals and occasionally new small crystals without much, if any, new precipitate. Finally, the capillary was left in 40% MPD, where the needles grew to their maximum size of 0·5 mm long within one week.

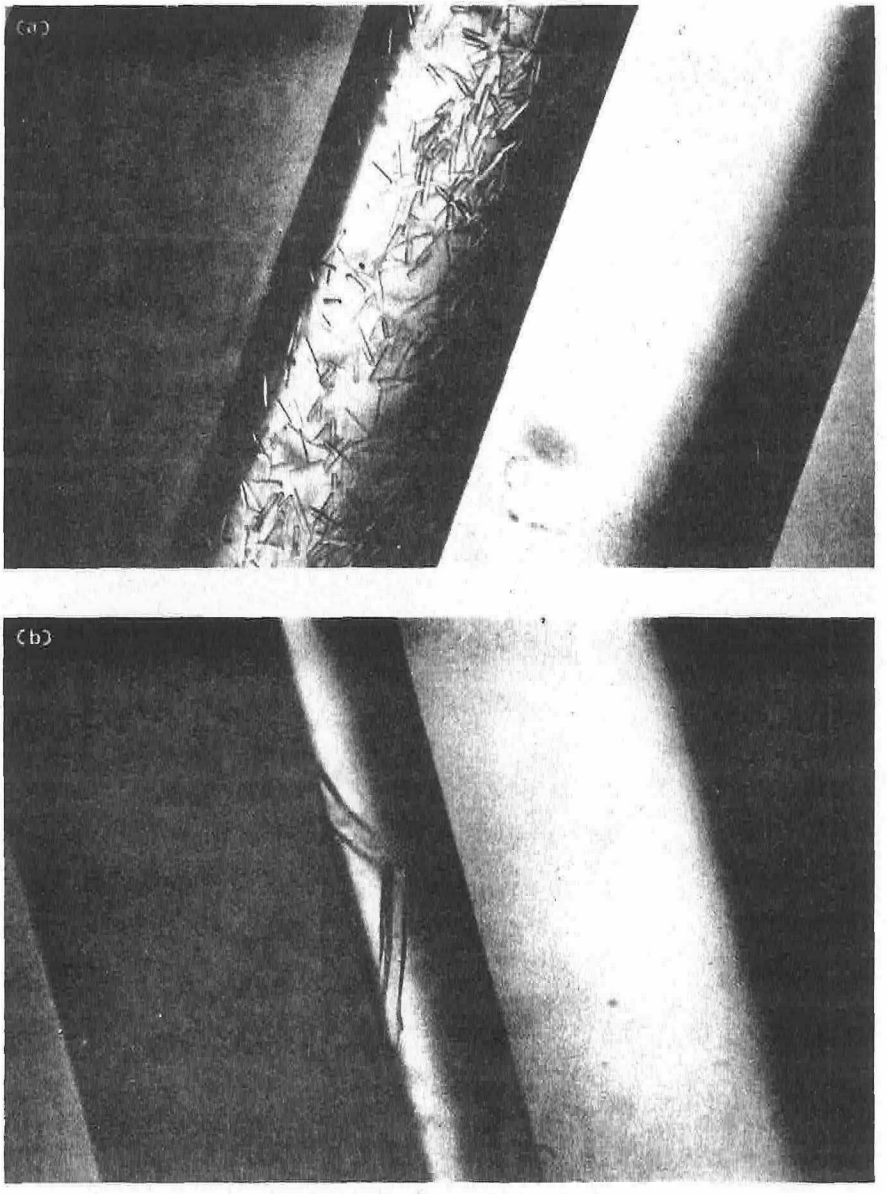


PLATE I. Crystals of thymidylate synthetase grown (a) by the conventional technique of Zeppezauer (1968) in a 2 mm internal diam. capillary, and (b) by pulsed-diffusion in a 1 mm capillary. Conditions for crystallization are described in the text.

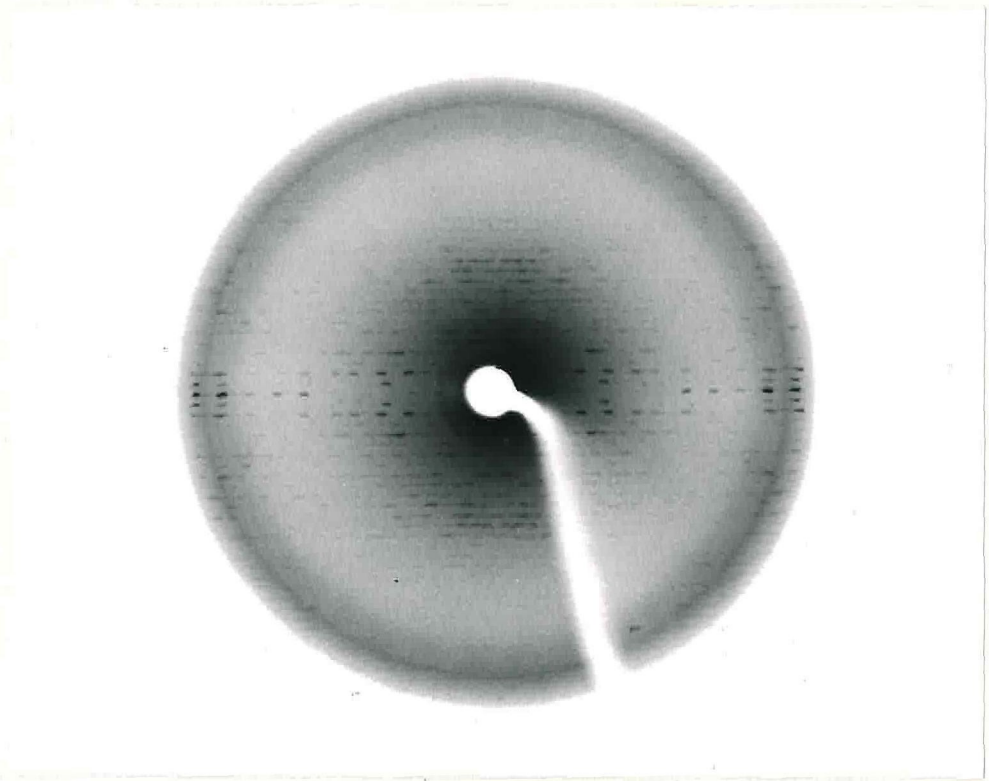


PLATE II. 10° precession photograph of a thymidy late synthetase crystal shown in Plate I(b).

The concentration changes and crystal or precipitate formation at each stage of this application are indicated in Figure 2. The curves represent MPD concentration in the half of the capillary near the membrane, and are estimated from conductivity measurements, which determined average $(NH_4)_2SO_4$ concentrations throughout cells, and from visual observation of protein precipitate at various times. Best results have been obtained when the capillary was kept horizontal in the beaker. In this case gravity is not a factor, since precipitated protein does not accumulate on the membrane surface.

(b) *Thymidylate synthetase*

Another application of the pulsed-diffusion method was made to the crystallization of thymidylate synthetase. Crystals were grown from 0·7% protein solution at 4°C using 0·8 to 1·1 M-ammonium sulfate and 50 mM-Tris buffer (pH 7·2). A one-step dialysis approach produced crystals too small for X-ray work (generally smaller than 0·05 mm in diameter). In this case there was no significant precipitate formed and use of the pulsed-diffusion method resulted in fewer and larger crystals in each capillary (Plate I). The crystals grow as hexagonal needles up to 2 mm long and 0·2 mm in diameter.

The 10° zero layer precession photograph of these crystals shown in Plate II was taken using nickel-filtered CuKα radiation from an Elliott rotating anode generator. The space group is $P6_122$ (or $P6_522$) and the spacings in Plate II correspond to dimensions of $a = 152$ Å and $c = 298$ Å. However, patterns have been obtained from other crystals with halving of the a axis to $a = 76$ Å, a change which decreases the unit cell volume by a factor of four. The crystal density is 1·26 g cm^{-3}. The active form of the enzyme consists of two identical subunits each of molecular weight 35,000 (Santi *et al.*, 1974). The smaller unit cell contains one dimer per asymmetric unit, has 36% solvent, and a crystal volume per unit protein molecular weight of 1·78 (Matthews, 1968). The latter crystal form is clearly more suitable for a detailed structure analysis and efforts are now directed toward preferentially stabilizing this form.

4. Discussion

The pulsed-diffusion technique has been shown to be useful for growing improved crystals of the above two enzymes. In the case of glutamate–aspartate transaminase, the method overcame the problem of amorphous precipitate which interfered with crystal formation and growth. With thymidylate synthetase, precipitate was not a problem, but the large number of microcrystals produced by conventional procedures was transformed into a few large crystals by use of the pulsed-diffusion method.

The success of pulsed-diffusion is attributed to the ability to nucleate the protein solution in an early supersaturating condition (Salemme, 1972), while maintaining accurate control of the crystallizing conditions. The time of each step in the experiment is controlled by the time constant of diffusion and by the concentration step sizes used in dialysis. High gradients of precipitant concentration, both increasing and decreasing, are used in what one may call the early "kinetic" stages of dialysis. Then the single crystals are grown to maximum size by dialysis to "equilibrium". The pulsed-diffusion technique may be of use in other situations where protein precipitate or the large number of nuclei is a problem.

We thank Dr Marino Martinez-Carrion, University of Notre Dame, for his gift of the cytoplasmic glutamate–aspartate transaminase. This work was carried out with the support of the National Institutes of Health (grant no. GM-19984 to R. M. S. and no. CA-14394 to D. V. S.), whose help is gratefully acknowledged. One of us (R. M. S.) is the recipient of a National Institutes of Health career development award, and another (R. E. K.) was the recipient of a National Science Foundation predoctoral fellowship and presently holds a National Institutes of Health predoctoral traineeship. This is contribution no. 5088 from the Norman W. Church Laboratory of Chemical Biology, California Institute of Technology.

REFERENCES

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Salemme, F. R. (1972). *Arch. Biochem. Biophys.* **151**, 533–539.
Samson, S. (1967). *Acta Crystallogr. sect. A*, **23**, 586–600.
Santi, D. V., McHenry, C. S. & Sommer, H. (1974). *Biochemistry*, **13**, 471–480.
Zeppezauer, M., Eklund, H. & Zeppezauer, E. S. (1968). *Arch. Biochem. Biophys.* **126**, 564–573.

APPENDIX B

Reprinted from *Proteases and Biological Control*, pp. 13–32

Structure-Function Relationships in the Serine Proteases

Robert M. Stroud, Monty Krieger, Roger E. Koeppe II, Anthony A. Kossiakoff and John L. Chambers

Norman W. Church Laboratory of Chemical Biology
California Institute of Technology, Pasadena, California 91109

Of the many ways available to control the biological activity of proteins, e.g., induction or repression of their synthesis at the translational (Jacob and Monod 1961) or transcriptional levels (Tomkins et al. 1969), specific modification or destruction are the most direct. Many biological systems are controlled by methods such as these, and the serine protease family of enzymes plays a major role in many of these systems (Stroud 1974). The pancreatic serine proteases are digestive enzymes which show optimal activity around the neutral pH region. Their function in hydrolyzing peptide bonds and the systems of physiological control over their activity have close homology in many other biological processes, e.g., blood clotting (Owren and Stormorken 1973; Magnusson 1971), bacterial sporulation (Leighton et al. 1973), fertilization (Stambaugh, Brackett and Mastroianni 1969), etc. Many of the enzymes of biological control have been recognized as serine proteases, which in nearly all documented cases have amino acid sequence homology to the pancreatic serine proteases. It is therefore to be expected that these enzymes will have closely homologous tertiary structures and will share the same catalytic mechanism of action. The mechanisms by which such enzymes are activated or inhibited will also share many common features with the digestive serine proteases. In many cases the degrees to which these principles can be extended may be predicted by recognition of the chemical and structural features of the pancreatic serine proteases which appear to define their properties. In this article we will discuss recent advances in the understanding of aspects of the structures and functions of the mammalian serine proteases.

THE PANCREATIC DIGESTIVE ENZYMES: TRYPSIN, CHYMOTRYPSIN, ELASTASE

Intrinsic to the process of digestion in mammals is the breakdown of dietary protein by the pancreatic serine proteases. These pancreatic digestive enzymes

are among the most thoroughly studied of all enzymes, principally because they are extracellular enzymes that are easily separated and purified in large quantities (Kunitz and Northrop 1935). They originate in the pancreas as inactive precursors, or proenzymes, which are secreted into the duodenum. There they are activated (Kunitz and Northrop 1936; Northrop, Kunitz and Herriot 1948; Maroux, Baratti and Desnuelle 1971) by the cleavage of one critical peptide bond near the amino-terminal end of the polypeptide chain (Davie and Neurath 1955). This cleavage in turn permits a conformational change (Neurath and Dixon 1957; Sigler et al. 1968) which converts the proenzymes to active enzymes. Once activated, these enzymes catalyze the breakdown of proteins, first into fragments and ultimately into individual amino acids.

Kinetic studies on a variety of amide and ester substrates have shown that the mechanism of serine protease catalysis (Eq. 1) involves a number of intermediates (Zerner, Bond and Bender 1964; Oppenheimer, Labouesse and Hess 1966; Caplow 1969; Hess et al. 1970; Fersht and Requena 1971a; Fastrez and Fersht 1973a,b; Fersht and Renard 1974):

$$E + S \rightleftharpoons ES \rightleftharpoons ES^{T} \rightleftharpoons \begin{matrix} EP_2^{a} \\ + \\ P_1 \end{matrix} \rightleftharpoons EP_2^{T} \rightleftharpoons EP_2 \rightleftharpoons E + P_2. \qquad (1)$$

Here, E represents free enzyme; S, the substrate; ES and EP_2, enzyme substrate and product complexes; ES^T and EP_2^T, tetrahedral intermediates; EP_2^a, an acyl enzyme; P_1, the amino or alcohol portion of the product; and P_2, the carboxylic acid portion of the product. For amides, the rate-determining step is generally acylation, $E + S \rightarrow EP_2^a$, whereas deacylation, $EP_2^a \rightarrow E + P_2$, is usually rate-determining for esters (Zerner and Bender 1964). The characteristic differences between each of the digestive serine proteases—trypsin, chymotrypsin and elastase—lie in their specificity for hydrolyzing the peptide bonds between different amino acids in the protein substrate. Trypsin, the most sharply specific of the digestive enzymes, hydrolyzes those peptide bonds that immediately follow either of the two basic amino acids, lysine or arginine. Chymotrypsin hydrolyzes peptide bonds that follow several of the amino acids with larger hydrophobic side chains, and elastase binds the small side chains of glycine, alanine or serine at the equivalent binding site (Naughton and Sanger 1961; Brown, Kauffman and Hartley 1967; Sampath Narayanan and Anwar 1969). The complete amino acid sequences and three-dimensional molecular structures have now been worked out for chymotrypsin (Sigler et al. 1968) and its proenzyme (Freer et al. 1970), elastase (Shotton and Watson 1970), and DIP (diisopropylphosphoryl)-trypsin (Stroud, Kay and Dickerson 1971, 1974) and the proenzyme (Kossiakoff, Kay and Stroud, unpubl.). These structures, along with that of the bacterial serine protease subtilisin (Wright, Alden and Kraut 1969; Alden, Wright and Kraut 1970), have been uniquely valuable in developing an understanding of how these enzymes bind a substrate and how they catalyze the subsequent chemical reaction.

Activating the Proenzyme

The first key to activation of the pancreatic proenzymes is enterokinase, an enzyme secreted in small amounts by the mucous membrane of the stomach. Its prime function is to convert some trypsinogen to trypsin, which then activates all of the proenzymes (including more trypsinogen) (Kunitz and Northrop 1936; Northrop, Kunitz and Herriot 1948; Maroux, Baratti and Desnuelle 1971). In each case, activation involves the cleavage of a few amino acid residues from the amino-terminal end of the proenzyme (Davie and Neurath 1955; Neurath and Dixon 1957).

With the formation of the new amino terminus at Ile-16[1] (Oppenheimer, Labouesse and Hess 1966), the protein undergoes conformational changes (Neurath, Rupley and Dreyer 1956) leading to a catalytically active configuration. A comparison of the high-resolution structure of chymotrypsinogen with that of chymotrypsin (Freer et al. 1970; Wright 1973) and of the high-resolution structure of DIP-trypsin with the recently determined high-resolution structure of trypsinogen (Kossiakoff, Kay and Stroud, unpubl.) helps us to understand the exact nature of these conformational changes. (A detailed description of the trypsinogin structure will be published later.) In both cases,

Figure 1
A comparison of the 5-Å models of trypsinogen (*left*) and DIP-trypsin (*right*) shows close structural homology in most areas of the molecule. Striking structural differences are observed in only two areas of the molecule. The first is in the binding pocket region, which is formed by residues 214–220 and 189–192, and the second is along a loop of chain containing residues 140–151 located on the right-hand side of the molecule. The small difference in size of the models is due to a difference in scale and orientation.

[1] The numbering system referred to is that of chymotrypsin, which will be adopted here as a standard for comparison of sequences.

clipping the proenzyme tail permits the new, positively charged α-amino terminus at residue 16 to fold into the interior of the globular structure and form an ion pair with the negative carboxyl group of Asp-194 (Matthews et al. 1967; Sigler et al. 1968). While this change is accompanied by movements in the region of the specificity binding pocket, there appears to be little change in the interaction between Asp-102, His-57 and Ser-195 at the catalytic site (Freer et al. 1970; see also Fig. 1). Thus the arrangement of these catalytic residues is preformed in the proenzyme. One major factor which contributes to the relative inactivity of the zymogen is that the binding of the normal substrates is impaired (Kassell and Kay 1973; Gertler, Walsh and Neurath 1974).

Enzyme Specificity and Substrate Binding

The serine proteases differ in their specificities because of differences in their substrate binding sites. Trypsin, chymotrypsin and elastase all have specific side-chain binding pockets on the surface of the protein close to the catalytic site (see Fig. 2). This pocket is lined by residues 214–220 and 189–192 and

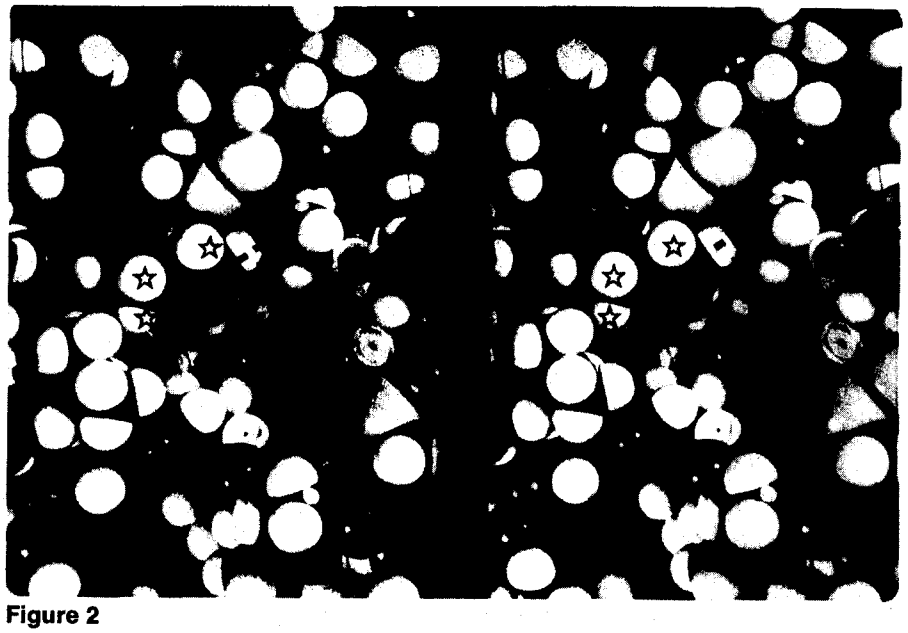


Figure 2

Stereoscopic photograph of a space-filling model showing the active site and specificity pocket in trypsin (see Stroud, Kay and Dickerson 1971, 1974). The imidazole side chain of His-57 is visible; however, the carboxylic acid side chain of Asp-102 is hidden from view by several amino acid residues. The viewing direction is approximately the same as that of Figures 4 and 7. The side-chain binding pocket is located beneath and to the right of the catalytic site.

Stars are placed near the active site as markers. They are located (reading from left to right) at the following positions: the Asp-102/His-57 hydrogen bond, the β-carbon protons of His-57, one of the ring protons on His-57, and the Ser-195 γ-hydroxyl.

defines the primary specificity toward substrate side chains immediately prior to the peptide bond which is to be cleaved. Cysteine residues 220 and 191 are linked by a disulfide bond. In trypsin, residue 189 is an aspartic acid, and its negatively charged carboxyl group (pK_a = 4.6) (East and Trowbridge 1968) lies at the bottom of the pocket (Stroud, Kay and Dickerson 1974). Trypsin has primary specificity for basic amino acids because their positively charged side chains bind tightly in this pocket (Mares-Guia and Shaw 1965; Ruhlmann et al. 1973; Blow, Janin and Sweet 1974; Sweet et al. 1974; Krieger, Kay and Stroud 1974). In an attempt to determine the manner in which amino acid side chains bind, we determined the structure of benzamidine trypsin. Benzamidine is a competitive, specific and reversible inhibitor of trypsin. Figure 3 shows how benzamidine, an amino acid side chain analog, binds in the specificity binding pocket (Krieger, Kay and Stroud 1974). In the case of trypsin, there is evidence that when side chains are bound in this pocket, they induce small conformational changes in the enzyme-substrate complex which help to accelerate catalysis (Inagami and Murachi 1964; Inagami and York 1968). In chymotrypsin, residue 189 is a serine (Hartley 1964). The pocket is now relatively hydrophobic and uncharged at neutral pH's, thus explaining chymotrypsin's specificity. In both trypsin and chymotrypsin, residue 216, a glycine, lies at the entrance to the binding pocket. In elastase, valine replaces glycine at position 216 (Shotton and Hartley 1970). The larger hydrophobic side chain blocks the entrance to the pocket and only allows the binding of amino acids with small side chains at the primary binding site (Shotton and Watson 1970).

Other parts of the enzyme are involved in binding other parts of the substrate molecule as well as the side chain, so that the susceptible substrate bond is aligned appropriately on the surface. Secondary specificity toward other

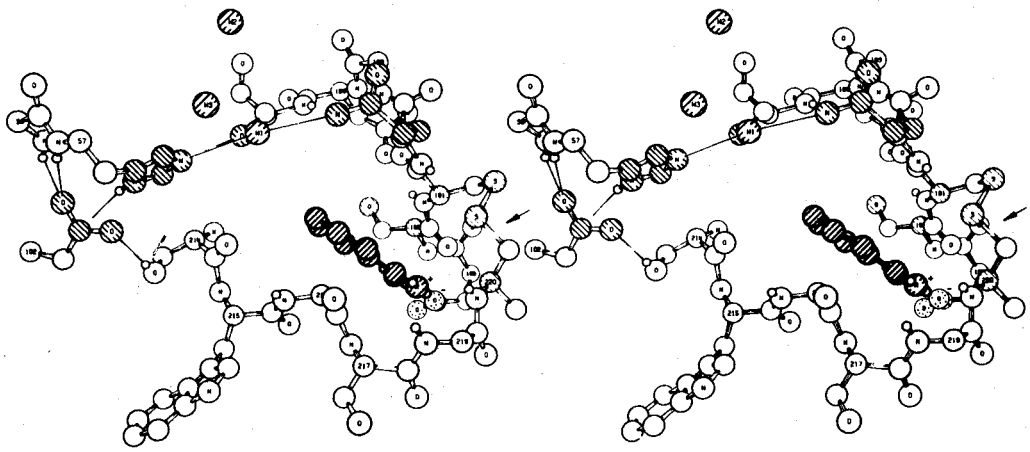


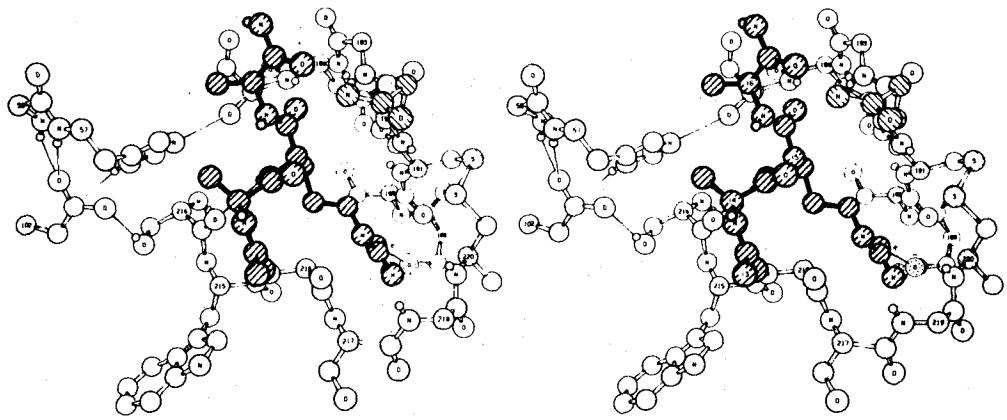
Figure 3

The figure shows the structure of the trypsin binding pocket in benzamidine-trypsin. (The phenyl amidinium is depicted by heavy shading.) This view is approximately the same as that of Figure 2. (Reprinted, with permission, from Krieger, Kay and Stroud 1974.)

side chains in the physiological substrate molecules can also be correlated with enzyme structure (Fersht, Blow and Fastrez 1973), although the role of secondary specificity in the digestive enzymes is clearly much less significant than it is in highly specific enzymes of biological control.

Ideally, one would like to study the three-dimensional structure of an enzyme-substrate complex by X-ray crystallography and in so doing, gain new insights into the mechanism of serine protease catalysis. Unfortunately, this has not yet been possible because the catalyzed reaction takes place almost immediately after the substrate is bound and the system becomes an enzyme-plus-product complex. Data for a three-dimensional structure analysis cannot generally be collected in so short a time. Fersht and Renard (1974) have pointed out, however, that it may be possible to use equilibrium methods to trap intermediates in the reaction pathway and study their structures. Until the structures of such intermediates have been determined, crystallographers will be limited to studying the binding of substrate analogs and inhibitors, which in some limited respects resemble true substrates. Nevertheless, from such studies inferences can be drawn about the structural transformations which occur during the catalyzed reaction.

Among the best analogs to true trypsin substrates are the naturally occurring trypsin inhibitors. They have evolved in parallel with the enzymes so that they bind extremely tightly to the active site. Such protein inhibitors are crucial to physiological control of the serine proteases (Tschesche 1974). For example, if pancreatic secretory trypsin inhibitors were not synthesized along with the pancreatic serine proenzymes, one prematurely activated molecule of trypsin could start an autocatalytic chain reaction which would activate the other serine proenzymes and destroy any nearby proteins. Inhibitors are present to prevent such catastrophes and to control physiological processes mediated by proteolytic enzymes. The structure of an intracellular 6500 molecular weight trypsin inhibitor (PTI) isolated from bovine pancreas and other organs was determined by R. Huber et al. (1970, 1971). Chemical modifications had already shown that Lys-15 of this inhibitor was involved in the trypsin-PTI association (Chauvet and Acher 1967; Kress and Laskowski 1967; Fritz et al. 1969). By combining models of PTI with the known structures of trypsin and chymotrypsin, substrate binding models (Fig. 4) were developed by us (Stroud, Kay and Dickerson 1971; Krieger, Kay and Stroud 1974) and independently by Huber et al. (1971) and Blow et al. (1972). High-resolution structures of the PTI-trypsin complex (Ruhlmann et al. 1973) and of a soybean trypsin inhibitor-trypsin complex (Blow, Janin and Sweet 1974; Sweet et al. 1974) have since been determined. Several hydrogen bonds and stereochemical complementarity between enzyme and inhibitor orient the susceptible bond at the active site. There is one very important difference between the predicted substrate binding models and the structures of the trypsin-inhibitor complexes; the inhibitor and trypsin are covalently bound together via an oxygen-carbon bond between the hydroxyl of Ser-195 and the carbonyl group of Lys-15 in the trypsin-PTI complex and between the serine hydroxyl and the carbonyl carbon of Arg-63 in the trypsin-STI (soybean trypsin inhibitor) complex. These complexes have been shown to exist as tetrahedral adducts which probably resemble normal intermediates (ES^T) in serine protease catalysis.



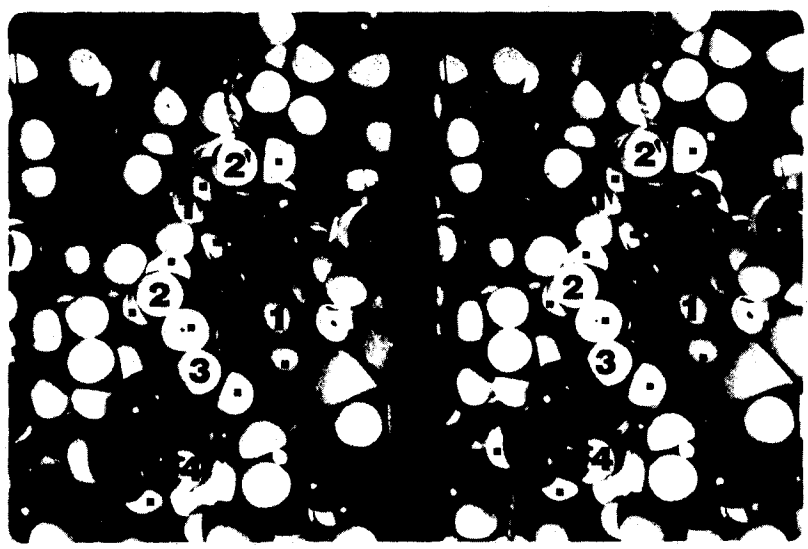


Figure 4

(*Top*) A model for the binding of a portion of the bovine pancreatic trypsin inhibitor (heavy shading) to trypsin in which the side chain of the inhibitor's Lys-15 has been replaced by an arginine side chain. The model was constructed so that the "Arg-15" side chain fitted the electron density for benazmidine in benzamidine-trypsin. This model for enzyme-substrate interaction embodies a substrate conformation that evolved to bind tightly to the enzyme and an enzyme conformation which is presumably like that induced by the binding of specific substrate side chains. (Reprinted, with permission, from Krieger, Kay and Stroud 1974.)

(*Bottom*) Stereoscopic view of a space-filling model of the trypsin-substrate complex described in the text. This view (approximately as above) should be compared with the same region of the enzyme alone shown in Figure 2. This figure beautifully demonstrates the intimate stereochemical compatibility between enzyme and substrate. The woolen threads indicate the end points of the portion of the oligopeptide substrate chain which is included in the model. Protons in the substrate are labeled with square dots, and atoms attached to the α-carbon atoms of the substrate fragment are labeled 4-3-2-1-1′-2′ (along the sequence in the substrate). The peptide bond between 1 and 1′ is the one to be hydrolyzed.

These and other crystallographic and chemical studies have produced detailed models for the association of enzyme and substrate prior to catalysis. Although based on inferences drawn from substrate analogs or inhibitor binding studies, such models do suggest ways that chemical groups on the enzyme may participate in catalysis.

The Active Center: pH-Activity Correlations

The X-ray crystal structures of the serine proteases have shown that their active sites are almost identical. The catalytic site of all serine proteases is characterized by a serine hydroxyl group (residue 195). Diisopropylfluorophosphate (DFP) (Jansen, Nutting and Balls 1949; Hartley 1960) and phenyl methane sulfonyl fluoride (PMSF) (Fahrney and Gold 1963; Kallos and Rizok 1964) react with this hydroxyl and irreversibly inhibit serine proteases, regardless of their substrate specificity. In the free enzyme, this hydroxyl is hydrogen bonded to the N-ϵ_2 of His-57. N-δ_1 of His-57 is hydrogen bonded to the carboxyl group of Asp-102, which in turn is not directly accessible to solvent (Blow, Birktoft and Hartley 1969; Wright, Alden and Kraut 1969; Alden, Wright and Kraut 1970; Birktoft and Blow 1972; Stroud, Kay and Dickerson 1974). The direct participation of these three groups in catalysis has been established, and chemical modification of any of them can greatly diminish or abolish catalysis (Jansen, Nutting and Balls 1949; Hartley 1960; Fahrney and Gold 1963; Kallos and Rizok 1964; Ong, Shaw and Schoellmann 1964; Shaw, Mares-Guia and Cohen 1965; Henderson 1971; Martinek, Savin and Berezin 1971; Chambers et al. 1974).

The pH-activity profiles for hydrolysis of peptides, amides or esters by trypsin or chymotrypsin are bell shaped and reflect maximal enzymatic activity at about pH 8. The high pH limb of the curve depends on an apparent pK_a of 8.8 for α-chymotrypsin (Fersht and Requena 1971b) or 10.1 for trypsin (Spomer and Wootton 1971). Fersht and Requena have demonstrated that this ionization, which controls enzyme conformation and substrate binding (K_m), is directly associated with titration of the α-amino terminus of Ile-16. The internal salt bridge formed between this amino group and the side chain of Asp-194 in the active enzyme is broken at high pH, where deprotonation of the amino group was shown to favor an alternate conformation for the enzyme.

The low pH limb of the profile depends on a single group of pK_a about 6.7 in both enzymes; protonation of this group adversely affects both acylation and deacylation (Bender and Kezdy 1964). It has often been assumed that this group corresponds to His-57. Jencks (1969), however, has pointed out that this group need not be His-57, but might be some other group on the enzyme either controlling conformation or effecting a change in rate-determining step near this pH. The aspartic-histidine-serine system *as a whole* has been shown to take up a single proton as the pH is lowered below 7.0. Fersht and Renard (1974) have also demonstrated that for the hydrolysis of acetyl phenylalanine-*p*-nitrophenyl ester by δ-chymotrypsin, k_{cat} or k_{cat}/K_m depends on a single ionization between pH 9.0 and pH 2.0. Thus it would seem that only one group at the active center has a pK_a in the range pH 2–9 which can be detected kinetically.

Richards and his colleagues (Hunkapiller et al. 1973) have shown by

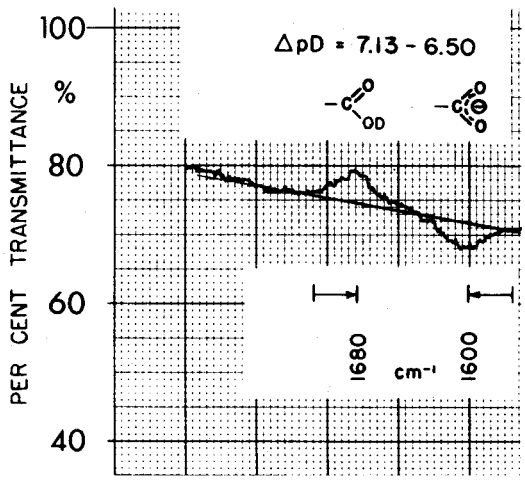
nuclear magnetic resonance studies that the imidazole of His-57 in α-lytic protease (a bacterial homolog of the pancreatic serine proteases) does not ionize until the pH is lowered below 4.0. Their research led them to propose that the group of pK_a 7.0 was Asp-102.

In order to study the ionization of Asp-102 in trypsin (Koeppe and Stroud, unpubl.), we monitored infrared absorbance arising from the carboxyl C=O stretch at 1570 cm^{-1} and 1710 cm^{-1} (Timasheff and Rupley 1972) as a function of pH. To diminish the number of titratable carboxyls, accessible carboxyl groups in trypsin were modified with semicarbazide (Fersht and Sperling 1973). The modified enzyme was found to contain 2.5 molar equivalents of free carboxyl groups. These were identified as Asp-102, Asp-194 (1.0 molar equivalent each) and the α-carboxyl terminus of Asn-245 (0.5 equivalent).

The spectrum shown in Figure 5 indicates differential absorbance at 1600 cm^{-1} and at 1680 cm^{-1} between semicarbazide-trypsin solutions of pD 6.50 and 7.13. Each peak is shifted toward the other by about 30 cm^{-1} from the value found for other carboxyls—a result which is to be expected for hydrogen-bonded carboxyls (Susi 1972) such as Asp-102. The titration curve of semicarbazide-trypsin compiled from a series of infrared difference spectra is shown in Figure 6. Based on the number of free carboxyl groups, we assume that the low pH titration of average pK_a 2.9 corresponds to titration of 1.5 carboxyl groups, while the titration of average pK_a 6.8 corresponds to one carboxyl. The gradient of the low pH titration is approximately 1.5 times that of the upper one, which is consistent with the assumption. Both titrations, however, appear sharper than expected for single or noninteracting groups.

Binding of Cu^{++} ions displaces the upper limb of the titration downward from pH 6.8. Martinek et al. (1969, 1971) have shown that Ag^{+} ions are powerful competitive inhibitors of trypsin, and that Cu^{++} and Ag^{+} compete with each other in inhibiting chymotrypsin. Either species competes with pro-

Figure 5
Infrared difference spectrum for semicarbazide-trypsin. The path length was 0.150 mm. The sample cell was at pD 7.13, and the reference cell at pD 6.50. pD values in all cases correspond to uncorrected pH meter readings. Concentrations were 1.5 mM enzyme, 6 mM $NaNO_3$, and 12 mM benzamidine. The peak positions at 1680 cm^{-1} for C=O in COOD and at 1600 cm^{-1} for C—O in COO^- are closer together than for other trypsin carboxyls due to the effect of hydrogen bonding. The detected difference in species concentration judged from peak heights corresponds to about 0.4 carboxyl equivalent.



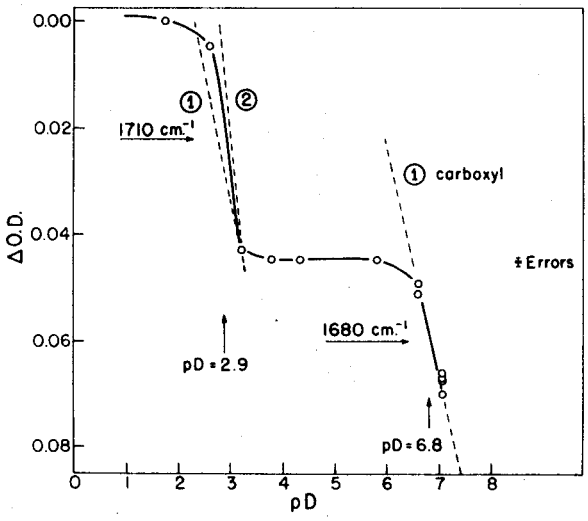


Figure 6
Titration curve for free carboxyls in semicarbazide-trypsin obtained by plotting differential absorbance at 1710 cm^{-1} and 1680 cm^{-1} from a series of infrared difference spectra. Reference solutions were at pD 3.2, 3.8, 4.4 and 6.6. The gradient of the titration at low pD was 1.5–2.0 times greater than the gradient for the group of pKa 6.8 and corresponds to titration of about 1.5 carboxyls.

tons for a site on the enzyme with $pK_{app} \simeq 7.0$. Since we have shown that Ag^+ binds specifically between Asp-102 and His-57 in the orthorhombic crystal form of DIP-trypsin (Chambers et al. 1974 and Fig. 7), in trypsinogen and in trigonal DIP-trypsin (Kossiakoff, Kay and Stroud, unpubl.), the pK_a of 6.8 has tentatively been assigned to the carboxyl of Asp-102. (A detailed description of the experimental procedures and results will be published elsewhere.) Copper ion binding shows that Asp-194 and Asn-245 cannot be responsible for this pK_a. These data suggest that the average apparent pK_a of Asp-194 and Asn-245 is 2.9.

Control experiments eliminate the possibility that imidazole-stretching frequencies could account for the infrared bands at 1680 cm^{-1} and 1600 cm^{-1} in Figure 5. However, imidazole titration may perturb a neighboring carbonyl and thereby conceivably be responsible for the infrared difference peaks observed around pD 6.8. This possibility is the subject of continuing investigations in our laboratory.

There are two arguments against this possibility and in support of the assignment of a pK_a of about 7.0 to Asp-102. First, using an average extinction coefficient for the other seven carboxyl groups in β-trypsin, derived from infrared difference spectra of the unmodified enzyme, the low pH limb of the titration shown in Figure 6 corresponds to 1.5 carboxyl equivalents. This implies that the remaining carboxyl group must titrate outside of the range pH 2–pH 5. Second, in the presence of Cu^{++} ions, there is no differential

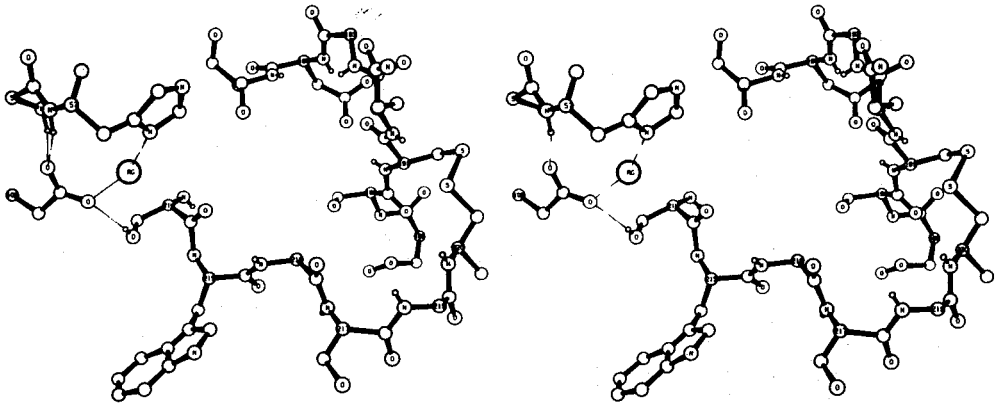


Figure 7
A stereo ORTEP drawing of the catalytic site of silver DIP-trypsin, shown in the same orientation as Figure 1. The DIP group has been omitted for clarity. The γ-oxygen of Ser-195 is in the position found for DIP-trypsin, close to that found in tetrahedral intermediates or in acyl enzymes (see text). (Reprinted, with permission, from Chambers et al. 1974.)

absorbance in the region of pH 7.0, which suggests that either of the two imidazoles at His-40 or His-91 which may titrate around pH 7.0 do not induce changes in their neighboring carbonyl groups which would be detected by the technique.

The Mechanism of Hydrolysis by Serine Proteases

Following their determination of the active-site structure of chymotrypsin, Blow, Birktoft and Hartley (1969) first proposed that proton transfers between His-57 and Asp-102 were important in catalysis. The studies of the microscopic pK_a's of His-57 and Asp-102 referred to in the previous section are consistent with this proposal. From their studies of the histidine ionization in α-lytic protease, Hunkapiller et al. (1973) explained the sequence of proton transfers between Asp-102 and His-57, discussed here and included in Figure 9 (below), in terms of pK_a's.

In this discussion we assume (see previous section) that the pK_a of Asp-102 is 6.8, and that the imidazole of His-57 is essentially neutral above pH 4.0. This leads to the ionization of the active center around pH 7.0 (Fig. 8). The mechanistic importance of these assignments is that the aspartate ion of residue 102 can act as a chemical base which can readily accept a proton from the histidine side chain during catalysis (Hunkapiller et al. 1973). Together, Asp-102 and His-57 shuttle protons back and forth from enzyme to substrate, and so the mechanism can best be described as nucleophilic attack with general base catalysis by His-57 (Bender and Kezdy 1964; Inward and Jencks 1965) and Asp-102. The important differences between this reaction and a non-enzymatic hydrolysis are the binding to the enzyme and the efficient proton shuttle.

Interfering with this shuttle inhibits catalysis. For example, by methylating the N-ϵ_2 of His-57 in chymotrypsin, the shuttle can no longer operate normally

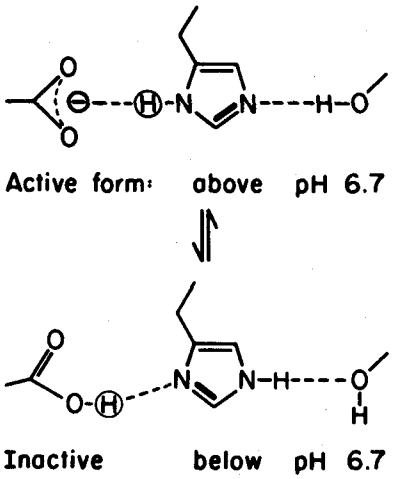


Figure 8
Ionization of the active center in the range pH 4.0–8.0 as discussed in the text.

and the rate of catalysis drops by a factor of 5000 to 200,000 for specific substrates (Henderson 1971). Silver ions bind specifically to trypsin between Asp-102 and His-57 (Chambers et al. 1974). By adding silver, the shuttle is blocked and catalysis is inhibited (Martinek, Savin and Berezin 1971).

The mechanistic scheme shown in Figure 9 is consistent with most experimental data relating to hydrolysis of peptides, esters or amides by trypsin or chymotrypsin. In the first step (I), substrate and enzyme form a Michaelis complex. Nucleophilic attack by the hydroxyl group of Ser-195 follows. As the reaction proceeds, the hydroxyl twists around the C_α—C_β bond and forms a covalent chemical bond to the substrate carbon at step I–II (Steitz, Henderson and Blow 1969). Concerted with this, a proton is transferred from the serine hydroxyl group to the N-ϵ_2 of His-57. From there, it is eventually delivered to the nitrogen of the peptide bond in the substrate. As a result of this proton transfer, the proton previously bound to the N-δ_1 of His-57 is transferred to the carboxyl group of Asp-102, which acts as a base in this reaction (Hunkapiller et al. 1973). Although these proton transfers are rapid, deuterium isotope effects show that proton transfer is involved in the rate-determining step of the catalysis (Bender et al. 1964; Pollock, Hogg and Schowen 1973).

Whether the Asp-His-Ser proton shuttle is concerted or stepwise remains in question. If the mechanism is concerted, the negative charge of Asp-102 would be neutralized while negative charge develops on the carbonyl oxygen of the substrate. The imidazole would remain neutral throughout the reaction; thus unstable intermediates due to charge separation would be avoided (Jencks 1971; Hunkapiller et al. 1973). In fact, charge development in the transition state in chymotrypsin catalysis does appear to be small (Jencks 1971). The shuttle may be stepwise if the energy requirements of charge separation (negative charges on the substrate and Asp-102 and a positive charge on His-57) are offset by a more favorable entropy of activation in a two-step process (Jencks 1972).

One might favor the concerted mechanism because it might be expected that the precise alignment of the shuttle, which has been observed in all serine protease structures, evolved so that the entropic advantage of the

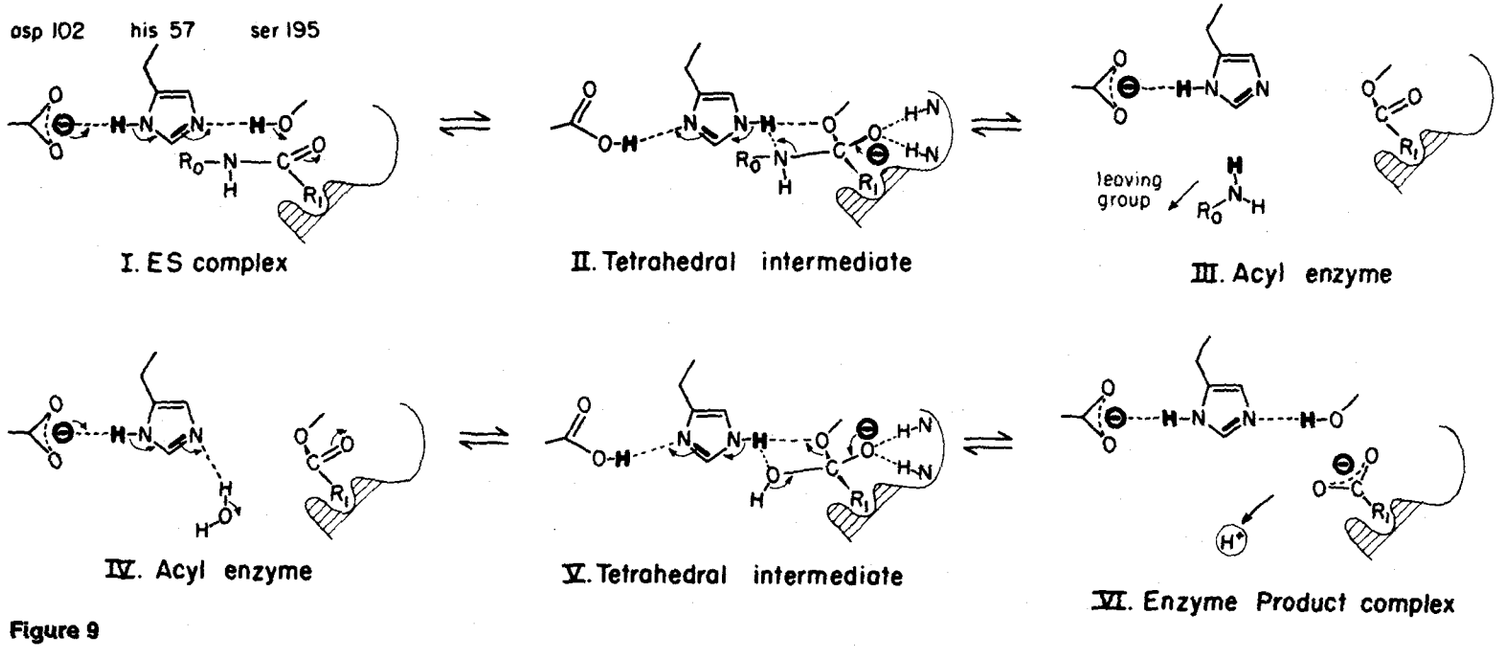


Figure 9
A mechanism for serine protease hydrolysis of peptides or amides. In this representation, the proton shuttle is concerted.

two-step process over the concerted process was minimized. Thus the enzyme could exploit for increased reaction rate the energy saved in eliminating charge separation. If this were not the case, it would seem unnecessary to use both an Asp and an His for the general base catalysis. The Asp could be eliminated and the His could act as the base.

After the attack by the serine on the substrate, a short-lived tetrahedral intermediate is formed (II). This intermediate is stabilized by the covalent bond to the enzyme and by a number of hydrogen bonds. The following structural features of the tetrahedral intermediate are primarily based on the crystallographic determination of many different protease-inhibitor structures.

The negatively charged substrate oxygen in the tetrahedral intermediate is stabilized by hydrogen bonds from the amide nitrogens of residues 195 and 193. The importance of the amides of Ser-195 and Gly-193 was first noted by Steitz, Henderson and Blow (1969), and their role in transition-state stabilization was postulated by Henderson (1970) and by Robertus et al. (1972). Another hydrogen bond forms between the carbonyl group of Ser-214 and the α-N of the substrate (Steitz, Henderson and Blow 1969; Segal et al. 1971). Comparison of the kinetics of hydrolysis of specific trypsin and chymotrypsin substrates with and without the hydrogen bonding capacity of the α-N suggests that the Ser-214—α-N bond may not form in the Michaelis complex (Ingles and Knowles 1968; Caplow and Harper 1972; Kobayashi and Ishi 1974). These results show, however, that this hydrogen bond does play a role in the transition states between intermediates and possibly in the tetrahedral and acyl enzyme intermediates.

One explanation for the exceptional catalytic powers of enzymes is that enzymes have evolved so that they can optimally bind the transition-state structures in the reactions they catalyze rather than the substrates themselves (Pauling 1946; Wolfenden 1972). The hydrogen-bonded structure in the serine protease-substrate transition state may be an example of transition-state stabilization, for the oriented hydrogen bonds can help to speed up the reaction by smoothing down the highest energy barriers between intermediate states. The stability of the tetrahedral adduct in the trypsin-trypsin inhibitor complexes is consistent with the transition-state stabilization hypothesis.

At step II–III, the now unstable carbon-nitrogen bond is broken, and the first product of hydrolysis, an amine, is free to diffuse away, taking with it a proton from the enzyme. At the same time, the bound part of the substrate rearranges to a covalently modified acyl enzyme intermediate (III). At pH 8, N^{14}/N^{15} kinetic isotope effects (O'Leary and Kluetz 1972) show that the C—N bond rupture is partially rate-determining for the hydrolysis of acetyl tryptophanamide by chymotrypsin. The rate-determining step for amide hydrolysis, however, may vary from the formation of the tetrahedral intermediate to its breakdown, depending upon the pH and the structure of the substrates (Fastrez and Fersht 1973a).

The breakdown of the acyl intermediate (IV–VI) is the microscopic reverse of steps I–III; this time water is the attacking group. At step V–VI, the second product is formed. It is an acid which loses a proton to the solution and becomes negatively charged. For the first time (if the proton shuttle is concerted), there are two charges in the system. These two negative charges repel each other and so help to dissociate the second product from the enzyme (Johnson and Knowles 1966), regenerating free enzyme.

The presence of a carboxyl group of high pK_a and a neutral side chain of His-57 with a low pK_a would suggest two compelling evolutionary reasons why the Asp-His-Ser arrangement should be universal to serine proteases (Hunkapiller et al. 1973). First, by neutralizing a negative charge on Asp-102, rather than generating a positive charge on His-57, during formation of the tetrahedral intermediate, there would be no unfavorable charge separation. This would contribute to reducing transition-state internal energies, and thus to rate enhancement (Jencks 1971). Second, if the charged Asp-102 is to be a proton acceptor at physiological pH values, its pK_a must be raised and it must have access to the proton donor. The imidazole of His-57 is ideally suited both to insulate Asp-102 from solvent (so raising the pK_a of the buried carboxyl group) and to serve as a proton conductor, transferring charge from the carboxyl group to the substrate. It is also important to note that both the reverse separation of the pK_a values of Asp-102 and His-57 and the structure of trypsin at pH 7 and pH 8 (Stroud, Kay and Dickerson 1974; Krieger, Kay and Stroud 1974; Huber et al. 1974; Sweet et al. 1974), which shows a symmetric interaction between the charge on Asp-102 and His-57 (see Fig. 3), are unlike the situation expected in aqueous solution and reflect a unique microenvironment for these groups.

As far as we know, all the serine proteases use the same three chemical groups to hydrolyze peptide bonds. They, like trypsin, are active catalysts only when the aspartic acid is negatively charged. Against the active site, the reaction goes on in an unique way as the enzyme smoothes the transition from one intermediate state to another. This emphasizes the importance of exact stereochemical fit and correct orientation (Koshland 1958) of the substrate as the reaction takes place, rather than simply the generation of an especially reactive site. After all, a very reactive site could react in many less specific ways. It is better to have a moderately efficient catalytic site coupled with a very selective binding requirement (Fersht and Sperling 1973). With this in mind, the subtle differences between serine proteases involved in biological control can be understood in terms of differences in their specific substrate binding properties.

Acknowledgments

This is contribution number 4990 from the Norman W. Church Laboratory of Chemical Biology. We would like to acknowledge our indebtedness to the National Institutes of Health for their support by means of grants GM-19984 and GM-70469. R.M.S. is the recipient of an NIH Career Development Award, M.K. is the recipient of a Danforth Foundation Fellowship, R.K. and J.C. are recipients of NIH Predoctoral Traineeships, and A.K. is the recipient of an NIH Postdoctoral Fellowship.

REFERENCES

Alden, R. A., C. S. Wright and J. Kraut. 1970. A hydrogen-bond network at the active site of subtilisin BPN'. *Phil. Trans. Roy. Soc. London* B **257**:119.

Bender, M. L. and F. J. Kezdy. 1964. The current state of the α-chymotrypsin mechanism. *J. Amer. Chem. Soc.* **86**:3704.

Bender, M. L., G. E. Clement, F. J. Kezdy and H. D'A. Heck. 1964. The correlation of the pH (pD) dependence of the stepwise mechanism of α-chymotrypsin-catalyzed reactions. *J. Amer. Chem. Soc.* **86:**3680.

Birktoft, J. J. and D. M. Blow. 1972. Structure of crystalline α-chymotrypsin. V. The atomic structure of tosyl-α-chymotrypsin at 2 Å resolution. *J. Mol. Biol.* **68:**187.

Blow, D. M., J. J. Birktoft and B. S. Hartley. 1969. Role of a buried acid group in the mechanism of action of chymotrypsin. *Nature* **221:**337.

Blow, D. M., J. Janin and R. M. Sweet. 1974. Mode of action of soybean trypsin inhibitor (Kunitz) as a model for specific protein-protein interactions. *Nature* **249:**54.

Blow, D. M., C. S. Wright, D. Kukla, A. Ruhlmann, W. Steigemann and R. Huber. 1972. A model for the association of bovine pancreatic trypsin inhibitor with chymotrypsin and trypsin. *J. Mol. Biol.* **69:**137.

Brown, J. R., D. L. Kauffman and B. S. Hartley. 1967. The primary structure of porcine pancreatic elastase: The N-terminus and disulphide bridges. *Biochem. J.* **103:**497.

Caplow, M. 1969. Chymotrypsin catalysis. Evidence for a new intermediate. *J. Amer. Chem. Soc.* **91:**3639.

Caplow, M. and C. Harper. 1972. Discrete effects of the acylamino proton in a chymotrypsin substrate on different processes in catalysis. *J. Amer. Chem. Soc.* **94:**6508.

Chambers, J. L., G. G. Christoph, M. Krieger, L. Kay and R. M. Stroud. 1974. Silver ion inhibition of serine proteases: Crystallographic study of silver trypsin. *Biochem. Biophys. Res. Comm.* **59:**70.

Chauvet, J. and R. Acher. 1967. The reactive site of the basic trypsin inhibitor of pancreas. *J. Biol. Chem.* **242:**4274.

Davie, E. W. and H. Neurath. 1955. Identification of a peptide released during autocatalytic activation of trypsinogen. *J. Biol. Chem.* **212:**515.

East, E. J. and C. G. Trowbridge. 1968. Binding of benzamidine and protons to trypsin as measured by difference spectra. *Arch. Biochem. Biophys.* **125:**334.

Fahrney, D. E. and A. M. Gold. 1963. Sulfonyl fluorides as inhibitors of esterases. *J. Amer. Chem. Soc.* **85:**997.

Fastrez, J. and A. R. Fersht. 1973a. Mechanism of chymotrypsin. Structure, reactivity, and nonproductive binding relationships. *Biochemistry* **12:**1067.

———. 1973b. Demonstration of the acyl-enzyme mechanism for the hydrolysis of peptides and anilides by chymotrypsin. *Biochemistry* **12:**2025.

Fersht, A. R. and M. Renard. 1974. pH dependence of chymotrypsin catalysis. *Biochemistry* **13:**1416.

Fersht, A. R. and Y. Requena. 1971a. Mechanism of the α-chymotrypsin hydrolysis of amides. pH dependence of k_c and K_m. Kinetic detection of an intermediate. *J. Amer. Chem. Soc.* **93:**7079.

———. 1971b. Equilibrium and rate constants for the interconversion of two conformations of α-chymotrypsin. *J. Mol. Biol.* **60:**279.

Fersht, A. R. and J. Sperling. 1973. The charge relay system in chymotrypsin and chymotrypsinogen. *J. Mol. Biol.* **74:**137.

Fersht, A. R., D. M. Blow and J. Fastrez. 1973. Leaving group specificity in the chymotrypsin catalyzed hydrolysis of peptides. A stereochemical interpretation. *Biochemistry* **12:**2035.

Freer, S. T., J. Kraut, J. D. Robertus, H. T. Wright and Ng.- H. Xuong. 1970. Chymotrypsinogen: 2.5 Å crystal structure, comparison with α-chymotrypsin, and implications for zymogen activation. *Biochemistry* **9:**1997.

Fritz, H., H. Schult, R. Meister and E. Werlo. 1969. Herstellung und eigenschaften

von derivaten des Trypsin-Kallikrein-Inhibitors aus Rinderorganen. *Z. Physiol. Chem.* **350:**1531.

Gertler, A., K. A. Walsh and H. Neurath. 1974. Catalysis by chymotrypsinogen. Demonstration of an acyl-zymogen intermediate. *Biochemistry* **13:**1302.

Hartley, B. S. 1960. Proteolytic enzymes. *Annu. Rev. Biochem.* **29:**45.

———. 1964. Amino acid sequence of bovine chymotrypsinogen A. *Nature* **201:**1284.

Henderson, R. 1970. Structure of crystalline α-chymotrypsin. IV. Structure of indoleacryloyl-α-chymotrypsin and its relevance to the hydrolytic mechanism of the enzyme. *J. Mol. Biol.* **54:**341.

———. 1971. Catalytic activity of α-chymotrypsin in which histidine-57 has been methylated. *Biochem. J.* **124:**13.

Hess, G. P., J. McConn, E. Ku and G. McConkey. 1970. Studies of the activity of chymotrypsin. *Phil. Trans. Roy. Soc. London* B **257:**89.

Huber, R., D. Kukla, A. Ruhlmann and W. Steigemann. 1971. Pancreatic trypsin inhibitor: Structure and function. *Cold Spring Harbor Symp. Quant. Biol.* **36:**141.

Huber, R., D. Kukla, A. Ruhlmann, O. Epp and H. Formanek. 1970. The basic trypsin inhibitor of bovine pancreas. I. Structure analysis and conformation of the polypeptide chain. *Naturwissenschaften* **57:**389.

Huber, R., D. Kukla, W. Steigemann, J. Deisenhofer and A. Jones. 1974. Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. In *Bayer Symposium V: Proteinase Inhibitors* (ed. H. Fritz et al.), p. 497. Springer Verlag, Heidelberg.

Hunkapiller, M. W., S. H. Smallcombe, D. R. Whitaker and J. H. Richards. 1973. Carbon nuclear magnetic resonance studies of the histidine residue in α-lytic protease. *Biochemistry* **12:**4732.

Inagami, T. and T. Murachi. 1964. The mechanism of the specificity of trypsin catalysis. *J. Biol. Chem.* **239:**1395.

Inagami, T. and S. S. York. 1968. The effect of alkylguanidines and alkylamines on trypsin catalysis. *Biochemistry* **7:**4045.

Ingles, D. W. and J. R. Knowles. 1968. The stereospecificity of α-chymotrypsin. *Biochem. J.* **108:**561.

Inward, P. W. and W. P. Jencks. 1965. The reactivity of nucleophilic reagents with furoyl-chymotrypsin. *J. Biol. Chem.* **240:**1986.

Jacob, F. and J. Monod. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3:**318.

Jansen, E. F., M. D. F. Nutting and A. K. Balls. 1949. Mode of inhibition of chymotrypsin by diisopropylfluorophosphate. I. Introduction of phosphorous. *J. Biol. Chem.* **179:**201.

Jencks, W. P. 1969. *Catalysis in Chemistry and Enzymology*, p. 218. McGraw-Hill, New York.

———. 1971. Reactivity correlations and general acid-base catalysis in enzymic transacylation reactions. *Cold Spring Harbor Symp. Quant. Biol.* **36:**1.

———. 1972. General acid-base catalysis of complex reactions in water. *Chem. Rev.* **72:**705.

Johnson, C. H. and J. R. Knowles. 1966. The binding of inhibitors to α-chymotrypsin. *Biochem. J.* **101:**56.

Kallos, J. and D. Rizok. 1964. Heavy atom labeling of the serine of the active center of chymotrypsin: Pipsyl-chymotrypsin. *J. Mol. Biol.* **9:**255.

Kassell, B. and J. Kay. 1973. Zymogens of proteolytic enzymes. *Science* **180:**1022.

Kobayashi, R. and S. Ishi. 1974. The trypsin-catalyzed hydrolysis of some L-α-amino-lacking substrates. *J. Biochem.* **75:**825.

Koshland, D. E., Jr. 1958. Application of a theory of enzyme specificity to protein synthesis. *Proc. Nat. Acad. Sci.* **44**:98.

Kress, L. F. and M. Laskowski. 1967. The basic trypsin inhibitor of bovine pancreas. *J. Biol. Chem.* **242**:4925.

Krieger, M., L. M. Kay and R. M. Stroud. 1974. Structure and specific binding of trypsin: Comparison of inhibited derivatives and a model for substrate binding. *J. Mol. Biol.* **83**:209.

Kunitz, M. and J. H. Northrop. 1935. Crystalline chymo-trypsin and chymotrypsinogen. I. Isolation, crystallization, and general properties of a new proteolytic enzyme and its precursor. *J. Gen. Physiol.* **18**:433.

———. 1936. Isolation from beef pancreas of cystalline trypsinogen, trypsin, a trypsin inhibitor, and an inhibitor-trypsin compound. *J. Gen. Physiol.* **19**:991.

Leighton, T. J., R. H. Doi, R. A. J. Warren and R. A. Kelln. 1973. The relationship of serine protease activity to RNA polymerase modifications and sporulation in *Bacillus subtilis*. *J. Mol. Biol.* **76**:103.

Magnusson, S. 1971. Thrombin and prothrombin. In *The Enzymes* (ed. P. D. Boyer), vol. 3, p. 277. Academic Press, New York.

Mares-Guia, M. and E. Shaw. 1965. Studies on the active center of trypsin: The binding of amidines and guanidines as models of the substrate chain. *J. Biol. Chem.* **240**:1579.

Maroux, S., J. Baratti and P. Desnuelle. 1971. Purification and specificity of porcine enterokinase. *J. Biol. Chem.* **246**:5031.

Martinek, K., Y. V. Savin and I. V. Berezin. 1971. Kinetic manifestations of trypsin's active center during inhibition of its enzymatic activity by Ag^+ ions. *Biokhimiya* **36**:806.

Martinek, K., Kh. Vill', Z. A. Strel'tsova and I. V. Berezin. 1969. Kinetic manifestations of the active-center structure of α-chymotrypsin during inhibition of enzyme activity by Ag^+ ions. *Mol. Biol.* (U.S.S.R.) **3**:554.

Matthews, B. W., P. B. Sigler, R. Henderson and D. M. Blow. 1967. Three-dimensional structure of tosyl-α-chymotrypsin. *Nature* **214**:652.

Naughton, M. A. and F. Sanger. 1961. Purification and specificity of pancreatic elastase. *Biochem. J.* **78**:156.

Neurath, H. and G. H. Dixon. 1957. Structure and activation of trypsinogen and chymotrypsinogen. *Fed. Proc.* **16**:791.

Neurath, H., J. A. Rupley and W. J. Dreyer. 1956. Structural changes in the activation of chymotrypsinogen and trypsinogen. Effects of urea on chymotrypsinogen and delta-chymotrypsinogen. *Arch. Biochem. Biophys.* **65**:243.

Northrop, J. H., M. Kunitz and R. M. Herriot. 1948. *Crystalline Enzymes*, p. 96. Columbia University Press, New York.

O'Leary, M. H. and M. D. Kluetz. 1972. Nitrogen isotope effects on the chymotrypsin catalyzed hydrolysis of *N*-acetyl-L-tryptophanamide. *J. Amer. Chem. Soc.* **94**:3585.

Ong, E. B., E. Shaw and G. Schoellmann. 1964. An active center histidine peptide of α-chymotrypsin. *J. Amer. Chem. Soc.* **86**:1271.

Oppenheimer, H. L., B. Labouesse and G. P. Hess. 1966. Implication of an ionizing group in the control of conformation and activity of chymotrypsin. *J. Biol. Chem.* **241**:2720.

Owren, P. A. and H. Stormorken. 1973. The mechanism of blood coagulation. *Ergeb. Physiol. Biol. Chem. Exp. Pharmacol.* **68**:1.

Pauling, L. 1946. Molecular architecture and biological reactions. *Chem. Eng. News* **24**:1375.

Pollack, E., J. L. Hogg and R. L. Schowen. 1973. One-proton catalysis in the deacetylation of acetyl-α-chymotrypsin. *J. Amer. Chem. Soc.* **95**:968.

Robertus, J. D., J. Kraut, R. A. Alden and J. J. Birktoft. 1972. Subtilisin; a stereochemical mechanism involving transition-state stabilization. *Biochemistry* **11:**4293.

Ruhlmann, A., D. Kukla, P. Schwager, K. Bartels and R. Huber. 1973. Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **77:**417.

Sampath Narayanan, A. and R. A. Anwar. 1969. The specificity of purified porcine pancreatic elastase. *Biochem. J.* **114:**11.

Segal, D. M., J. C. Powers, G. H. Cohen, D. R. Davies and P. E. Wilcox, 1971. Substrate binding site in bovine chymotrypsin Aγ. A crystallographic study using peptide chloromethyl ketones as site-specific inhibitors. *Biochemistry* **10:**3728.

Shaw, E., M. Mares-Guia and W. Cohen. 1965. Evidence for an active-center histidine in trypsin through use of a specific reagent, 1-chloro-3-tosylamido-7-amino-2-heptanone, the chloromethyl ketone derived from N_a-tosyl-L-lysine. *Biochemistry* **4:**2219.

Shotton, D. M. and B. S. Hartley. 1970. Amino-acid sequence of porcine pancreatic elastase and its homologies with other serine proteases. *Nature* **225:**802.

Shotton, D. M. and H. C. Watson. 1970. The three-dimensional structure of crystalline porcine pancreatic elastase. *Phil. Trans. Roy. Soc. London* B **257:**111.

Sigler, P. W., D. M. Blow, B. W. Matthews and R. Henderson. 1968. Structure of crystalline α-chymotrypsin. II. A preliminary report including a hypothesis for the activation mechanism. *J. Mol. Biol.* **35:**143.

Spomer, W. E. and J. F. Wootton. 1971. The hydrolysis of a α-*N*-benzoyl-L-argininamide catalyzed by trypsin and acetyl trypsin. Dependence on pH. *Biochim. Biophys. Acta* **235:**164.

Stambaugh, R., B. Brackett and L. Mastroianni. 1969. Inhibition of *in vitro* fertilization of rabbit ova by trypsin inhibitors. *Biol. Reprod.* **1:**223.

Steitz, T. A., R. Henderson and D. M. Blow. 1969. Structure of crystalline α-chymotrypsin. III. Crystallographic studies of substrates and inhibitors. *J. Mol. Biol.* **46:**337.

Stroud, R. M. 1974. A family of protein-cutting proteins. *Sci. Amer.* **231:**74.

Stroud, R. M., L. M. Kay and R. E. Dickerson. 1971. The crystal and molecular structure of DIP-inhibited bovine trypsin at 2.7 Å resolution. *Cold Spring Harbor Symp. Quant. Biol.* **36:**125.

———. 1974. The structure of bovine trypsin: Electron density maps of the inhibited enzyme at 5 Å and 2.7 Å resolution. *J. Mol. Biol.* **83:**185.

Susi, H. 1972. The strength of hydrogen bonding: Infrared spectroscopy. In *Methods in Enzymology* (ed. S. P. Colowick and N. O. Kaplan), vol. 26, p. 381. Academic Press, New York.

Sweet, R. M., H. T. Wright, J. Janin, C. H. Chothia and D. M. Blow. 1974. Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6 Å resolution. *Biochemistry* **13:**4212.

Timasheff, S. N. and J. A. Rupley. 1972. Infrared titration of lysozyme carboxyls. *Arch. Biochem. Biophys.* **150:**318.

Tomkins, G. M., T. D. Gelehrter, D. Granner, D. Martin, Jr., H. H. Samuels and E. B. Thompson. 1969. Control of specific gene expression in higher organisms. *Science* **166:**1474.

Tschesche, H. 1974. Biochemistry of natural proteinase inhibitors. *Angew. Chem. Int. Ed. Eng.* **13:**10.

Wolfenden, R. 1972. Analog approaches to the structure of the transition state in enzyme reactions. *Acc't. Chem. Res.* **5:**10.

Wright, C. S., R. A. Alden and J. Kraut. 1969. Structure of subtilisin BPN′ at 2.5 Å resolution. *Nature* **221:**235.

Wright, H. T. 1973. Comparison of the crystal structures of chymotrypsinogen-A and α-chymotrypsin. *J. Mol. Biol.* **79:**1.

Zerner, B. and M. L. Bender. 1964. The kinetic consequences of the acyl-enzyme mechanism for the reactions of specific substrates with chymotrypsin. *J. Amer. Chem. Soc.* **86:**3669.

Zerner, B., R. P. M. Bond and M. L. Bender. 1964. Kinetic evidence for the formation of acyl-enzyme intermediates in the α-chymotrypsin-catalyzed hydrolysis of specific substrates. *J. Amer. Chem. Soc.* **86:**3674.

APPENDIX C

The Effect of Pre-Incubation on Trypsin Kinetics at Low pH*

Roger E. Koeppe II, Monty Krieger, and Robert M. Stroud

ABSTRACT: The deacylation rate for hydrolysis of N_{α}-carbobenzoxy-L-lysine-p-nitrophenyl ester by trypsin depends on ionization of a single group of pK_a 6.8 on the enzyme which is required in its deprotonated form for activity (Bender, M. L., Kezdy, F. J., and Feder, J. (1965), J. Amer. Chem. Soc. 87, 4953-4954). It is shown here that the Dixon plot for this reaction is linear between pH 3.0 and 5.0 and has a slope of ~1; therefore, there is no second group on the enzyme with a pK_a between 3 and 7 which affects the rate of this reaction. Since there are two groups in the active center of serine proteases which might be expected to affect the rate of hydrolysis in this pH range, it is clear that the second group must have a pK_a less than 3. Thus, it would seem that spectroscopic studies which do

---

*Contribution No. 5253 from the Norman W. Church Laboratory of Chemical Biology, California Institute of Technology, Pasadena, California 91125. Supported by National Institutes of Health Grant GM-19984; a National Institutes of Health Career Development Award (RMS); a National Science Foundation Predoctoral Fellowship (REK); a National Institutes of Health Predoctoral Traineeship (REK); and a Danforth Foundation Fellowship (MK).

detect ionization of a second group with a pK_a in the pH range 3.0–7.0 are monitoring a group which does not control catalysis.

A possible source of discrepancy between kinetic and spectroscopic studies could arise if a slow pH-dependent conformational change took place during the long incubations required for most spectroscopic studies, and affected the pK_a's of groups at the active center. No such effect is observed within the time scale of one minute to three hours when enzyme pre-incubated at pH 2.0 or at pH 6.9 precedes kinetic studies. Therefore, spectroscopic studies which do detect a second ionization with a pK_a in the range 3–7 are clearly detecting the ionization of a group which does not affect the rate of hydrolysis, and so cannot be that of His 57 in the active enzyme.

Introduction

The active site in the trypsin-like enzymes contains both an imidazole (His 57) and a carboxylic acid (Asp 102). Because the normal pK_a's of the side chains of histidine and aspartic acid in solution are about 6.0 and 3.6 (Greenstein and Winitz, 1961), respectively, it is surprising that there is only one group (of pK_a 6.8) whose ionization affects the rate of catalysis by chymotrypsin or

trypsin in the pH range 2.0 - 7.0 (Kezdy et al., 1964; Bender et al., 1965; Fersht and Renard, 1974).

It has sometimes been assumed that if individual pK_a's could be assigned to the two groups, then His 57 would most likely be the group of pK_a ~7. However, some spectroscopic studies provide evidence that the group ionizing at pH ~7 is Asp 102 (Hunkapiller et al., 1973; Koeppe and Stroud, 1976). Because only one pK_a on the enzyme is detected in the rate profile, the second group (His 57) may have a remarkably low pK_a (less than 2.0), although there is as yet no precedent for such a large perturbation of imidazole ionization in proteins.

Any apparent conflict between kinetic data which do not detect a rate-controlling pK_a between 2.0 and 7.0, and spectroscopic studies which lead to an assignment for the pK_a of the second group within this range (Markley and Porubcan, 1975) might be resolved if the ionization of the imidazole of His 57 were dependent on a very slow pH-dependent conformational change. This ionization might not be detected in kinetic experiments in which aliquots of a stock solution of enzyme at one pH are diluted into reaction mixtures at various pH's. If a conformational change did not occur before the reaction, but did occur during the long incubations used in spectroscopic studies, then the two methods could give different results.

One way to approach the question of whether a slow conformational change affects the pH dependence of serine protease hydrolysis would be to conduct kinetic experiments with enzyme pre-incubated at different pH's. We therefore measured the rate of hydrolysis of N_α-carbobenzoxy-L-lysine-p-nitrophenyl ester by trypsin which had been pre-incubated for up to three hours at pH 2.0, or at pH 6.9. This work also presents the pH-activity profile for hydrolysis of an ester by trypsin down to pH 2.0.

Materials and Methods

Materials: Bovine trypsin (three times crystallized, lyophilized, salt free; lot 73M339) was purchased from Worthington Biochemical Corporation, and N_α-carbobenzoxy-L-lysine-p-nitrophenyl ester (CLNE)[1] was obtained from Cyclo Chemicals (lot D-1308). PIPES was purchased from Calbiochem, and all other chemicals were reagent grade. Water was distilled and deionized with a Barnstead ultrapure cartridge.

Kinetics: Stock solutions of trypsin (28 μM active sites (Chase and Shaw, 1967)) were pre-incubated in low pH (HCl, pH = 2.0) or

[1] Abbreviations used: CLNE, N_α-carbobenzoxy-L-lysine-p-nitrophenyl ester; PIPES, piperazine-N,N'-bis(2-ethane-sulfonic acid).

high pH (2mM PIPES-HCl, pH = 6.9) buffers at 0°C for up to three hours. Reaction buffers in the pH range 2.0 - 5.0 were prepared by mixing the appropriate amounts of buffer A (200 mM glacial acetic acid - 12 mM HCl, pH = 2.01) with buffer B (200 mM sodium acetate - 12 mM HCl, pH = 5.91).

The rates of CLNE hydrolysis were followed by observing the rate of liberation of p-nitrophenol in a Gilford Model 240 spectrophotometer at 340 nm. The reactions were run at 30.6 ±0.1°C, and the pH of the reaction solutions remained constant throughout the reaction. In a typical run, 2.0 ml of reaction buffer and 20 µl of pre-incubated enzyme solution were mixed in a 3 ml cuvette, and the reaction was initiated by the addition of 100 µl of CLNE (2.19 mM) in ethanol. The reagents were rapidly mixed by inverting the cuvette several times, and the change in absorbance at 340 nm was recorded on a Honeywell strip-chart recorder. The relative rate of hydrolysis (k_r) was determined by calculating the slope of the best straight line relating the change in absorbance at 340 nm to time.

Results

The relative rate constants for the hydrolysis of CLNE by trypsin are listed in Table I. There is essentially no difference between the results for enzyme pre-incubated at pH 2.0 and at pH 6.9;

therefore, a slow conformational change with a time constant between one minute and three hours does not affect the pH dependence of CLNE hydrolysis by trypsin. Figure 1 is a Dixon plot (Dixon, 1953) of the rate data together with overall standard deviations.

The data in Figure 1 fall on a straight line between pH 5.0 and pH 3.0; therefore, there is no second ionization which affects the deacylation rate in trypsin between pH 5.0 and pH 3.0. The slope of this line is 0.83, whereas a slope of 1.0 would be expected if the rate depended on a single ionizing group (Dixon, 1953). Decreased slopes such as these may be an effect of low ionic strength of the solution, and the overall charge on the protein (Edsall and Wyman, 1958; Kezdy et al., 1964). Data for the hydrolysis of N_{α}-acetyl-DL-tryptophan-p-nitrophenyl ester by α-chymotrypsin at ionic strength 0.05 are reported by Kezdy et al. (1964) in their Table III. A least-squares fit of a straight line to their data yields a slope of 0.86.

Between pH 3.0 and pH 2.5, the slope changes. Due to the uncertainty in measuring the very slow hydrolysis rates at pH 2.0, the slope of the line is not well-determined between pH 2.7 and pH 2.0. Nevertheless, the slope changes by a factor of approximately 2, the change expected for an additional rate-determining ionization.

Discussion

The rate of hydrolysis of CLNE by trypsin is extremely fast ($k_{3_{lim}}$ = 170 sec^{-1} (Bender et al., 1966)). This is one of the fastest hydrolysis rates by this enzyme; therefore CLNE is one of the best substrates for low pH studies of trypsin-catalyzed hydrolysis. The hydrolysis of CLNE by trypsin has been shown to proceed via an acyl enzyme (ES') intermediate (Bender and Kezdy, 1964). Thus the reaction may be represented by:

$$E + S \rightleftharpoons ES \xrightarrow{k_2} \begin{matrix} ES' \\ + \\ P_1 \end{matrix} \xrightarrow{k_3} E + P_2$$

where E is the free enzyme, S the substrate, ES the Michaelis complex, P_1 is the p-nitrophenol released, and P_2 is the free acid.

Over most of the pH range in our experiments, CLNE was at an enzyme-saturating concentration of 1.17×10^{-4} M ($K_m = 1 \times 10^{-5}$ M at pH 5.80 (Bender and Kezdy, 1965)). Therefore, the steady-state rate of release of p-nitrophenol was proportional to k_3, the deacylation rate constant. Under these conditions, pK_a's determined from Dixon plots can be assigned to groups in the acyl enzyme intermediate (ES').

Our results show that in the acyl enzyme there is no group with

a pK_a between 3 and 5 which affects the deacylation rate, k_3. Furthermore, since the rate has been shown to depend on a group of pK_a 6.8 (Bender et al., 1965) and since the slope between pH 3.0 and 5.0 in Figure 1 is characteristic of a single ionizing group, there can be no second rate-affecting pK_a between 3.0 and 7.0. For the same reaction Bender et al. (1965) reported that k_{cat} depends on a single basic group between pH 2.0 and 7.4. These results are also consistent with the findings of Stewart and Dobson (1965), who showed that a Dixon plot was linear between pH 3.6 and 4.4, and had a slope near unity for the deacylation of N_α-benzoyl-L-arginyl trypsin. Similar conclusions can be drawn from the data of Bender et al. (1964) for the hydrolysis of the non-specific substrate N_α-carbobenzoxy-L-tyrosine-p-nitrophenyl ester by trypsin between pH 3.0 and pH 7.0.

After pre-incubation of enzyme, the tryptic hydrolysis of CLNE was run at several different pH's between 2.5 and 5.05. The rates of hydrolysis by trypsin pre-incubated at pH 2.0 or at pH 6.9 were identical. Thus, there is no slow pH-dependent conformational change which affects activity in this pH range. This means that equilibrium and kinetic methods of detecting ionization changes in the active site should be equivalent in this range. There is no kinetic evidence for a second group of pK_a between 3 and 7 which affects the

acylation or deacylation rates in hydrolyses by serine proteases. Therefore, spectroscopic methods which detect pK_a's in the region between 3 and 6 would seem to be monitoring groups which have no effect on the acylation or deacylation rates.

One explanation for the curvature in the Dixon plot between pH 3.0 and pH 2.5 is that the overall rate may depend on steps in the reaction other than deacylation, due to diminished substrate binding at low pH. It is known that the K_m for CLNE increases from 1.0×10^{-5} at pH 5.80 to 7.95×10^{-4} at pH 2.66 (Bender et al., 1965). Thus, at low pH the CLNE concentration was not saturating, the relative rate may not have been equivalent to k_3 alone, and the change in slope in the Dixon plot may not represent a true pK_a. Furthermore, Bender et al. (1965) stated that their analysis of tryptic hydrolysis of CLNE showed no such curvature above pH 2.0. Their substrate concentration was presumably higher, and their report again suggests that substrate binding, rather than a rate-controlling ionization, is responsible for the curvature we see.

Another possibility might be that the change of slope between pH 2.5 and pH 3.0 may be due to the titration of a group in the acyl enzyme. It is unlikely that His 57 could be this group, because there is no corresponding slope change in hydrolysis by chymotrypsin (Kezdy et al., 1964; Fersht and Renard, 1974).

The work of Hunkapiller et al. (1973) on α-lytic protease has shown that His 57 is neutral above pH 4.0, and therefore cannot be responsible for the pK_a around 7. Koeppe and Stroud (1976) have directly established that Asp 102 in trypsin is the group with a pK_a of 6.8. Kinetics at low pH exclude the possibility that His 57 titrates between 3.0 and 6.0. Thus the pK_a of His 57 must be below 3 in trypsin and below 2 in chymotrypsin. Although the unique structure of the active site must somehow be involved in stabilizing the neutral form of the imidazole ring, there is as yet no satisfactory explanation as to why the pK_a of His 57 should be so low.

References

Bender, M. L., Begué-Cantón, M. L., Blakeley, R. L., Brubacher, L. J., Feder, J., Gunter, C. R., Kezdy, F. J., Killheffer, J. V. Jr., Marshall, T. H., Miller, C. G., Roeske, R. W., and Stoops, J. K. (1966), J. Amer. Chem. Soc. 88, 5890-5913.

Bender, M. L. and Kezdy, F. J. (1964), J. Amer. Chem. Soc. 86, 3704-3714.

Bender, M. L. and Kezdy, F. J. (1965), J. Amer. Chem. Soc. 87, 4954-4955.

Bender, M. L., Kezdy, F. J., and Feder, J. (1965), J. Amer. Chem. Soc. 87, 4953-4954.

Bender, M. L., Killheffer, J V. Jr., and Kezdy, F. J. (1964), J. Amer. Chem. Soc. 86, 5330-5331.

Chase, T. Jr. and Shaw, E. (1967), Biochem. Biophys. Res. Commun. 29, 508-514.

Dixon, M. (1953), Biochem. J. 55, 161-170.

Edsall, J. T. and Wyman, J., Biophysical Chemistry, Vol. I, pp. 512-518, Academic Press, Inc., New York, 1958.

Fersht, A. R. and Renard, M. (1974), Biochemistry 13, 1416-1426.

Greenstein, J. P. and Winitz, M., Chemistry of the Amino Acids, Vol. 1, p. 487, John Wiley and Sons, Inc., New York, 1961.

Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. R., and Richards, J. H. (1973), Biochemistry 12, 4732-4743.

Kezdy, F. J., Clement, G. E., and Bender, M. L. (1964), J. Amer. Chem. Soc. 86, 3690-3696.

Koeppe, R. E. II and Stroud, R. M. (1975), Biochemistry--in press.

Markley, J. L. and Porubcan, M. A. (1975), private communication of manuscript submitted to J. Mol. Biol.

Stewart, J. A. and Dobson, J. E. (1965), Biochemistry 4, 1086-1091.

TABLE I

Relative rates of CLNE hydrolysis by trypsin which was pre-incubated at pH 2.0, or at pH 6.9.

Reaction pH	Pre-incubation pH	Average k_r ΔA_{340} min^{-1}	Overall k_r [a] ΔA_{340} min^{-1}	Overall σ [b]
2.0	2.0	2.2×10^{-4} [c]	2.2×10^{-4}	2.8×10^{-4} [c]
	6.9			
2.5	2.0	.0023	.00245	2.2×10^{-4}
	6.9	.0026		
3.08	2.0	.0142	.0126	.0012
	6.9	.0121		
3.95	2.0	.0597	.0571	.0063
	6.9	.0613		
4.2	2.0	.1008	.1014	.0008
	6.9	.1019		
4.39	2.0	.1206	.1256	.0086
	6.9	.1295		
4.6	2.0	.2090	.2060	.0052
	6.9	.2029		
4.81	2.0	.3148	.3065	.020
	6.9	.2981		
5.05	2.0	.491	.479	.049
	6.9	.466		

[a] The "overall" rate is the average value of the rates observed at both conditions of pre-incubation.

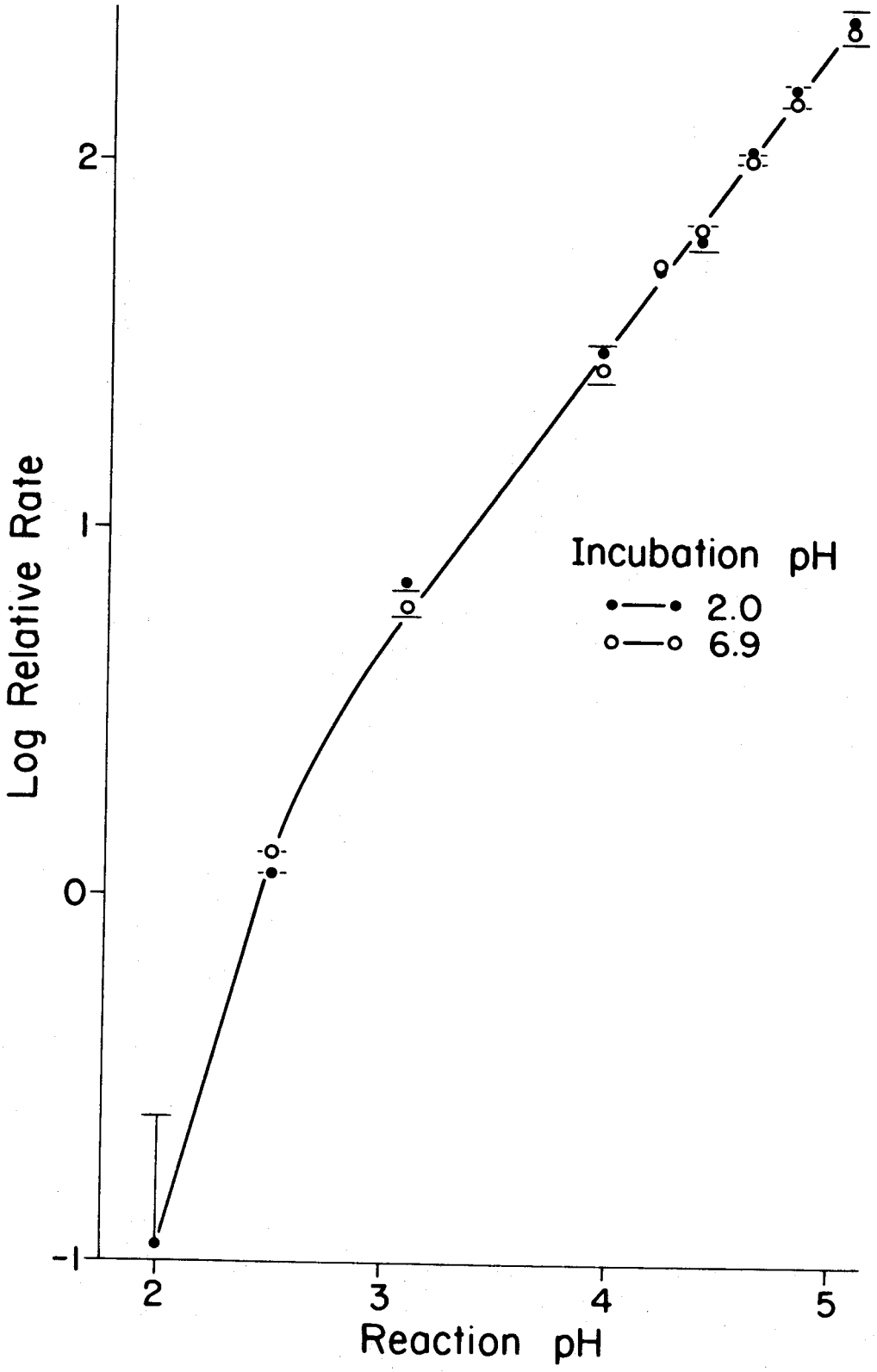
[b] $$\sigma = \left(\frac{\sum_{i=1}^{N} (k_r \text{ (overall)} - k_{r_i})^2}{N - 1} \right)^{\frac{1}{2}}$$

[c] One datum only. In this case, the overall σ is an estimate.

Figure Captions

FIGURE 1. pH dependence of CLNE hydrolysis by trypsin, which was pre-incubated at pH 2.0 (●——●) or at pH 6.9 (o——o). The error bars represent the overall standard deviations listed in Table I.

Figure 1



APPENDIX D

The pH Dependence of Tritium Exchange with the C-2 Protons of the Histidines in Bovine Trypsin*

Monty Krieger, Roger E. Koeppe II and Robert M. Stroud

ABSTRACT: At pH 8.9 and 37°C the half-times for tritium exchange with the C-2 protons of the histidines of trypsin are 73 days for His 57, and greater than 1000 days for His 40 and His 91. These half-times are much longer than the half-life of exchange for the C-2 proton of free histidine (2.8 days at pD 8.2), and longer than any previously reported half-time of exchange at pH > 8. These very low rates of exchange are discussed with reference to the refined structure of trypsin. The tritium exchange of His 57 depends on an apparent pK_a of 6.6. This pK_a may represent the pK_a of the imidazole of His 57 in an inactive conformation of the enzyme.

---

*Contribution No. 5247 from the Norman W. Church Laboratory of Chemical Biology, California Institute of Technology, Pasadena, California 91125. Supported by National Institutes of Health Grant GM-19984; a Danforth Foundation Fellowship (MK); a National Science Foundation Predoctoral Fellowship (REK); a National Institutes of Health Predoctoral Traineeship (REK); and a National Institutes of Health Career Development Award (RMS).

Introduction

The catalytic sites of all serine proteases contain three amino acid side chains which are essential for enzymatic activity: Ser 195, His 57, and Asp 102.[1] The mechanism by which these enzymes hydrolyze peptides, amides, or esters involves nucleophilic attack by the serine hydroxyl group on the susceptible carbonyl carbon of the substrate. The histidine and aspartic acid side chains may be regarded as a coupled hydrogen-bonded system which promotes the reaction by general base catalysis (Bender and Kezdy, 1964; Inward and Jencks, 1965). The base facilitates proton transfers among the reacting species, first accepting the serine hydroxyl proton during the nucleophilic attack, and later donating a proton to one of the products of hydrolysis.

This paper reports an attempt to determine the pK_a's of the three histidines of bovine trypsin using the isotope exchange method of Ohe et al. (1974). This technique involves incubating the enzyme in tritiated water at various pH's, digesting the protein, separating the histidine-containing peptides and determining the extent of the isotope incorporation into each histidine.

The kinetics of deuteration of imidazole (Vaughan et al., 1970)

---

[1] The numbering system referred to is that of chymotrypsinogen.

and of N-acetyl-histidine (Matsuo et al., 1972) in aqueous solutions have been studied as a function of pH. The pH dependence of the exchange rate can be explained by a mechanism which involves a rate-determining abstraction of the C-2 proton by OH^- or by H_2O to form an ylide intermediate, followed by a fast protonation of the ylide (Vaughan et al., 1970). Evidence that proton abstraction is involved in the rate-determining step is provided by a deuterium isotope effect of 7.5 which has been observed when comparing the rates of exchange of tritium into the tripeptide Gly-His-Gly with ^1H or ^2H in the C-2 position of the imidazole (Markley and Cheung, 1973). The pK_a of the imidazole ring can be determined by fitting the pH-exchange data to the rate equation derived from the ylide mechanism.

The rate of exchange at the C-2 of imidazole is intermediate between the rates for fast-exchanging O—H and N—H protons and non-exchanging C—H protons. The uniqueness of the rate of C-2 exchange allows the specific labeling of the histidine rings of a protein in tritiated water (Matsuo et al., 1972). The tritium can be incorporated under mild, non-denaturing conditions (37°C, pH 2-10), and the labile protons can be back-exchanged, thus leaving only histidines labeled. Studies of the pH dependence of tritium incorporation into the histidines of lysozyme (Matsuo et al., 1972) and ribonuclease (Ohe et al., 1974) have been used to determine the individual pK_a's of

the histidines in these proteins.

Our experiments have involved measuring the pH dependence of the rate of tritium incorporation into each histidine of bovine trypsin: His 57 at the active site and His 40, which are also present in chymotrypsin, and His 91. The observed rates for each of the three histidines are slower than any previously reported rates for tritium exchange into imidazole.

Methods

Materials: Some reagents employed in this work and their suppliers are: Benzamidine hydrochloride hydrate (Aldrich Chemical Company), PIPES[2] (Calbiochem), p-nitrophenyl-p'-guanidino benzoate (Cyclo Chemical Company), cyanogen bromide (Kodak), Aquasol scintillation fluid and tritiated water, 1 Ci/ml (New England Nuclear), chromatographic paper, 3 MM (Whatman), DL-norleucine (Sigma), L-histidine (Sigma), D_2O (Stohler). Sephadex G-25, G-75, and Sepharose 4B were products of Pharmacia Fine Chemicals. Standard mixtures of amino acids were purchased from Beckman Instruments Company. Trypsin (Lot 73M339, 3x recrystallized,

---

[2] Abbreviations used: PIPES, piperazine-N,N'-bis(2-ethanesulfonic acid); STI, soybean trypsin inhibitor; DIP-trypsin, diisopropylphosphoryl trypsin; BA-trypsin, benzamidine trypsin.

sterile), chymotrypsin (Lot 340698), and soybean trypsin inhibitor (Lots 54J358 and 1AA) were purchased from Worthington Biochemical Corp. Water was distilled and deionized before use, and all other chemicals were reagent grade.

Procedures: The concentration of active trypsin was determined by active site tritration with p-nitrophenyl-p'-guanidino benzoate (Chase and Shaw, 1967) and protein absorbance at 280 nm (ϵ = 1.54 mg^{-1} ml cm^{-1}; 23,891 g/mole (Robinson et al., 1971)). A Gilford Model 240 spectrophotometer was used for all spectrophotometric measurements and a Beckman Model LS-350 scintillation counter was used for tritium counting. All experiments were conducted at room temperature unless otherwise indicated. All data were assigned unit weights for the least-squares curve fitting.

Amino Acid Analysis: A Beckman Model 120-C amino acid analyzer was used for all analyses. Amino acid compositions were determined from samples hydrolyzed in distilled 6 N HCl for twenty hours at 110°C. Norleucine was used as an internal standard to determine the absolute amounts of the amino acids in the samples. The color constants used were measured from analyses of standard amino acid mixtures.

Soybean Trypsin Inhibitor-Sepharose Affinity Chromatography: Sepharose 4B was activated using a slightly modified version of the method of Cuatrecasas (1970). In the modified procedure, cyanogen bromide was dissolved in dioxane (1 g/ml) before addition to Sepharose 4B. The dioxane was purified by passage through activated alumina immediately before use (M. Ross, personal communication). The STI was coupled to the activated Sepharose 4B and the complex was subsequently washed according to the procedure of Light and Liepnieks (1974). The STI-Sepharose was packed into a 3.14 cm^2 x 22 cm column and equilibrated with a 0.1 M tris, 50 mM $CaCl_2$, 0.5 M KCl buffer, pH 8.0. Trypsin was dissolved in pH 7.14 incubation buffer (see below) before being applied to the column and eluted by step or continuous pH gradients using 0.1 M formate with 50 mM $CaCl_2$ at pH 4.5, and 0.1 M formate at pH 2.6.

Protein Purification: Sephadex G-25 and G-75 chromatography (column dimensions: 4.91 cm^2 x 90 cm; buffer: HCl, pH 2.5 with and without 40 mM $CaCl_2$), and STI-Sepharose affinity chromatography were used in attempts to isolate trypsin with a high concentration of active sites. The number of active sites in the purified trypsin was usually about 95% of the theoretical value. Control incubations without tritium showed that the benzamidine included in the

incubation buffer (see below) completely prevented autolysis during incubations of 0, 7, 10, and 14 days.

It was necessary to further purify STI by Sephadex G-75 chromatography (column dimensions: 4.91 cm^2 x 120 cm; buffer: 0.1 M NH_4HCO_3, pH 7.8, 4°C). The commercial STI was contaminated by materials with both higher and lower molecular weight.

NMR Spectroscopy: Deuterium exchange with the protons of L-histidine was measured with a Varian HR-220 spectrometer modified for Fourier transform operation. The spectrometer was interfaced with a Varian 16K 620/i computer and a Sykes compucorder 120. A solution of 0.05 M L-histidine in 99% D_2O, pD = 8.2 (Glasoe and Long, 1960) was Millipore filtered under sterile conditions, flame sealed in a 5 mm NMR tube, and incubated at 37°C. The proton resonances were assigned according to Schutte et al. (1966). The relative extent of exchange was measured by comparing the areas of the observed peaks and the C_α proton resonance was used as an internal standard. Both continuous wave and Fourier transformed NMR spectra were observed.

Tritium Exchange: The procedure used was based on the method described by Ohe et al. (1974). Incubation buffers with pH's

below 4 were prepared by adding concentrated HCl to buffer A (50 mM benzamidine, 50 mM PIPES, 100 mM HCl, pH = 4.1). Incubation buffers with pH's above 4 were prepared by adding the appropriate amount of buffer B (50 mM calcium acetate, 50 mM benzamidine, 50 mM PIPES, 50 mM NaOH, 50 mM NaCl, pH = 10.6) to buffer A. 50 mM benzamidine, a competitive reversible inhibitor of trypsin (Mares-Guia and Shaw, 1965), was included to prevent autolysis. Benzamidine binds to trypsin in the specificity binding pocket, but does not block solvent accessibility to histidine in the catalytic site (Krieger et al. 1974). 10 mg of trypsin and 25 μl of tritiated water were added to 200 μl of incubation buffer and incubated at 37°C. The final specific activity of the water was 186 cpm/nmole. After 38 hours the reaction was quenched with 25 μl of glacial acetic acid. Protein was separated from the incubation buffer and the bulk of the tritiated water by Sephadex G-25 chromatography (column dimensions: 4.91 cm^2 x 90 cm; buffer: HCl, pH 2.5). Whenever possible, samples were maintained at pH 2-3 at room temperature or below, where back exchange of tritium was quenched.

Following Sephadex G-25 chromatography, the protein was prepared for digestion to separate the histidine-containing peptides. Each sample was lyophilized twice from the pH 2.5 buffer and redissolved in 0.5 ml of HCl, pH 2.0. To improve the results of the

subsequent digestion, each sample was denatured by immersion in boiling water for five minutes. The heat-treated sample was lyophilized and redissolved in 250 μl of 20 mM ammonium bicarbonate buffer (pH = 7.8). This solution was incubated at 37°C. 10 μl aliquots of chymotrypsin (1 mg/ml in 20 mM NH_4HCO_3) were added successively at 0, 45, and 85 minutes. The chymotrypsin was less than 3% of the sample by weight. After 180 minutes, the digestion was quenched by addition of one drop of glacial acetic acid. The digestion mixture was lyophilized, and redissolved in 50 μl of water. The entire sample was spotted on chromatographic paper.

Two-dimensional peptide maps of the digests were prepared by two-phase descending chromatography (acetic acid:1-butanol:water, 1:3.375:5, v/v, pH = 3.5) followed by electrophoresis (pyridine:acetic acid:water, 1:20:280, v/v, pH = 3.5; 53 volts/cm; 80 min.) (Bennett, 1967). The maps were lightly stained with ninhydrin spray (0.1% in acetone, w/v). The locations of the spots containing histidine peptides were found by amino acid analysis of peptides from control maps using similarly digested trypsin. The regions of the map containing histidine peptides were cut out and eluted with 10% formic acid. 25 nmoles of norleucine were added to each sample as an internal standard. 75% of each sample (375 μl in 10% formic acid) was added to 7 ml Aquasol and counted for a minimum of 300 minutes.

Peptides which did not contain histidine were treated identically and used to determine the background counts. The remaining 25% of each sample was lyophilized and analyzed for amino acids.

The origin of each peptide map was similarly analyzed to quantitate total tritium incorporation into trypsin. For each origin sample, 50 nmoles of norleucine were added and 5% of the mixture was amino-acid analyzed; 75% was added to 7 ml Aquasol for scintilation counting.

The specific activity (cpm/nmole) of each sample was used to calculate a first-order rate constant for exchange using the expression:

$$k = -\frac{1}{t} \ln \left(\frac{e - sa}{e}\right) \qquad (1)$$

where sa is the observed specific activity after an incubation of time t (38 hours), and e is the specific activity of the tritiated water used in the incubation (186 cpm/nmole). e corresponds to the expected specific activity of the histidines in the protein when equilibrium with the buffer is reached.

In a separate exchange experiment, 10 mg of trypsin was incubated at 37°C in pH 7.14 buffer for fourteen days. The specific activity of the incubation solution (225 µl) was 1840 cpm/nmole. The

exchange was quenched and the protein separated from the buffer and from the bulk of the tritiated water as described above. 4 mg of the tritiated trypsin was divided into two portions. 80% of the sample (3.2 mg) was mixed with 30 mg of purified STI in 1 ml of 0.1 M NH_4HCO_3, pH 7.8, and the STI-trypsin complex was purified on a Sephadex G-75 column (4.91 cm^2 x 120 cm; buffer: 0.1 M NH_4HCO_3, pH 7.8) at 4°C. The complex was then hydrolyzed in 6 N HCl and the individual amino acids were separated by high voltage paper electrophoresis (pH 1.7, 7600 volts, ~350 mA, 2 hrs. (Dreyer and Bynum, 1967)). A standard amino acid mixture was co-electrophoresed. The positions of the amino acids in the standard mixture were located by spraying with a solution of ninhydrin (0.3% in acetone, w/v). The histidine sample was eluted from the paper with 10% formic acid and lyophilized. 50 nmoles of norleucine were added in 250 µl of H_2O and 10% of this mixture was analyzed to determine the histidine content. The specific activity was determined by adding 200 µl of the sample to 7 ml Aquasol and counting for 60 minutes in a scintillation counter. Arginine from the hydrolysate was treated similarly and used to determine the background counts.

The remaining 0.8 mg of the tritiated trypsin was hydrolyzed and the specific activity of the isolated histidine was determined as described above for the STI-trypsin complex.

Results

After incubating bovine trypsin in various buffer solutions containing tritiated water between pH 2 and pH 9, the labeled protein was separated from the bulk of the tritiated incubation buffer by gel chromatography (Figure 1). The majority of the tritium in the protein at this stage is due to groups other than the C-2 of histidine.

In order to separate the three histidines of trypsin, each sample was enzymatically digested, and the peptides were separated by chromatography and electrophoresis. Figure 2 is a photograph and a tracing of the ninhydrin staining pattern which was typical of all of the two-dimensional peptide maps. The isolated peptides were products of digestion by both the added chymotrypsin and the fraction of the trypsin in the sample which was still active after boiling. The three histidines were found in different peptides: one containing His 40, one containing His 57, and two peptides of overlapping sequence containing His 91. The amino acid compositions of these peptides were used to determine their locations in the known sequence of bovine trypsin (Table I).

Part of each histidine peptide sample was used to count the incorporated tritium, and part was used to quantitate the amount of histidine present. The specific activities determined from these measurements were then used to calculate pseudo-first-order rate

constants for the exchange, k, according to Equation 1. The rate constants for each histidine at each incubation pH are listed in Table II. These rate constants were used to extract pK_a values for the imidazole side chains by least-squares fitting to a theoretical curve based on the exchange mechanism proposed by Vaughan et al. (1970). This mechanism involves the rate-determining formation of an ylide intermediate through abstraction of a proton from an imidazolium cation by water (k_a) or by hydroxide (k_b) ion (Figure 3), and leads to the following rate equation:

$$k = \left(\frac{k_b[H^+][OH^-] + k_a[H_2O][H^+]}{K_a + [H^+]} \right) \qquad (2)$$

where k is the pseudo-first-order rate constant and K_a is the ionization constant for imidazole.

The experimental data for the pH dependence of the exchange rates for histidines 40, 57, and 91, together with the least-squares fits to the above rate equation, are illustrated in Figure 4. The least-squares determined values of k_a, k_b and pK_a for each histidine are listed in Table III. k_b is the rate constant for the reaction pathway catalyzed by OH^-, and determines the maximum exchange rate which will be observed at high pH. The values of k_b for the histidines of trypsin are much smaller than for the histidines in ribonuclease or

lysozyme, or for small molecules which contain imidazole (Table III). The best fit to the data for His 57, the active site histidine, gives an apparent pK_a of 6.55. Because the specific activities for histidines 40 and 91 were only slightly above background, the pK_a assignments for these groups are tentative. The data suggest that exchange into His 40 may depend on two pK_a's.

Tritium exchange rate data can frequently be fit by a simple function which describes the ionization of a single group; for example,

$$k' = \frac{C \cdot 10^{(pH-pK_a)}}{1 + 10^{(pH-pK_a)}} \tag{3}$$

or

$$k' = C \cdot \left(1 - \frac{10^{(pH-pK_a)}}{1 + 10^{(pH-pK_a)}}\right) \tag{4}$$

where C is a constant. The pK_a's determined from such a fit will be similar, but not identical, to the apparent pK_a's calculated by fitting the rate Equation 2 derived from the ylide mechanism of exchange to the data (Table III).

The pH-exchange profile for intact trypsin isolated from the origins of the peptide maps is shown in Figure 5. The extent of tritium incorporation into the intact molecule was somewhat greater

than the sum of the counts incorporated into all three histidines. The difference is presumably due to incorporation of tritium into other groups in the intact molecule, which also back exchange slowly. The rate constants for intact trypsin exchange are also listed in Table III. The origin exchange exhibits an apparent pK_a of 6.9.

Because the tritium-exchange rates for the histidines in trypsin are very low, lower than any previously observed, only a maximum of 1% of the trypsin had undergone exchange during the 38-hour incubation. This low incorporation raised the question of whether only denatured trypsin had undergone exchange at the histidines because the trypsin used in the experiment was only ~95% active. To eliminate this possibility, a second exchange experiment was performed. Water with higher specific activity (1840 cpm/nmole) and a longer incubation time (14 days) were used to increase the amount of label incorporated in the histidines. Soybean trypsin inhibitor was used to isolate "active" trypsin from inactive protein.

The trypsin labeled during the 14-day incubation at pH 7.14 was split into two portions. (Control experiments showed that the benzamidine in the incubation buffer completely inhibited autolysis.) One portion of the trypsin was converted to a trypsin-soybean trypsin inhibitor complex and was purified by gel filtration (Figure 6). Complex formation was used as an assay for the "active" or inhibitor-

binding trypsin in the sample. A comparison of the specific activities of the histidines in the uncomplexed and complexed trypsin (Table IV) demonstrates how much of the tritium label had been incorporated into active trypsin during the incubation. These results show that approximately 80% of the tritiated histidine was in active trypsin.

A comparison of the observed rate constant for the uncomplexed trypsin at pH 7.14 with the expected value based on the histidine exchange rates during the 38-hour incubation (Table IV) reconfirms the very slow rate of exchange seen in the 38-hour experiment, although the rates are not identical. The difference may be due to back exchange of tritium out of imidazole during the 20-hour acid hydrolysis (110°C). The samples in the 14-day experiment were acid hydrolyzed before scintillation counting, whereas the samples which were counted in the 38-hour experiment had not been acid hydrolyzed. An "observed" rate constant for back exchange during the acid hydrolysis step can be calculated from the expected value of the specific activity, 193.1 cpm/nmole (Table IV), and the observed value after acid hydrolysis, 87.5 cpm/nmole, according to Equation 5:

$$k_{back} = \frac{-1}{t} \ln \left(\frac{87.5}{193.1}\right) \qquad (5)$$

where $\underline{t}$ is the time of hydrolysis (20 hrs.). This "observed" rate

constant is 39.6×10^{-3} hour^{-1}.

Alternatively, the rate constant for back exchange during acid hydrolysis can be roughly estimated from the rate of deuterium exchange at 65°C and pH 0.26 (0.090 hour^{-1}; average value from Vaughan et al., 1970) by applying corrections for the tritium isotope effect $\left(\frac{k_H}{k_T} \approx 18.2\right)^3$ and the temperature dependence of the rate.[4] The estimated value based on this calculation was

---

[3] This value is estimated from the deuterium isotope effect of 7.5 (Markley and Cheung, 1973) using:

$$\log \frac{k_H}{k_T} = 1.44 \log \frac{k_H}{k_D} \quad \text{(Jencks, 1969).}$$

[4] The data of Vaughan et al. (1970) for deuterium exchange at pH 10.56 at 60°, 65°, and 70°C were fit to the Arrhenius equation:

$$\ln k = -\frac{E_a}{RT} + \ln A.$$

The value of k at T = 110°C was estimated from this fit. The data of Vaughan et al. (1970) suggest that the temperature dependence of k may vary with pH, so that the extrapolated value of k at 110°C in 6 N HCl is only a rough estimate.

46×10^{-3} $hour^{-1}$. Because of the similarity between the "observed" and estimated rates of back exchange, it is reasonable to assume that the difference between the rate constants for the 14-day and 38-hour experiments (Table IV) is due to back exchange during the acid hydrolysis of the 14-day sample. If this is the case, then the rate constants for the histidines in ribonuclease and lysozyme (Table III) may be too low because the samples were hydrolyzed before counting (Matsuo et al., 1972; Ohe et al., 1974).

Because of the slow exchange rates, there was some question as to which proton on His 57 was exchanging. Although ^{2}H or ^{3}H exchange for the C-2 proton of imidazole is normally faster than exchange with the proton in the C-4 position (Matsuo et al., 1972), the C-2 proton of His 57 in trypsin is relatively solvent inaccessible while the C-4 proton is solvent accessible (Stroud et al., 1971). Therefore, it seemed possible that the C-4 rather than the C-2 protons of His 57 might have exchanged during the incubations at 37°. Isotope exchange at the C-4 position of imidazole at 180°C had previously been studied (Vaughan et al., 1970). To determine the relative rates of exchange at the C-2 and C-4 positions under conditions similar to those used in the tritium exchange experiment, deuterium exchange into L-histidine at 37°C was investigated using NMR

spectroscopy. The time dependence of the deuterium exchange is shown in Figure 7. The first-order rate constant for C-2 exchange was calculated by fitting an exponential to the data. The associated half-time of exchange, 2.8 days, is compared to the exchange rates for other histidines in Table II. In contrast to the C-2 exchange, there was essentially no C-4 exchange for up to 115 days of incubation. Thus, the tritium exchange into His 57 of trypsin, which occurs with an apparent pK_a of 6.55 and a half-time of 73 days, takes place at the C-2 position.

Discussion

All previously observed pH-rate profiles for ^3H exchange at the C-2 of histidine show an increase in the rate of exchange as the imidazole titrates (Vaughan et al., 1970; Matsuo et al., 1972; Ohe et al., 1974). These profiles look similar to a pH titration curve and were so treated by Ohe et al. (1974). One might expect that if the ylide mechanism of exchange (Figure 3) were valid, the rate of exchange would decrease as the concentration of imidazolium ion decreased with increasing pH. The exchange rate, however, depends not only on the imidazolium ion concentration, but also on the hydroxide ion concentration. Because of this dependence, the pseudo-first-order rate constant, k, can either increase, decrease, or remain unchanged as the pH

increases. Its behavior will depend upon the ratio of the hydroxide-mediated (k_b) and the water-mediated (k_a) rates of ylide formation. This dependence is illustrated in Figure 8 where the pH-rate profiles for a fixed pK_a and varying $k_a:k_b$ ratios are plotted. The curves in the figure, which were calcaulated using Equation 2, clearly show that the water-mediated exchange rate controls the low pH region of the profile while the high pH region depends on the hydroxide-mediated exchange. The $k_a:k_b$ ratios for His 40 and His 91 are sufficiently small that the decrease in imidazolium ion concentration appears to dominate the overall exchange and the pH dependence is similar to that seen in curves 4 and 5 of Figure 8. His 57 follows the trend set in previous studies and k increases between pH 6 and 8.

The data for the exchange into His 40 (Figure 4, Table II) suggest that a second pK_a near 2.7 affects the exchange of this group. Either the high pK_a (6.9) or the low pK_a (2.7) could be that of the imidazole ring of His 40 while the other apparent pK_a could be due to a group or groups which perturb the exchange reaction. The curve representing the pH dependence of tritium exchange into His 40 in Figure 4 was calculated by fitting the data to an equation for the ylide exchange mechanism (Equation 2) which includes a term to account for a perturbation by an ionizable group:

$$k = \frac{10^{(pH-pK_a)}}{1 + 10^{(pH-pK_a)}} \cdot \left(\frac{k_b[H^+][OH^-] + k_a[H_2O][H^+]}{K_a + [H^+]} \right) \quad (6)$$

Because the specific activities for His 40 were only slightly above background, the conclusion that the exchange is affected by a second pK_a is tentative. If the apparent pK_a of 2.7 does represent an ionizing group which perturbs the exchange, Asp 194 is a likely candidate for the assignment of this pK_a. Asp 194 is near to His 40 in the three-dimensional structures of both DIP- and BA-trypsin (Stroud et al., 1971; Krieger et al., 1974), the adjacent carbonyl group of Gly 193 is hydrogen bonded to the imidazole of His 40, and infrared spectroscopic studies of trypsin show that the pK_a of Asp 194 is near 2.9 (Koeppe and Stroud, 1975).

The histidines in trypsin exhibit the slowest imidazole tritium exchange rates yet reported. The kinetic parameters for exchange into the histidines of trypsin, lysozyme, ribonuclease, N-acetyl-L-histidine and L-histidine are listed in Table III. These parameters were derived by fitting Equation 2 to the rate data. The rate constants for the hydroxide-mediated exchange at the imidazoles of trypsin are approximately one to four orders of magnitude slower than those of the other examples. The rate constants for the water-mediated exchange listed in the table are based on the very slow

exchange which occurs in the low pH region, and are too uncertain to be used for a detailed comparison. The half-time for exchange, $t_{\frac{1}{2}}$, at pH 9 (Table III) depends both on the hydroxide-mediated rate constant (k_b) and on the concentration of imidazolium ion. The pK_a's of the histidines of ribonuclease are lower than the pK_a of N-acetyl-L-histidine; therefore, the values of $t_{\frac{1}{2}}$ for the histidines in ribonuclease are longer than the $t_{\frac{1}{2}}$ for N-acetyl-L-histidine even though their k_b values are up to 12 times larger.

The differences in the hydroxide-mediated rate constants can generally be understood in terms of the solvent accessibility, orientations, and mobilities of the imidazole rings.

Ohe et al. (1974) have shown that there is a correlation between the exchange rates for His 119, His 105, and His 12 of ribonuclease and the static solvent accessibilities of their side chains (Lee and Richards, 1971). While solvent accessibility must be crucial in determining the exchange rate, the orientation and mobility of the imidazole ring are also important. The side chain of His 48 is more solvent accessible than that of His 12, but His 12 exchange is six times faster at pH 8. His 48 exchange is slower because its side chain orientation at the surface of the molecule does not permit solvent to attack the C-2 position in the plane of the imidazole ring. This interpretation is based on a wire model of ribonuclease S which

which was built using the coordinates of Wyckoff et al. (1970). The structure suggests that water or hydroxide must approach in a direction normal to the ring, and thus proton extraction and addition are expected to be energetically unfavorable. Also, from an examination of the Kendrew model, one would conclude that the imidazole ring cannot reorient to facilitate C-2 exchange by simple, low-energy rotation about either the C_α-C_β or C_β-C_γ bonds. A more substantial conformational change would be required. The side chains of His 119, 105 and 12 in the model all extend edgewise into the solvent so that exchange takes place much more readily on these rings.

The C-2 exchange rate for His 40 in trypsin can also be explained by considering the accessibility, orientation and flexibility of the histidine side chain. The following analysis is based upon the partially refined 1.76 Å resolution structure of bovine trypsin inhibited with diisopropylfluorophosphate (Chambers and Stroud, in preparation). His 40 is buried beneath the surface of the protein and is quite inaccessible to solvent. The imidazole ring is surrounded by the side chains of Trp 141, Ser 32, and the peptide backbone of residues 41-43, 142-143, and 193-194. The ϵ-nitrogen of His 40 is hydrogen bonded to the carbonyl group of Gly 193, and the C-2 proton points toward the center of the molecule. Thus, the exceptionally slow tritium exchange rate at His 40 is

consistent with the three-dimensional structure of trypsin. The side chain of His 91 lies in a depression on the surface of trypsin; however there is no obvious explanation for why the exchange is very slow.

The slow exchange rate of His 57 can be explained by considering its local environment in the active site of benzamidine-inhibited trypsin (Krieger et al., 1974). The plane of the imidazole ring of His 57 is approximately normal to the surface of the protein. The N-1 and N-3 nitrogens are hydrogen bonded to the carboxylate group of Asp 102 and the side chain of Ser 195, respectively. The C-2 proton points toward the center of the molecule and is sequestered from solvent by the Cys 42-58 disulfide bridge and the backbone of residues 57-58 above the ring and by Leu 99 and the backbone of 214-215 from below. One might expect that tritium exchange at the buried C-2 position of His 57 in the active conformation of the enzyme would be much slower than the exchange of His 91 because of greater steric hindrance. Exchange into the solvent-accessible C-4 position of the ring seems very unlikely because there was no observable incorporation of deuterium into the C-4 position of L-histidine in D_2O after a 115-day incubation at 37°C, pD 8.2.

The imidazole of His 57, however, can swing out into solution by disrupting the hydrogen bonds to Asp 102 and Ser 195 and by rotating about the C_α-C_β bond. Precedence for such a conformational

change can be found in the structure of silver-diisopropylphosphoryl trypsin. The silver ion induces a rotation of the imidazole about the C_α-C_β bond with little perturbation of the rest of the structure (Chambers et al., 1974).

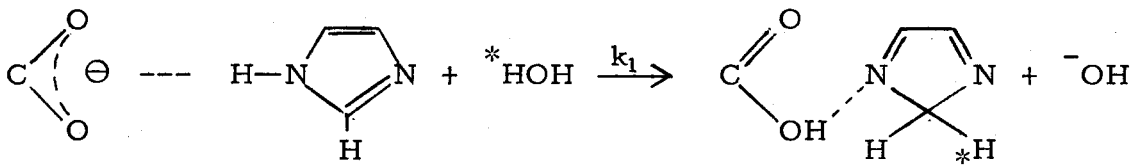
If the side chain of His 57 is involved in a conformational equilibrium between the active "in" position and an inactive "out" position, it might only exchange when it is in the "out" conformation. Exchange into His 57 in the "in" position seems unlikely because of the structure of the active site and the stereochemistry of the reaction. The exchange rate in the "out" position could be similar to that for small molecules containing imidazole; for example, N-acetyl-L-histidine, L-histidine, or for His 15 in denatured lysozyme (Table III). The overall slow rate of exchange could be explained by an "in-out" equilibrium in which only a small percentage of the trypsin would have His 57 populating the "out" conformation. The tritium exchange would then detect the pK_a of His 57 in the "out" position rather than in the active conformation.

Because of the long duration of the tritium exchange experiment, the exchange due to the "in-out" swinging of the histidine imidazole ring cannot be distinguished from the slow exchange of the "in" form. This ambiguity is a characteristic drawback of a reaction which is as slow as tritium-hydrogen exchange at the C-2 position of histidine,

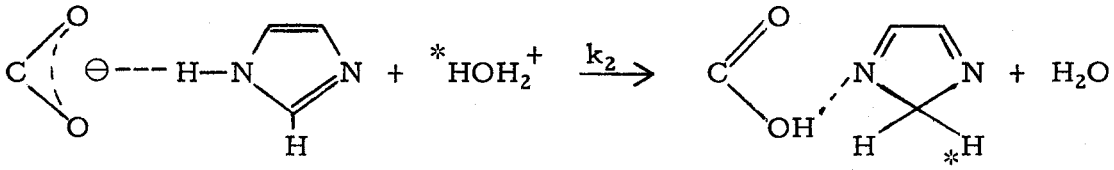
and is an especially serious problem for the case of His 57 of trypsin, which exchanges at a very slow rate.

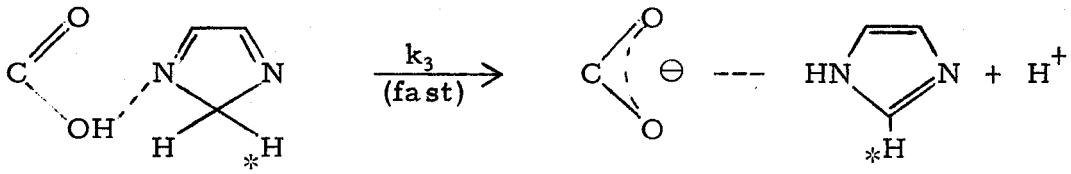
In addition to the ylide mechanism, there are other possible explanations for the tritium exchange into His 57. Of particular interest are mechanisms in which the ionization of a neighboring group affects the rate of tritium exchange of the imidazole. These mechanisms could be of two types: 1) those in which the "primary" pH dependence at pH 6.55 is due to the titration of the neighboring group, and the titration of the imidazole itself is either not seen at all or is seen merely as a small perturbation on top of the "primary" effect; and 2) those in which the ylide mechanism is operative with the "primary" titration at pH 6.55 due to the imidazole titration, but with an additional perturbation in the pH dependence due to the titration of a neighboring group.

The following mechanism is an example of the first possibility--that of a neighboring group titrating with a pK_a near 6.55. In this mechanism, neutral imidazole rather than positively-charged imidazolium is the reacting species, and neighboring group base catalysis assists in the proton transfers.



"iso-
imidazole"





In this scheme the base is represented as a carboxylate group.

The rate equation (Eq. 7) which can be derived from this "iso-imidazole" mechanism has the same form as Eq. 2, the rate equation for the ylide mechanism:

$$k_{iso} = \frac{A_1 + A_2[H^+]}{K_a + [H^+]} \qquad (7)$$

For the ylide mechanism, K_a is the ionization constant of the imidazole, whereas in the "iso-imidazole" mechanism K_a is the ionization

constant of the neighboring base. Although terms A_1 and A_2 represent different collections of constants for the two cases, the mechanisms cannot be distinguished on the basis of the rate equations.

For free histidine in solution, the rate-determining step has been shown to involve removal of a proton (Markley and Cheung, 1973). This is compatible with the ylide mechanism but not with an "iso-imidazole" mechanism which involves rate-determining protonation at the C-2 of imidazole. However, the exchange mechanism in the enzyme could possibly be different from histidine free in solution.

Regardless of which of the above mechanisms is preferred, it is possible that an additional pK_a might perturb the pH dependence of exchange. For the ylide mechanism one might expect the ionization of Asp 102 to affect the exchange rate, while the "iso-imidazole" mechanism would depend on the ionization of imidazole itself as well as on the neighboring proton acceptor (Asp 102). To test whether the data were consistent with this possibility, the kinetic equations applicable to either of the mechanisms (Equations 2 and 7) were modified to incorporate a simple additional positive or negative titration curve (see Equations 3, 4, and 6). No acceptable fit was found when the perturbations were included in the least-squares analysis. The fitting led to either unreasonable rate constants for the individual steps or physically impossible (negative) K_a's. This conclusion is

valid only for the pH region 5 to 9 because of the paucity of data below pH 5 and the uncertainty in the very low-specific activities of the low pH data. The data points for His 57 between pH 3.5 and 1.7 in Figure 4 indicate a downward trend which could be consistent with a second pK_a in that region, although uncertainties in the data do not permit this conclusion to be drawn. The tritium exchange data for His 57 in bovine trypsin in the pH range of 5-9 can be explained by the titration of only one ionizable group which affects the exchange rate. This ionization could be due either to His 57 with the "in" or with the "out" conformation, or to Asp 102.

Conclusion

Several studies have provided evidence that Asp 102 is the group at the active site of the serine proteases which has a pK_a near 6.7 (Hunkapiller et al., 1973; Koeppe and Stroud, 1976). The pH dependence of the kinetics of substrate hydrolysis by the serine proteases shows that only the single pK_a of 6.7 affects the rate of hydrolysis between pH 2 and pH 8 (Kezdy and Bender, 1964; Fersht and Renard, 1974).

This would lead to the conclusion that His 57 of trypsin has a pK_a below 2 in the active conformation. The tritium-exchange data would be consistent with this conclusion if histidine in the active form

had a low pK_a and did not measurably exchange, but had a pK_a of 6.55 when swung out into solution in the exchangeable form. This interpretation seems to be the most reasonable one because the "in" conformation of His 57 appears completely inaccessible to proton exchange at the C-2 position, and because the imidazole ring can rotate out into solution with minimal disruption of the rest of the structure. The base-assisted "iso-imidazole" mechanism of exchange with the ring "in" would be consistent with the pK_a of 6.55 assigned to Asp 102, but is inconsistent with the known isotope effect in C-2 exchange for free histidine in solution (Markley and Cheung, 1973), and is unlikely because of steric hindrance.

Acknowledgements

We thank Dr. L. Hood for allowing us to use his chromatographic and electrophoretic equipment, Dr. J. Baldeschwieler for the use of a scintillation counter, and Drs. M. A. Raftery and W. Dreyer for use of their amino acid analyzer. The NMR spectrometer is a National Science Foundation supported (GP-8540) Southern California facility, and we thank Dr. D. Live and Mr. N. Petersen for advice on use of the NMR spectrometer.

References

Bender, M. L. and Kezdy, F. J. (1964), J. Amer. Chem. Soc. 86, 3704-3714.

Bennett, J. C. (1967), Methods in Enzymology 11, 330-339.

Chambers, J. L., Christoph, G. G., Krieger, M., Kay, L. M., and Stroud, R. M. (1974), Biochem. Biophys. Res. Commun. 59, 70-74.

Chase, T. Jr. and Shaw, E. (1967), Biochem. Biophys. Res. Commun. 29, 508-514.

Cuatrecasas, P. (1970), J. Biol. Chem. 245, 3059-3065.

Dreyer, W. J. and Bynum, E. (1967), Methods in Enzymology 11, 32-39.

Fersht, A. R. and Renard, M. (1974), Biochemistry 13, 1416-1426.

Glasoe, P. K. and Long, F. A. (1960), J. Phys. Chem. 64, 188-190.

Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. R., and Richards, J. H. (1973), Biochemistry 12, 4732-4743.

Inward, P. W. and Jencks, W. P. (1965), J. Biol. Chem. 240, 1986-1996.

Jencks, W. P., Catalysis in Chemistry and Enzymology, p. 248. McGraw-Hill, New York, 1969.

Kezdy, F. J. and Bender, M. L. (1964), J. Amer. Chem. Soc. 86, 937-938.

Koeppe, R. E. II and Stroud, R. M. (1976)--Submitted to Biochemistry.

Krieger, M., Kay, L. M., and Stroud, R. M. (1974), J. Mol. Biol. 83, 209-230.

Lee, B. and Richards, F. M. (1971), J. Mol. Biol. 55, 379-400.

Light, A. and Liepnieks, J. (1974), Methods in Enzymology 34, 448-451.

Mares-Guia, M. and Shaw, E. (1965), J. Biol. Chem. 240, 1579-1585.

Markley, J. L. and Cheung, S.-M., Proceedings of the International Conference on Stable Isotopes in Chemistry, Biology and Medicine, Argonne, Illinois, May 1973. U.S. Atomic Energy Commission Conference, No. 730525, pp. 103-118.

Matsuo, H., Ohe, M., Sakiyama, F., and Narita, K. (1972), J. Biochem. 72, 1057-1060.

Ohe, M., Matsuo, H., Sakiyama, F., and Narita, K. (1974), J. Biochem. 75, 1197-1200.

Robinson, N. C., Tye, R. W., Neurath, H., and Walsh, K. A. (1971), Biochemistry 10, 2743-2747.

Schutte, L., Provó Kluit, P., and Havinga, E. (1966), Tetrahedron Suppl. 7, 295-306.

Stroud, R. M., Kay, L. M., and Dickerson, R. E. (1971), Cold Spring Harbor Symposia on Quantitative Biology 36, 125-140.

Titani, K., Ericsson, L. H., Neurath, H., and Walsh, K. A. (1975), Biochemistry 14, 1358-1366.

Vaughan, J. D., Mughrabi, Z., and Wu, E. C. (1970), J. Org. Chem. 35, 1141-1145.

Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B., and Richards, F. M. (1970), J. Biol. Chem. 245, 305-328.

TABLE I

Relative molar ratios of amino acids in chymotryptic-tryptic peptides of trypsin

Amino Acid	Peptide 40	57	91-A	91-B
Asx			1.84(2)†	
Thr			0.81(1)	
Ser		0.98(1)	2.29*(3)	1.42*(2)
Pro			1.47(1)	.89(1)
Gly	0.14	.23	.63	
Ala		1.95(2)	.76	
Val		1.34*(2)	1.00(1)	1.23(1)
Ile			1.26(1)	1.00(1)
Leu		.13	0.69(1)	
Tyr			1.00(1)	0.52(1)
Phe	1.00(1)			
His	0.68*(1)	1.00(1)	0.90(1)	0.76(1)
Lys		.11	0.39	
Sequence Assignment	His_{40} to Phe_{41}	Val_{52} to His_{57}	Ser_{88} to Leu_{99}	Ser_{88} to Tyr_{94}

†The numbers in parentheses are the values expected from the known sequence of bovine trypsin (Titani et al., 1975).

*The N-terminal residue is partially destroyed by ninhydrin.

TABLE II

The pseudo-first-order rate constants for the incorporation of tritium into the histidines of bovine trypsin

pH	$k \times 10^5$ (hr.$^{-1}$)[a]			
	His 57	His 40	His 91	Origin[b]
1.73	0.00	3.04	0.84	4.05
2.80	0.74	2.12	1.19	15.87
3.67	1.63	5.86	1.22	27.63
4.10				12.63
5.08	2.69	3.99	1.30	11.71
6.07	14.08		1.76	27.63
6.56	16.11	4.43	1.03	39.74
6.87	25.84		1.30	50.00
7.27		1.44	0.71	53.16
8.98	37.26	0.99	0.74	87.03

[a] The pseudo-first-order rate constants, k, were calculated using Equation 1 with t = 38 hours and e = 186 cpm/nmole. All incubations were at 37°C.

[b] These values reflect the exchange of an unknown number of protons and therefore are not pseudo-first-order rate constants.

TABLE III. Kinetic parameters for hydrogen isotope exchange at the C-2 position of imidazole.[a]

Sample		pK_a[b]	k_a $M^{-1}hr^{-1} \times 10^6$	k_b $M^{-1}hr^{-1} \times 10^{-5}$	$t_{\frac{1}{2}}$[c] (days)	'pK'[d]
Trypsin	His 40	6.9	0.4	0.0003	>1000	---
	His 57	6.6	0.2	0.1	73	6.5
	His 91	7.2	0.2	0.0004	>1000	---
Ribonuclease[e]	His 12	5.7	1.4	5.6	10	5.7
	His 48	5.7	0.3	1.0	58	5.7
	His 105	6.1	3.0	7.0	3	6.2
	His 119	5.7	1.6	11.7	5	5.7
Lysozyme,[f] native	His 15	5.3	22.0	90.0	1.6	5.2
Lysozyme, denatured	His 15	7.0	29.0	1.9	1.5	6.9
N-acetyl Histidine[f]		7.3	8.2	1.0	1.5	7.3
L-Histidine					2.8[g]	

TABLE III (continued)

Footnotes:

[a] Measurements for trypsin and L-histidine were made at 37°C; all others were at 36.5°C.

[b] pK_a, k_a, and k_b were determined by fitting the exchange data to Eq. 2 using non-linear least squares.

[c] $t_{\frac{1}{2}} = \frac{-\ln (0.5)}{k_9}$, where k_9 is calculated from Eq. 2 using the parameters in this table and a pH of 9.

[d] These values were determined by fitting the exchange data to the titration curve for a monobasic acid (Eq. 3).

[e] The pseudo-first-order rate constants which were used in calculating these parameters were taken from Figure 2 of Ohe et al. (1974).

[f] The pseudo-first-order rate constants were taken from Figures 2 and 3 of Matsuo et al. (1972).

[g] Half time for deuterium exchange at pD = 8.2.

TABLE IV

Tritium incorporation into the histidines of "active" trypsin[a] at pH 7.14 and 37°C

	Incubation time (days)	Specific activity (cpm/nmole)	Relative tritium incorporation (%)	k ($hr.^{-1} \times 10^5$)
Total trypsin sample	14	87.5[b]	100.0	14.5
STI binding component	14	72.5[c]	85.7	12.0
His 57 (expected value)[d]		193.1		33.0

[a] "Active" trypsin is operationally defined as that portion which is recovered complexed with STI after chromatography on Sephadex G-75 (see Figure 6).

[b] This value is per nmole of trypsin, assuming three histidines per molecule.

[c] This value is per nmole of trypsin-STI complex, assuming five histidines in the complex.

[d] These values were calculated from the data for the 38-hour incubation (Fig. 3, Table II). k was calculated using Eq. 2, a pH of 7.14 and the kinetic parameters in Table III. The specific activity (sa) was determined from Eq. 1 assuming t = 14 days and e = 1840 cpm/nmole.

Figure Legends

FIGURE 1. Separation of ^{3}H-trypsin from the bulk of the tritiated incubation buffer. After a 38-hour incubation at 37°C, the ^{3}H-trypsin was separated from the tritiated incubation buffer by gel chromatography on Sephadex G-25 (column dimensions: 4.91 cm^2 x 90 cm; buffer: HCl, pH 2.5; fraction volume: 6.5 ml; flow rate: 3 ml/min.). 50 µl aliquots of each fraction were added to 5 ml of Aquasol for scintillation counting. The arrows denote the elution volumes for control samples of trypsin and tritiated water. The pH of the incubation buffer used in this example was 7.27.

FIGURE 2. A pH 3.5 peptide map of trypsin digested by chymotrypsin and partial autolysis. Descending chromatography was followed by electrophoresis in which the positive pole was on the left and the negative pole on the right. The photograph (top) and tracing (bottom) are examples of the ninhydrin staining pattern typically seen in such maps. The histidine containing peptides (40, 57, 91A, 91B) were identified by amino acid analysis (see Table I).

FIGURE 3. The ylide mechanism of tritium exchange into the C-2 position of imidazole (Vaughan et al., 1970).

FIGURE 4. pH dependence of the pseudo-first-order rate constants for tritium exchange with the C-2 protons in the histidines of bovine trypsin: ● - histidine 40; ○ - histidine 57; Δ - histidine 91. The curves are least-squares fits to the data as described in the text.

FIGURE 5. pH dependence of the pseudo-first-order rate constants for tritium exchange into intact trypsin isolated from the origins of the two-dimensional peptide maps. The amino acid composition of each origin corresponded to the known composition of trypsin. The curve is a least-squares fit to the data as described in the text.

FIGURE 6. Purification of a complex of soybean trypsin inhibitor with tritiated trypsin. An approximately ten-fold excess of STI (30 mg) was mixed with 3.2 mg of ^{3}H-trypsin and the components of the mixture were separated by gel chromatography on Sephadex G-75 (column dimensions: 4.91 cm^2 x 120 cm; buffer: 0.1 M ammonium bicarbonate, pH 7.8; fraction volume: 6.5 ml; flow rate: 0.5 ml/min.; 4°C). 50 μl aliquots of each fraction were added to 5 ml of Aquasol for scintillation counting. The arrows denote the elution volumes of control samples of the complex, trypsin and STI. The shaded region indicates the fractions which were pooled and analyzed for tritium incorporation.

FIGURE 7. Time dependence of deuterium exchange into L-histidine. The exchange was monitored at pD 8.2 using 220 MHz NMR spectroscopy. The intensity of the C_α proton resonance was used as an internal standard. The ratio of the C-4 to the C_α intensities (○) remained constant throughout the 115-day incubation indicating that there was essentially no deuterium exchange with the C-4 proton. The ratio of the C-2 to the C-4 peak intensities (●) decreased exponentially with time (the solid line is a least-squares fit of an exponential to the data where $k = 1 \times 10^{-2}$ hr.$^{-1}$, $t_{\frac{1}{2}} = 2.8$ days).

FIGURE 8. Variation in the shape of the pH-exchange rate profile with changes in the hydroxide mediated exchange rate constant. The pseudo-first-order rate constant for hydrogen isotope exchange into the C-2 position of imidazole calculated using Equation 2 can either increase (curves 1 and 2), decrease (curves 4 and 5), or remain constant (curve 3) with an increase in the pH.

Figure 1

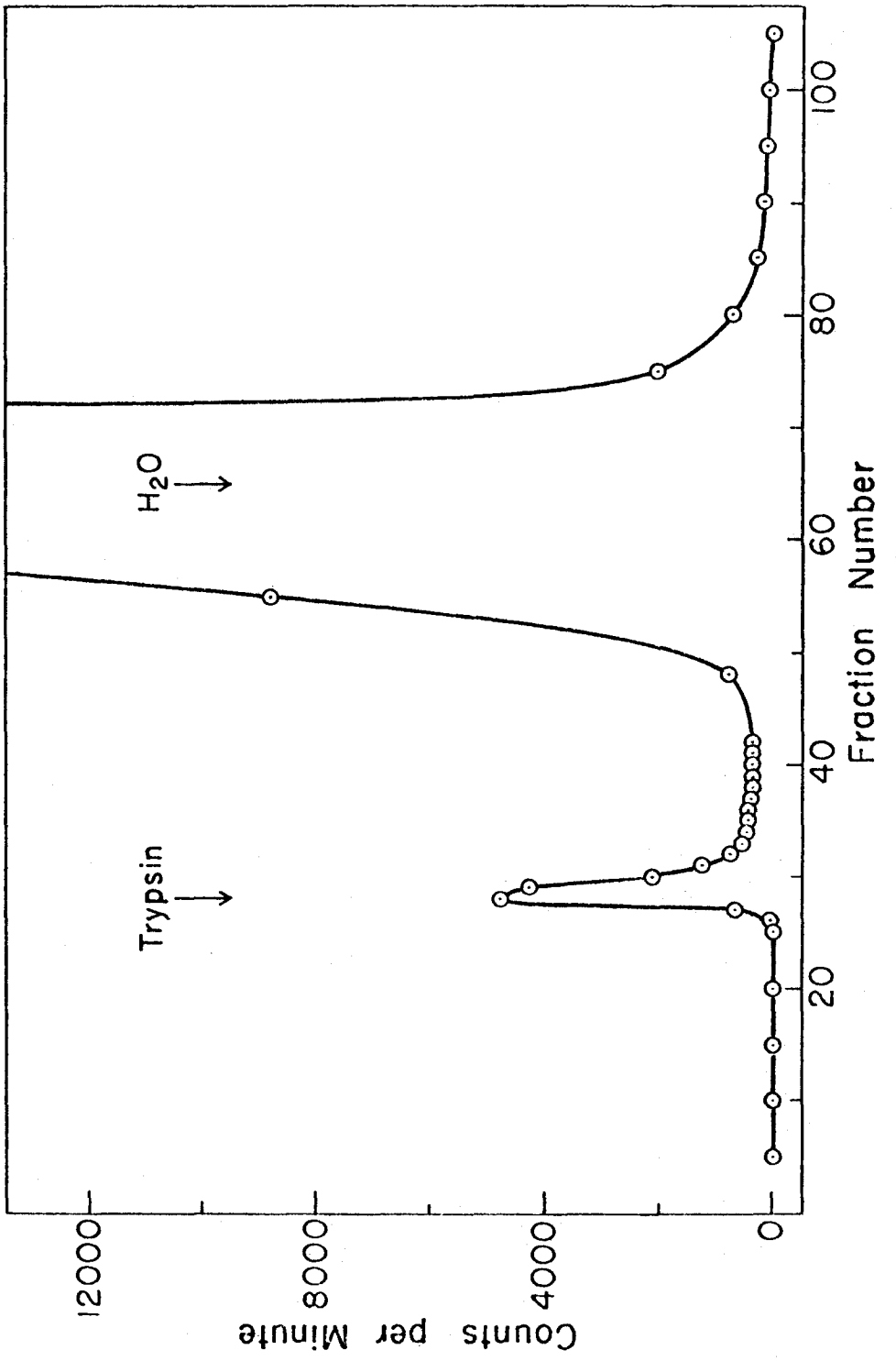

Counts per Minute
12000
8000
4000
0
20
40
60
80
100
Fraction Number
Trypsin
H_2O

Figure 2

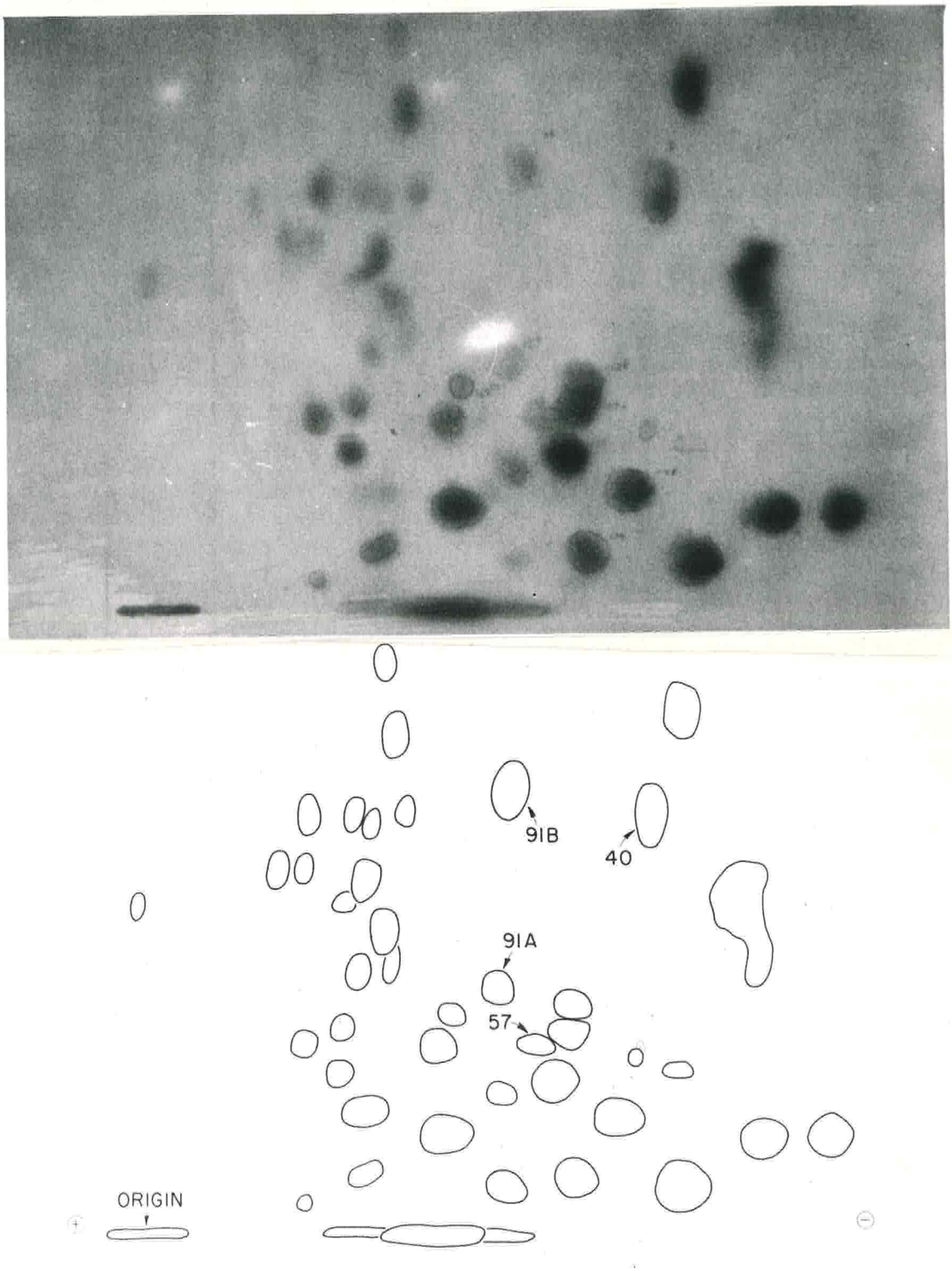


Figure 3

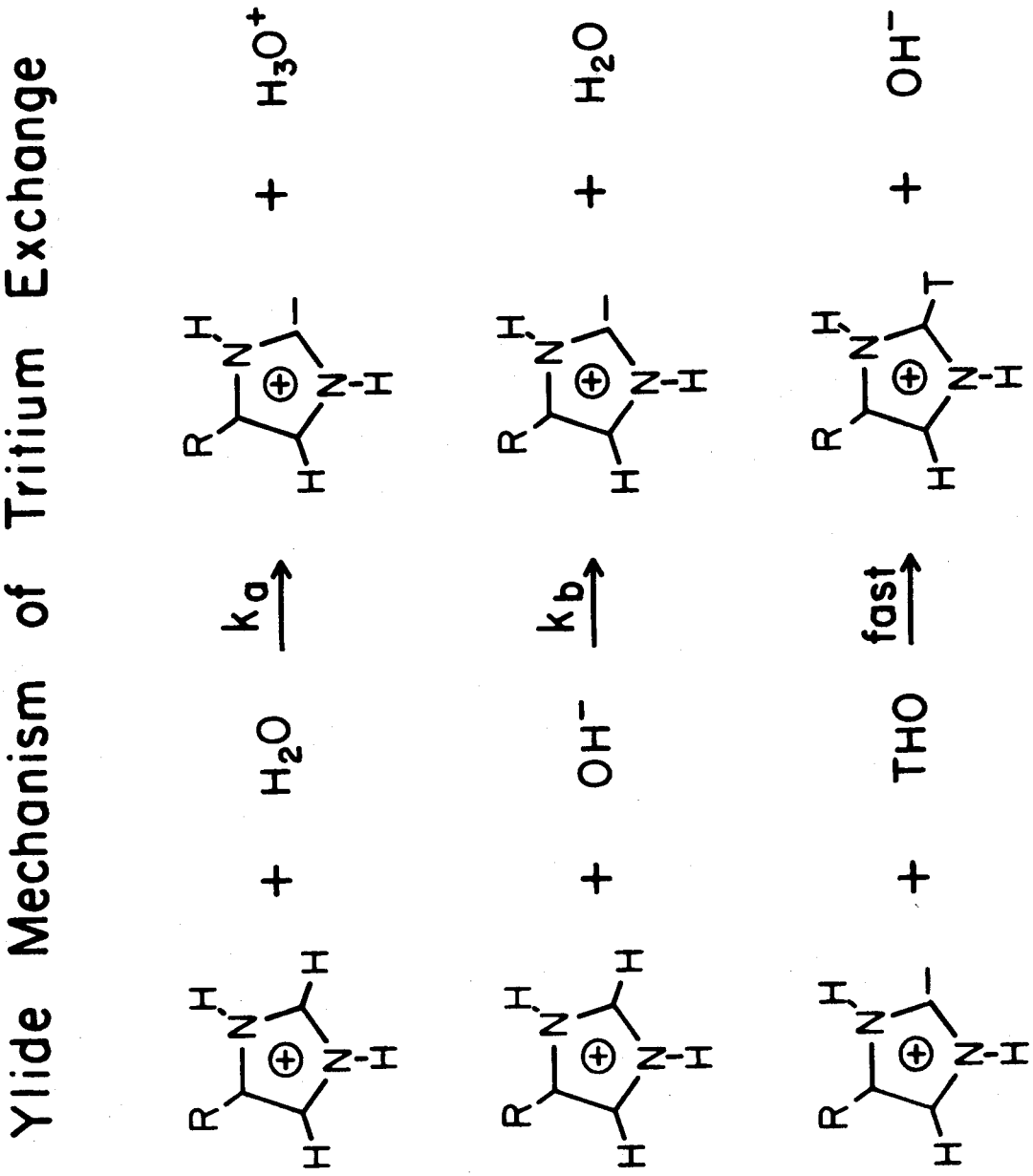

Ylide Mechanism of Tritium Exchange
H_3O^+
+
k_a
H_2O
+
H_2O
+
k_b
OH^-
+
OH^-
+
fast
THO
+

Figure 4

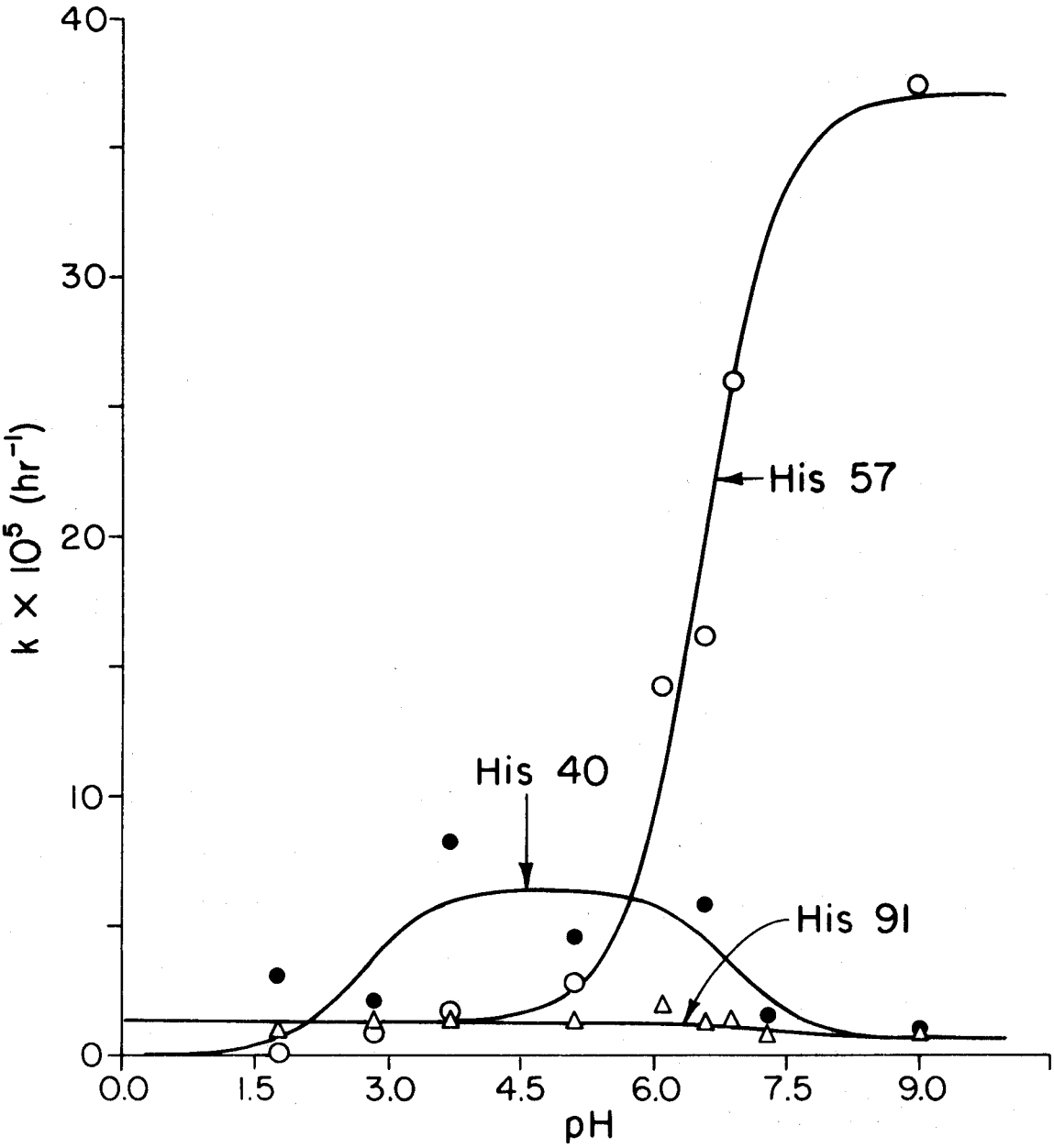


Figure 5

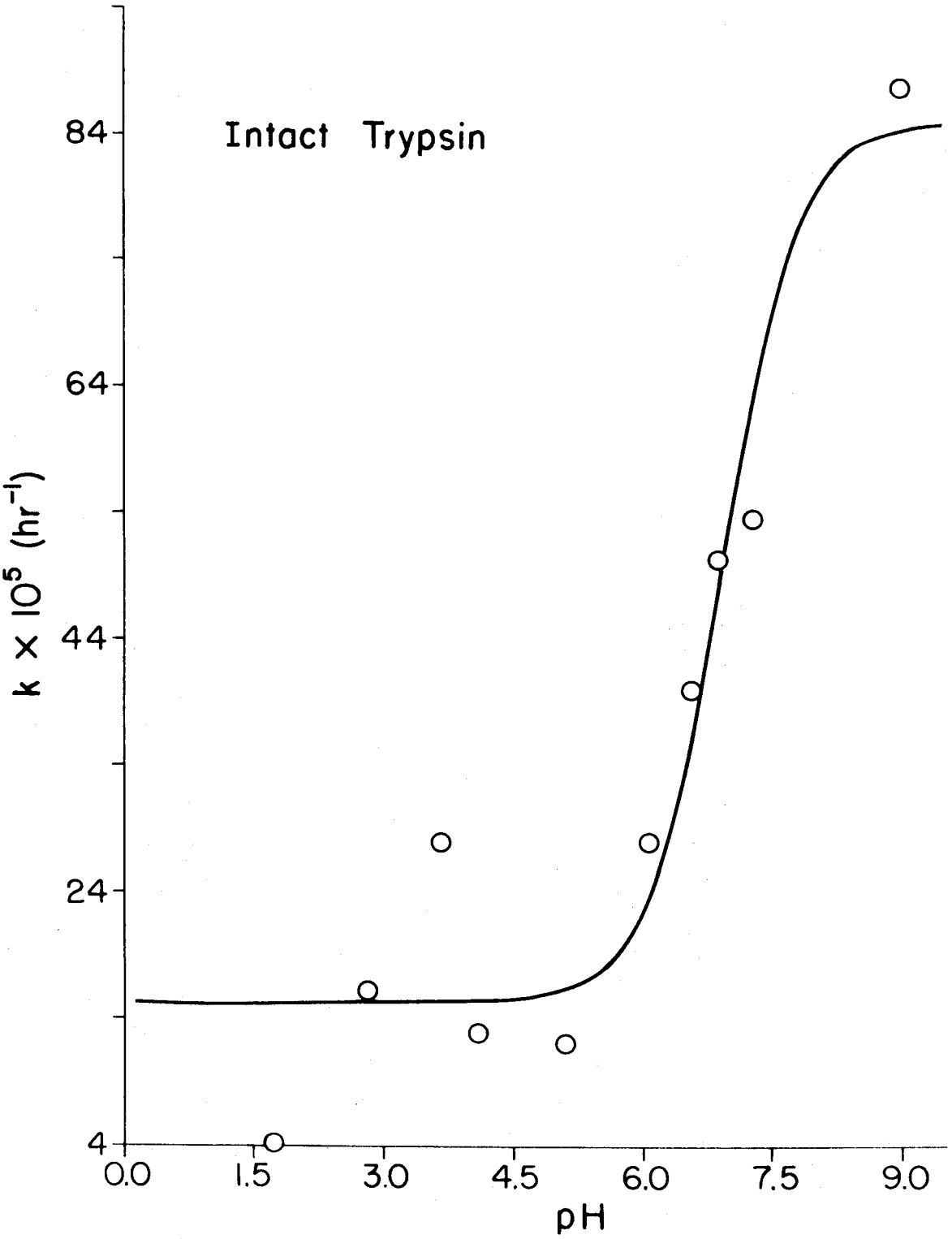


Figure 6

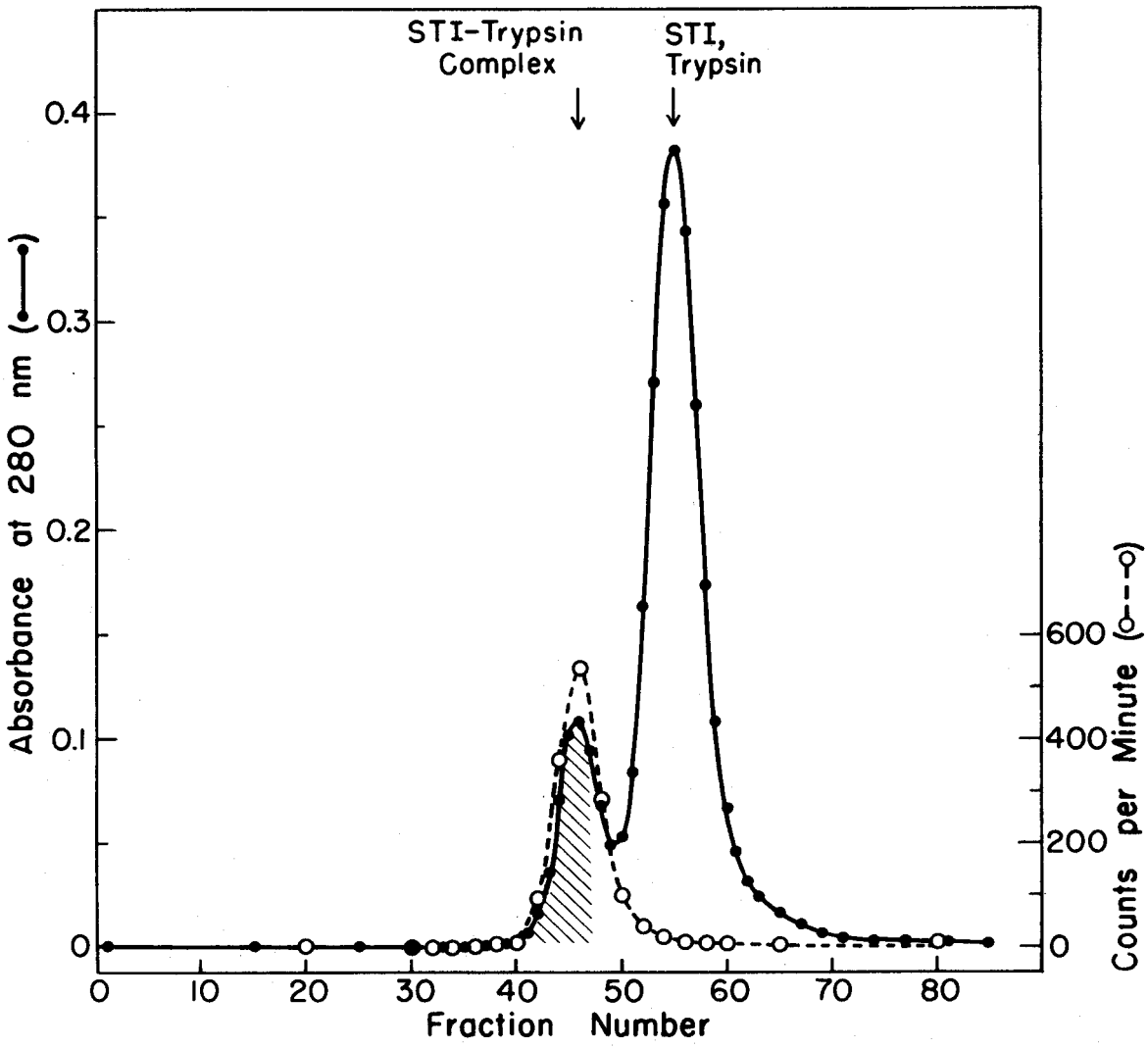

STI-Trypsin Complex
STI, Trypsin
Absorbance at 280 nm (●—●)
Counts per Minute (○--○)
Fraction Number
0.4
0.3
0.2
0.1
0
600
400
200
0
0
10
20
30
40
50
60
70
80

Figure 7

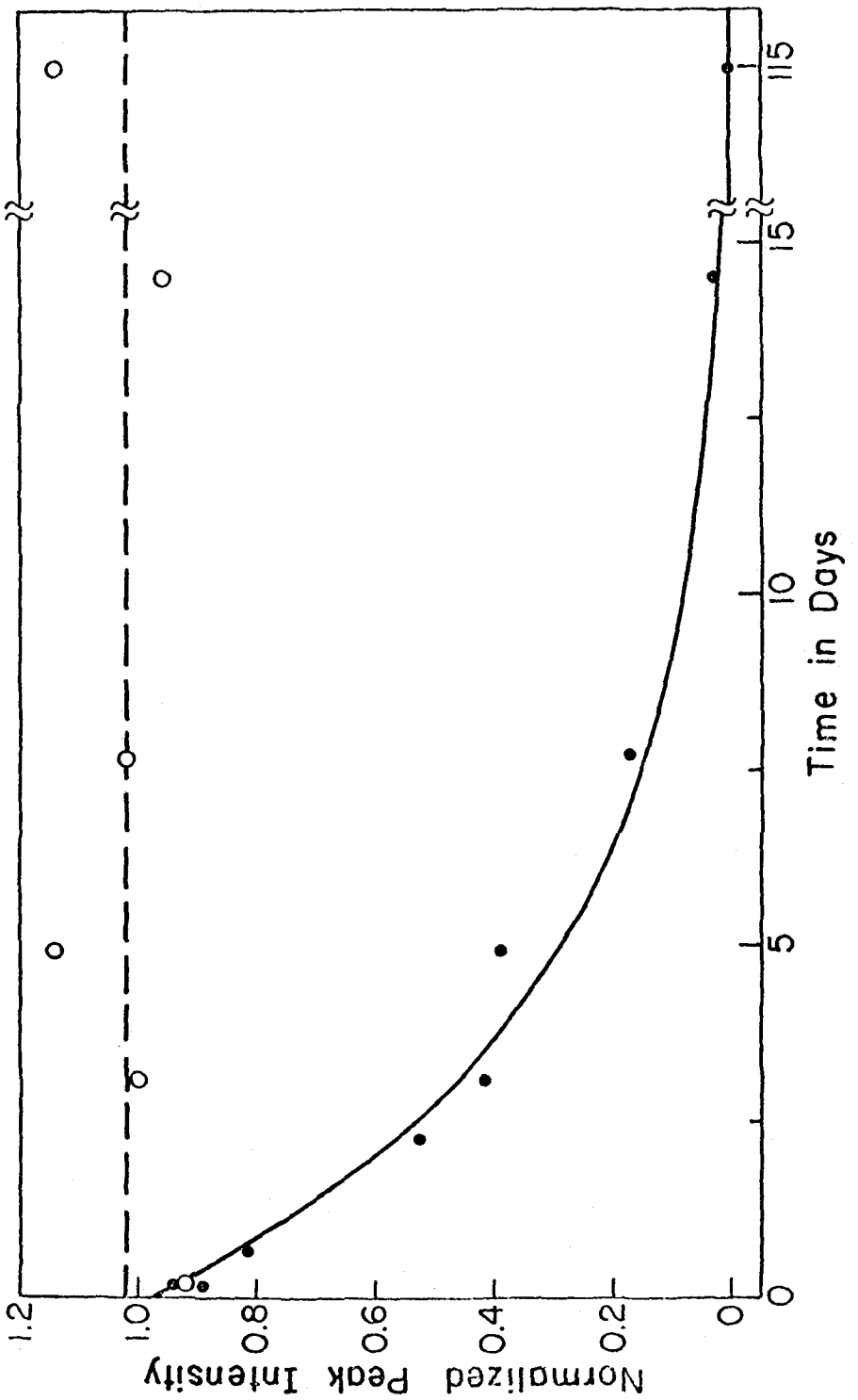
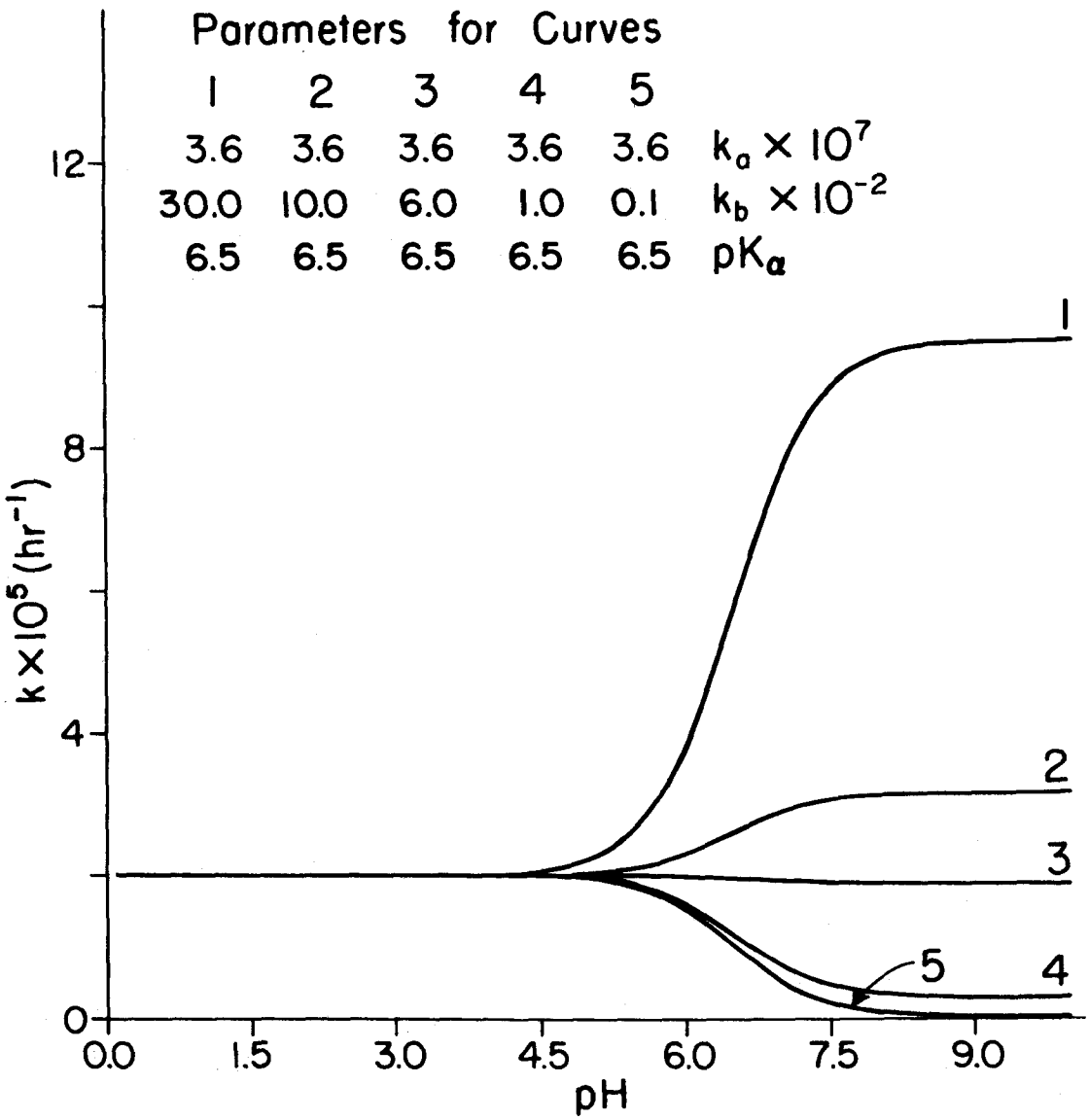

Normalized Peak Intensity
1.2
1.0
0.8
0.6
0.4
0.2
0
Time in Days
0
5
10
15
115

Figure 8


Parameters for Curves
1 2 3 4 5
3.6 3.6 3.6 3.6 3.6 $k_a \times 10^7$
30.0 10.0 6.0 1.0 0.1 $k_b \times 10^{-2}$
6.5 6.5 6.5 6.5 6.5 pK_a
$k \times 10^5 (hr^{-1})$
12
8
4
0
0.0
1.5
3.0
4.5
6.0
7.5
9.0
pH
1
2
3
4
5