SERINE PROTEASES: STRUCTURE, MECHANISM, AND ROLE IN EUCARYOTIC CELL TRANSFORMATION

Thesis by

Monty Krieger

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1976

(Submitted May 26, 1976)

TO MY PARENTS

ACKNOWLEDGMENTS

The colleagues of mine who collaborated on the work presented in this thesis are identified in the Appendices. Many persons at Caltech contributed directly to the work, and their particular contributions are acknowledged in the Appendices. Many other persons at Caltech contributed indirectly. Among them are Drs. Larry I. Grossman, Tom Broker, Hans-Peter Vosberg, Paul Johnson, Dick Vandlen, Sten Samson, Ray Teplitz, Michael Raftery, and Mrs. Marlyn Teplitz, who freely shared with me their knowledge of crystallography, biochemistry, or cell biology, and my fellow graduate students Mike Ross, Roger Koeppe II, Steve Spencer, John Chambers, Diane Kent, and Jerry Tobler.

The staff members of the Instrument and Mechanical Shops provided excellent technical support, and Braxton Evans and Fran Bennett always managed to procure promptly the items that were required.

Dee Barr patiently and carefully typed this thesis, including the drafts and manuscripts of the Appendices, and her advice and assistance has always been helpful.

My parents Mildred and Jay, my brother Terry, the other members of my family, my friends in Pasadena, and my friends Jeff

and Karen Greif and Julia Stroud at Caltech supported and encouraged me throughout my work.

I am grateful to all these people and to the Danforth Foundation which generously supported me and, through its conferences, enabled me to meet many persons who share my interest in teaching and learning, but I am grateful above all to my teacher, my colleague, and my dear friend, Robert Stroud.

ABSTRACT

The structure and catalytic mechanism of the serine protease trypsin was investigated using x-ray crystallography, enzyme kinetics and hydrogen isotope exchange. The role of serine proteases in eucaryotic cell transformation was examined by studying the binding of ^{125}I-labeled dog plasminogen to normal and SV40-transformed Balb/c 3T3 cells.

X-ray structural analysis using the difference Fourier technique showed that benzamidine, a non-covalent, reversible inhibitor of trypsin, binds in the specificity binding pocket of the enzyme and blocks substrate access to the active center. Benzamidine binding provided a reasonable model for the binding of specific side chains and permitted us to construct a model for specific substrate binding. The difference Fourier technique was also used to show that silver ions inhibit trypsin by binding between the aspartic acid and histidine side chains in the catalytic site. This observation was used to assign difference infrared spectral peaks to the side chain of Asp 102 in the catalytic site. A new level in understanding the catalytic mechanism was achieved using this information.

In the course of these crystallographic investigations, new, more efficient and accurate techniques of data collection were

developed based upon a detailed study of the sources of background radiation in x-ray diffractometry. These techniques involve new methods for reducing and accurately accounting for background effects.

A pH-dependent conformational change could introduce artifacts in the determination of the pK_a's of catalytic groups using either rapid kinetic or slower spectroscopic techniques. The rate of N_α-carbobenzoxy-L-lysine-p-nitrophenyl ester (CLNE) hydrolysis by trypsin over the range pH 2 to pH 5 was observed to test for such a conformational change. Pre-incubating the enzyme at pH 2 and pH 6.9 for up to three hours had no effect on the hydrolysis rate; therefore, within these limits, there was no pH-dependent conformational change which directly affected catalysis.

In an attempt to determine the pK_a of His 57 directly, the exchange of tritum with the C-2 protons in the histidine side chains of trypsin was measured. This exchange was found to be slower than any previously measured C-2 tritium exchange. The half-time for exchange into His 57 was 73 days, and this apparently took place when the active site imidazole had rotated away from its active conformation. The apparent pK_a of His 57 in this inactive "out" conformation was 6.6.

Increased proteolytic activity exhibited by many cells after transformation is due to plasminogen activation to the trypsin-like enzyme plasmin. Plasmin is responsible for a number of morphological changes which accompany transformation and some experiments have suggested that plasmin or plasminogen might bind to transformed cell surfaces in preference to normal cell surfaces. We therefore studied the time dependence of plasminogen binding to 3T3 and SV 3T3 cells. The binding to 3T3 cells on a per-cell basis either decreased during the course of a three-day incubation or decreased between the first and second days and then rose again on the third day ("V" shaped binding curve). The binding to the SV 3T3 cells was usually "V" shaped. After three days of incubation, all of the plasminogen bound to the SV 3T3 cells had been degraded while there was still substantial plasminogen associated with the 3T3 cells. The degradation by the transformed cells was independent of the plasmin-dependent morphological changes normally exhibited by these cells. The uptake of serum plasminogen by the cells apparently involved several processes and may be accompanied by cell-surface proteolysis in the case of the transformed cells.

PREFACE

This thesis summarizes work which is reported in detail in reprints and preprints appearing in the Appendices.

TABLE OF CONTENTS

	PAGE
Dedication	ii
Acknowledgments	iii
Abstract	v
Preface	viii
Table of Contents	ix
Serine Proteases: Structure, Mechanism, and Role in Eucaryotic Cell Transformation	1
Appendix I: Structure and Specific Binding of Trypsin: Comparison of Inhibited Derivatives and a Model for Substrate Binding	24
Appendix II: The Effect of Pre-Incubation on Trypsin Kinetics at Low pH	48
Appendix III: The pH Dependence of Tritium Exchange with the C-2 Protons of the Histidines in Bovine Trypsin	64
Appendix IV: Silver Ion Inhibition of Serine Proteases: Crystallographic Study of Silver-Trypsin	114
Appendix V: Data Collection in Protein Crystallography: Capillary Effects and Background Corrections	120
Appendix VI: Data Collection in Protein Crystallography: Experimental Methods for Reducing Background Radiation	130
Appendix VII: Structure-Function Relationships in the Serine Proteases	154

Appendix VIII: Plasminogen Binding to Balb/c 3T3 and SV 3T3 Cells 175

Abstracts of Propositions 214

INTRODUCTION

Of the many ways to control the biological activity of proteins (e.g., induction or repression of their syntheses at the translational (1) or transcriptional levels (2)), specific modification or destruction are the most direct. Biological systems are frequently controlled by these simple methods, and the serine protease family of enzymes plays a major role in many of these systems (3).

The serine proteases are enzymes which hydrolyze peptide bonds via a series of base catalyzed nucleophilic substitutions. In many respects trypsin is the archetypical serine protease. Trypsin is synthesized in the pancreas of mammals as an inactive precursor which is later activated in the duodenum (4) by the cleavage of one critical peptide bond near the amino-terminal end of the polypeptide chain (5). After activation, trypsin's primary function is to digest food protein. While the serine proteases are most frequently recognized as the primary tools in dietary protein digestion, they are also essential components in many other biological processes; for example, blood clotting (6), bacterial sporulation (7), and fertilization (8). Trypsin's sharp specificity for basic side chains of the correct lengths (9, 10) has meant that a large number of physiologically important reactions are performed by enzymes

resembling trypsin. These trypsin-like enzymes of physiological control attack just one or two specific peptide bonds in a whole system of proteins.

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 (11a - 11h).

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

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 amine 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, $E + S \rightarrow EP_2^{a}$, while deacylation, $EP_2^{a} \rightarrow E + P_2$ is usually rate-determining for esters. 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 specific of the digestive enzymes, hydrolyzes those peptide bonds that immediately follow either of the two basic amino acids, lysine or arginine (12). 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 (13). The complete amino acid sequences and three-dimensional molecular structures have now been worked out for trypsin (14) and its proenzyme (15), chymotrypsin (16) and its proenzyme (17), and elastase (18). These structures along with that of subtilisin (19), a bacterial serine protease, have been valuable in developing an understanding of how these enzymes bind their substrates and how they catalyze the subsequent chemical reactions.

I have studied the serine proteases by examining the structure and catalytic mechanism of bovine trypsin, the archetype, and by investigating the role that the trypsin-like enzyme plasmin plays in the manifestation of transformed properties by eucaryotic cells in culture.

THE STRUCTURE OF INHIBITED TRYPSIN

In 1971, Bob Stroud, Lois Kay, and Richard Dickerson solved the three-dimensional structure of Diisopropyl-fluorophosphate inhibited trypsin (DIP-trypsin) at 2.7 Å resolution (14). Their work presented a picture of the folding of the polypeptide chain and amino

acid side chains along with a view of how the DIP-inhibiting group blocked the active center of the enzyme. The structure of DIP-trypsin was very similar to the structures of bovine α-chymotrypsin and bovine elastase, two other pancreatic digestive serine proteases. The structures of these three proteins differed most sharply in the regions of their specificity binding pockets--the portion of the active site of the serine proteases which binds the side chains of substrates. In order to determine how the side chains of specific substrates bind to trypsin, we determined the structure of trypsin with the specific substrate side chain analogue benzamidine bound in the pocket (9; Appendix I). Benzamidine binding provided a reasonable model for the binding of specific side chains and permitted us to construct a model for specific substrate binding.

Benzamidine-trypsin allowed us to study trypsin's binding pocket while DIP-trypsin probed the serine hydroxyl group in the active center, which is responsible for the nucleophilic substitution during catalysis. Recently, the structure of DIP-trypsin has been extensively refined in our laboratory at 1.5 Å resolution using a new technique for automated Fourier and real-space refinement developed by John L. Chambers. The trypsin structure is currently the most accurately determined enzyme structure.

ENZYME SPECIFICITY AND SUBSTRATE BINDING

The x-ray crystal structures of the serine proteases have shown that their active sites are almost identical, and can be subdivided into catalytic and substrate binding sites.

The serine proteases differ in their specificities not because of their catalytic sites, which are all identical, but because of differences in their substrate binding sites. Trypsin, chymotrypsin and elastase all have specific side chain binding pockets on their surfaces close to the catalytic sites. These pockets are lined by residues 214-220 and 189-192. 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 (20)) lies at the bottom of the pocket (14). Trypsin is specific for basic amino acids because the positively charged side chains bind tightly in this pocket (9, 21, 22, 23). It has been suggested that side chain binding in this pocket induces conformational changes in the enzyme-substrate complex which help to accelerate catalysis (24). In chymotrypsin, residue 189 is a serine (25). The pocket is relatively hydrophobic and uncharged at neutral pH's, thus explaining chymotrypsin's specificity. In both trypsin and chymotrypsin, residue 216, a glycine, is at the entrance to the binding pocket. In elastase,

a valine replaces the glycine at 216 (26); its large side chain blocks the pocket and only allows the binding of amino acids with small side chains (18).

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. 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, it 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 be collected in so short a time. Allan Fersht and M. Renard (11g) have pointed out, however, that it may be possible to use equilibrium methods to trap intermediates in the reaction pathway and study their structures. Dr. G. Petsko is currently trying to trap crystalline intermediate structures using very low temperatures (private communication). Until the structures of such intermediates have been determined, crystallographers will be limited to studying substrate analogue binding. From such studies, we can make inferences about the structural transformations which occur during the reaction.

The best analogues 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 in the physiological control of the serine proteases (27). For example, if pancreatic trypsin inhibitors were not synthesized with the 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. The inhibitors are present to prevent such catastrophes. The structure of a 6500 molecular weight bovine pancreatic trypsin inhibitor (PTI) was determined by R. Huber and his associates (28, 29). Chemical modifications had shown that Lys 15 of this inhibitor was involved in the trypsin-PTI association (30). By combining models of the PTI with the known structures of trypsin and chymotrypsin, substrate binding models were developed in our laboratory (9, 14), and independently by Huber et al. (29) and Blow et al. (31).

To construct our model an arginine side chain was built into the inhibitor structure replacing the side chain of Lys 15. Models of trypsin and the modified inhibitor were brought together so that the "Arg 15" side chain was superposed onto the benzamidine density in the difference map. The side chain of Gln 192 was moved to

accommodate the inhibitor. Otherwise, the trypsin active-site configuration was that of benzamidine trypsin. The resulting complex (Figure 10 of Appendix I) was essentially the same as that proposed earlier by Stroud et al. (14).

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 binding of specific substrate side chains. Benzamidine binding is, therefore, a good model for substrate side chain binding, showing a number of interactions that we expect to be present during the specific interaction between real substrates and trypsin. For example, several hydrogen bonds and a ready-made hand-in-glove association orient the susceptible peptide bond close to the active site. The identical models proposed for the enzyme-substrate-inhibitor complexes have now been verified by the determination of the crystal structure of the PTI-trypsin complex (22). The soybean trypsin inhibitor-trypsin complex has also been determined (23).

THE CATALYTIC SITE

The catalytic site of all serine proteases is characterized by a serine hydroxyl group (residue 195, in chymotrypsinogen number-

ing system). Diisopropyl-fluorophosphate (DIFP)(32) and phenyl-methane-sulfonyl-fluoride (PMSF) (33) react with this hydroxyl and irreversibly inhibit serine proteases, regardless of their substrate specificity. This hydroxyl is hydrogen-bonded to the $N^{\epsilon 2}$ of His 57 which in turn hydrogen-bonds via $N^{\delta 1}$ to Asp 102 (14, 16, 17, 19, 34). Their direct participation in catalysis has been established unequivocally. Chemical modification of any of these three residues can greatly diminish or abolish catalysis (32, 33, 35, 36, 37, 38).

The ionization properties of the side chains of His 57 and Asp 102 have been the subject of considerable controversy. The pH activity profile for trypsin is bell-shaped with a maximum near pH 8. The high pH limb of the profile is due to an active-inactive conformational equilibrium which depends on the titration of the amino terminus (11d). The low pH limb depends on only a single group between pH 3 and pH 7 which titrates with a pK_a of 6.8. The pK_a's of the histidine and aspartic acid side chains in solution are normally 6.0 and 3.6 (39), respectively; therefore, the active site pK_a of 6.8 has frequently been assigned to His 57.

Hunkapiller et al. (40) have used ^{13}C-NMR to show that the active site imidazole in α-lytic protease, a bacterial serine protease homologous to the pancreatic enzymes, remains neutral above pH 4. They inferred from their work that Asp 102 was the group

with the pK_a of 6.8. Roger Koeppe II and Robert Stroud have used infrared difference spectroscopy to directly show that the pK_a of Asp 102 in trypsin is 6.8 (41).

Any discrepancy between pK_a's determined using enzyme kinetics and spectroscopy might be attributable to slow pH-dependent conformational changes which could occur during the long incubation times required for the spectroscopic measurements. One way to approach the question of whether a slow conformational change affects the pH dependence of serine protease hydrolysis is to conduct kinetic experiments with the enzyme pre-incubated at different pH's. Using this method, Rodriguez and Hollaway (42) detected a pH-dependent conformational change for phosphofructokinase having a half-time of about four minutes. We have looked for such a conformational change in trypsin and found that pre-incubation of trypsin for up to three hours at 0°C at pH 2.0 or at pH 6.9 had no effect on the rate of hydrolysis of N_α-carbobenzoxy-L-lysine-p-nitrophenyl ester (CLNE) by trypsin (43; Appendix II). We also found that CLNE hydrolysis depended on the titration of a single group between pH 3 and pH 7. This group is Asp 102 (41); therefore, the pK_a of His 57 in trypsin must be below 3.

To further probe the ionization behavior of the histidines in trypsin, we studied the pH dependence of tritium incorporation into

the C-2 position of the imidazole side chains (44; Appendix III). We found that the C-2 exchange rates in trypsin are slower than any previously reported imidazole exchange rates and that the side chain of His 57 appears to be involved in a conformational equilibrium. Apparently, when the imidazole ring of His 57 resides in its active "in" conformation, tritium exchange at the C-2 position is sterically hindered. When it is in its inactive "out" conformation, in which it is swung out into the surrounding solution, it can undergo the exchange. The apparent pK_a of His 57 in the "out" conformation is 6.6. The exchange occurs with a half-time of 73 days which suggests that 2% of the molecules have His 57 in the "out" conformation.

The structure of trypsin around the catalytic site perturbs the pK_a's of Asp 102 and His 57. The mechanistic importance of this perturbation is that aspartic acid acts as a chemical base which can readily accept a proton from the histidine side chain during catalysis. Aspartic acid 102 and histidine 57 shuttle protons back and forth from enzyme to substrate, and so the mechanism can best be described as nucleophilic attack with base catalysis by His 57 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 has disasterous effects on catalysis. For example, by methylating the $N^{\epsilon 2}$ of His 57 in chymotrypsin, the shuttle can no longer operate normally, and the rate of catalysis drops by a factor of 5000 to 200,000 for specific substrates (36). Silver ions are known to be potent inhibitors of trypsin and chymotrypsin (37); therefore, we prepared crystals of silver-bound DIP-trypsin in order to determine the site of silver binding (38; Appendix IV). Our crystallographic study showed that the primary silver ion binding site lies between Asp 102 and His 57 at the catalytic site. Thus, silver blocks the proton shuttle and therefore inhibits catalysis.

During the course of the silver-binding study, I realized that a new technique of background data collection which was used in other laboratories to speed the data collection process was fraught with systematic errors. We investigated the sources of background radiation and developed two new techniques which eliminated the systematic errors (45, 46), but still allowed substantial savings in data collection time. This first background study (45; Appendix V) led to a number of suggested improvements in the experimental apparatus which we later incorporated into our data collection system. The modifications in our equipment yield improved signal-to-noise ratios and permit further reductions in the data collection time and the use

of smaller specimens (46; Appendix VI).

MECHANISM OF HYDROLYSIS BY SERINE PROTEASES

Crystallographic studies, together with a vast amount of chemical information, have produced detailed models of the activation of the proenzymes and the mechanism of serine protease hydrolysis. The most recent developments in serine protease enzymology are summarized along with our interpretation of the mechanism in the review in Appendix VII (47).

SERINE PROTEASES AND EUCARYOTIC CELL TRANSFORMATION: PLASMINOGEN BINDING

With the substantial insight into the structures of serine proteases and their mechanism of catalysis provided by biophysical and biochemical research, I shifted my attention to the many physiological functions mediated by serine proteases. Along with Jerry Tobler and Robert Stroud, I have studied one problem associated with the question of how serine proteases act in controlling the proliferation of animal cells in culture; namely, how do the serine protease precursor plasminogen and the active enzyme plasmin bind to

established lines of normal and transformed fibroblasts in culture? Our interest in the plasminogen system was generated by several experiments. Plasminogen, an approximately 95,000 molecular weight protein found in blood, is normally activated in response to blood clot formation and plasmin's primary function is the dissolution of the clot. Burger found that gentle proteolysis can stimulate quiescent cells to initiate a round of replication (48), and others have shown that proteolysis can occasionally mimic or potentiate the activity of peptide hormones which are known to stimulate cell growth (49). Dr. E. Reich and his colleagues have found that the enhanced proteolytic activity exhibited by many cells transformed by oncongenic DNA, RNA viruses, or chemical carcinogens is important in mediating certain morphological changes characteristic of transformation (50, 51). Unkeless et al. (52) have shown that transformed cells synthesize an enzyme which activates serum plasminogen to plasmin. The increased proteolytic activity of transformed cells due to plasmin is responsible for these morphological changes (53, 54). Cells which have undergone these morphological changes are spherical, tend to clump, and retract from the substratum. When cells are grown in plasminogen-free serum or in serum containing high concentrations of plasmin inhibitors, the morphological changes do not occur (54). Ossowski et al. (55) reported that simian

virus 40 (SV 40) transformed hamster embryo fibroblasts bind more ^{125}I-labeled plasminogen than their normal counterparts. These binding data and reports of proteolytic activity associated with the plasma membranes of transformed cells (50, 55, 56) suggested the possibility of a specific cell surface-plasmin interaction. We have studied the binding of ^{125}I-labeled dog plasminogen to normal and SV 40 transformed Balb/c 3T3 fibroblasts and its relationship to the plasmin-dependent morphological changes (PDMC) exhibited by the transformed cells (Appendix VIII).

Our results show a significant and rather remarkable difference between the binding and processing of plasminogen by normal and SV 40 transformed mouse Balb/c 3T3 cells. Both cell types bind substantial amounts of plasminogen ($\sim 10^5$ molecules/cell). The binding to 3T3 cells on a per-cell basis either decreases dramatically during the course of a three-day incubation or decreases between the first and second days and then rises again on the third day ("V" shaped binding curve). The binding to the SV 3T3 cells is usually "V" shaped. The transformed cells degrade the plasminogen much more rapidly than the normal cells. After three days of growth in the presence of ^{125}I-labeled plasminogen, all of the plasminogen associated with the transformed cells was digested into small polypeptides and individual amino acids while a substantial portion of the

plasminogen associated with the normal cells was still intact. Our inhibitor studies showed that the degradation was essentially independent of the activation of plasminogen in the growth medium by the plasminogen activator which is synthesized by the cells. Therefore, the degradation is cell-specific and independent of PDMC. The striking differences in plasminogen degradation may depend on cell surface proteases present on the transformed but not the normal cells, or on differences in endocytosis and lysozomal proteolysis, or both. Clearly, the phenomenon is complex and will require further study before it can be explained by models which incorporate the present, rather limited, understanding of endocytosis. We are are continuing our investigations of plasminogen binding using autoradiography in order to localize the sites of plasminogen degradation and further characterize the processes involved in the binding and degradation.

REFERENCES

1. F. Jacob and J Monod, J. Mol. Biol. 3, 318 (1961).
2. G. M. Tomkins, T. D. Gelehrter, D. Granner, D. Martin Jr., H. H. Samuels, E. B. Thompson, Science 166, 1474 (1969).
3. R. M. Stroud, Scientific American 231, 74 (1974).
4. M. Kunitz and J. H. Northrop, J. Gen. Physiol. 19, 991 (1936); J. H. Northrop, M. Kunitz, and R. M. Herriot, Crystalline Enzymes Columbia University Press, New York, 1948 ; S. Maroux, J. Baratti, and P. Desnuelle, J. Biol. Chem. 246, 5031 (1971).
5. E. W. Davie and H. Neurath, J. Biol. Chem. 212, 515 (1955).
6. For detailed reviews of the blood-clotting system, see P. A. Owren and H. Stormorken, Ergeb. Physiol. Biol. Chem. Exp. Pharmakol. 68, 1 (1973); and S. Magnusson in The Enzymes (P. D. Boyer, Ed.) Academic Press, New York, 1971, Vol. 3, p. 277.
7. T. J. Leighton, R. H. Doi, R. A. J. Warren, and R. A. Kelln, J. Mol. Biol. 76, 103 (1973).
8. R. Stambaugh, B. Brackett, and L. Mastroianni, Biol. Reprod. 1, 223 (1969).

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

10. N. J. Baines, J. B. Baird, and D. T. Elmore, Biochem. J. 90, 470 (1964); J. B. Baird, E. G. Curragh, and D. T. Elmore, ibid., 96, 733 (1965).

11. a. B. Zerner, R. P. M. Bond, and M. L. Bender, J. Amer. Chem. Soc. 86, 3674 (1964).

b. J. Fastrez and A. R. Fersht, Biochemistry 12, 2025 (1973).

c. G. P. Hess, J. McConn, E. Ku, and G. McConkey, Phil. Trans. Roy. Soc. Lond. Ser. B. 257, 89 (1970).

d. A. R. Fersht and Y. Requena, J. Amer. Chem. Soc. 93, 7079 (1971).

e. J. Fastrez and A. R. Fersht, Biochemistry 12, 1067 (1973).

f. M. Caplow, J. Amer. Chem. Soc. 91, 3639 (1969).

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

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

12. H. Neurath and B. S. Hartley, J. Cell. Comp. Physiol. 54 (1), 179 (1959).

13. M. A. Naughton and F. Sanger, Biochem. J. 78, 156 (1961).

A. Sampath Naranyanan and R. A. Anwar, ibid., 114, 11 (1969).

J. R. Brown, D. L. Kauffman, and B. S Hartley, ibid., 103,

497 (1967).

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

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

15. A. Kossiakoff, L. Kay, and R. M. Stroud, unpublished results.

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

b. J. J. Birktoft and D. M. Blow, J. Mol. Biol. 68, 187 (1972).

17. S. T. Freer, J. Kraut, J. D. Robertus, H. T. Wright, and Ng. H. Xuong, Biochemistry 9, 1997 (1970).

H. T. Wright, J. Mol. Biol. 79, 1 (1973).

18. D. M. Shotton and H. C. Watson, Phil. Trans. Roy. Soc. Lond. Ser. B. 257, 111 (1970).

19. C. S. Wright, R. A. Alden, and J. Kraut, Nature 221, 235 (1969).

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

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

22. A. Ruhlmann, D. Kukla, P. Schwager, K. Bartels, and R. Huber, J. Mol. Biol. 77, 417 (1973).

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

R. M. Sweet, H. T. Wright, J. Janin, C. H. Chothia, and

D. M. Blow, Biochemistry 13, 4212 (1974).

24. T. Inagami and T. Murachi, J. Biol. Chem. 239, 1395 (1964).

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

25. B. S. Hartley, Nature 201, 1284 (1964).

26. D. M. Shotton and B. S. Hartley, Nature 225, 802 (1970).

27. For a review of protease inhibitors, see H. Tschesche, Angew. Chem. Int. Ed. Engl. 13, 10 (1974).

28. R. Huber, D. Kukla, A. Ruhlmann, O. Epp, and H. Formanek, Naturwiss. 57, 389 (1970).

29. R. Huber, D. Kukla, A. Ruhlmann, and W. Steigemann, Cold Spring Harbor Symp. Quant. Biol. 36, 141 (1971).

30. K. A. Wilson and M. Laskowski Sr., J. Biol. Chem. 246, 3555 (1971).

M. Rigbi in Proceedings of the International Conference on Proteinase Inhibitors, Munich, 1970, (H. Fritz and H. Tschesche, Eds.), Walter deGruyter, Berlin, 1970, p. 74.

H. Fritz, H. Schult, R. Meister, and E. Werlo, Z. Physiol. Chem. 350, 1531 (1969).

31. D. M. Blow, C. S. Wright, D. Kukla, A. Ruhlmann, W. Steigemann, and R. Huber, J. Mol. Biol. 69, 137 (1972).

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

B.S. Hartley, Ann. Rev. Biochem. 29, 45 (1960).

33. D.E. Fahrney and A.M. Gold, J. Amer. Chem. Soc. 85, 997 (1963).

J. Kallos and D. Rizok, J. Mol. Biol. 9, 255 (1964).

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

35. E.B. Ong, E. Shaw, and G. Shoellmann, J. Amer. Chem. Soc. 86, 127 (1964).

E. Shaw, M. Mares-Guia, and W. Cohen, Biochemistry 4, 2219 (1965).

36. R. Henderson, Biochem. J. 124, 13 (1971).

37. K. Martinek, Y.V. Savin, and I.V. Berezin, Biokhimiya 36, 806 (1971).

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

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

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

41. R.E. Koeppe II and R.M. Stroud, Biochemistry--in press, 1976.

42. H.B. Rodriguez and M.L. Hollaway, Bioch. Soc. Trans. 4, 83 (1976).

43. R. E. Koeppe II, M. Krieger, and R. M. Stroud. Unpublished work.

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

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

46. M. Krieger and R. M. Stroud, Acta Crystallogr. --in press, 1976.

47. R. M. Stroud, M. Krieger, R. E. Koeppe II., A. A. Kossiakoff, and J. L. Chambers, Cold Spring Harbor Symposium on Proteases and Biological Control, p. 13 (1975).

48. M. M. Burger, Nature 227, 170 (1970).

49. R. A. Yadley, D. Rodbard and A. Chramback, Endocrynology, 93, 866 (1973).

U. J. Lewis, S. J. Pence, R. N. P. Singh, and W. P. VanderLaan, Biochem. Biophys. Res. Commun. 67, 617 (1975).

L. A. Greene, J. T. Tomita and S. Varon, Exp. Cell Res. 64, 387 (1971).

R. O. Hynes, Cell 1, 147 (1974).

50. J C. Unkeless, A. Tobia, L. Ossowski, J. P. Quigley, D. B. Rifkin, and E. Reich, J. Exp. Med. 137, 85 (1973).

51. L. Ossowski, J. C. Unkeless, A. Tobia, J. P. Quigley, D. B. Rifkin, and E. Reich, J. Exp. Med. 137, 112 (1973).

52. J. Unkeless, K. Dano, G. M. Kellerman, and E. Reich, J. Biol. Chem. 249, 4294 (1974).

53. J. P. Quigley, L. Ossowski, and E. Reich, J. Biol. Chem. 249, 4306 (1974).

54. L. Ossowski, J. P. Quigley, and E. Reich, J. Biol. Chem. 249, 4312 (1974).

55. L. Ossowski, J P. Quigley, G. M. Kellerman, and E. Reich, J. Exp. Med. 138, 1056 (1973).

56. M. M Burger, in Growth Control in Cell Cultures (G. E. W. Wolstenholme and J. Knight, Eds.), pp. 45-63. Churchill Livingstone, Edinburgh and London, 1971.

APPENDIX I

Structure and Specific Binding of Trypsin: Comparison of Inhibited Derivatives and a Model for Substrate Binding

M. Krieger, L. M. Kay, and R. M. Stroud

J. Mol. Biol. 83, 209-230 (1974).

J. Mol. Biol. (1974) **83**, 209–230

Structure and Specific Binding of Trypsin : Comparison of Inhibited Derivatives and a Model for Substrate Binding

M. Krieger, L. M. Kay and R. M. Stroud

Norman W. Church Laboratory of Chemical Biology
California Institute of Technology
Pasadena, Calif. 91109, U.S.A.

(*Received 21 June 1973, and in revised form 12 October 1973*)

The high-resolution structure of bovine trypsin inhibited with DFP† was determined by Stroud *et al.* (1971 and R. M. Stroud, L. M. Kay, A. Cooper & R. E. Dickerson, *Abstr. 8th Int. Congr. Biochem.* 1970). The experiments reported here were designed to study the specific side-chain binding pocket of trypsin using benzamidine, which is a competitive, specific inhibitor of trypsin. High-resolution electron density syntheses and difference syntheses unambiguously identify the side-chain binding pocket, which normally recognizes and binds the side chains of arginine or lysine during proteolysis. Several important conformational differences in the protein structure are apparent between DIP- and BA-trypsins, and these are discussed with particular reference to inhibition, the binding of lysine and arginine, subsequent orientation of the target at the active site, and the enhancement of tryptic activity towards non-specific substrates seen on binding small alkyl amines or guanidines in the specific binding pocket.

The BA-trypsin structure provides a good model for the binding of real substrate side chains to trypsin during catalysis, explaining the sharp trypsin specificity for lysine or arginine side chains (Weinstein & Doolittle, 1972) and the lack of specificity for stereochemically different basic side chains. Benzamidine is shown to inhibit trypsin by steric interference with the inferred position of good substrates, even when they do not carry any side chain.

Apart from the substitution of benzamidine and DIP, the most significant differences between DIP-trypsin and BA-trypsin involve complete repositioning of the side chain of Gln192, alterations in the side chains of Asp102, His57 and Ser195 at the active site, and changes in the solvent structure around this region. The carboxyl group of Asp189, which is responsible for trypsin specificity, shows no movement on binding benzamidine. The amidinium cation of benzamidine forms a salt bridge with Asp189 in BA-trypsin; a similar salt bridge can be constructed between the side chains of model substrates with lysyl or arginyl side chains and Asp189. The γ-oxygen of Ser190 is displaced by a 120° rotation about its α–β bond on binding benzamidine and the binding pocket closes to sandwich the inhibitor ring between the peptide planes of 190–191 and 215–216. These contacts are presumably found in the enzyme–substrate complex with specific substrates.

† Abbreviations used: DFP, diisopropyl fluorophosphate; DIP, diisopropyl phosphoryl; DIP-trypsin, diisopropyl fluorophosphate-inhibited trypsin; BA-trypsin, benzamidine-inhibited trypsin. In equations the latter are further abbreviated to BAT and DIPT. The amino acid abbreviations follow standard convention, and the amino acid residue numbering is that adopted by Stroud *et al.* (1971).

The active site structure at pH 8·0 is discussed with particular reference to the microscopic pK_a values of Asp102 and His57, the pK_a of the Asp–His system, and the mechanistic consequences of these assignments.

1. Introduction

Trypsin shows its greatest activity for proteolysis between pH 7·0 and 9·0 (Northrop & Kunitz, 1932). It has been one of our goals to study the structure of trypsin within this pH range in order to minimize the possibility of examining an inactive conformation of the enzyme. Trypsin undergoes at least three pH-dependent conformational changes between pH 0·5 and 7·0, which are detectable by optical rotary dispersion (Lazdunski & Delaage, 1967). More subtle, but nonetheless mechanistically significant pH-dependent changes in the conformation, particularly around the active site, have been detected in the serine proteases using more sensitive techniques (Vandlen & Tulinsky, 1973). One of the difficulties of studying native trypsin within the most active pH range is its rapid rate of autolysis. Firmly bound inhibitors reduce the possibility of autolysis. For this reason, our first crystallographic studies were carried out on DIP-trypsin, in which the enzyme had been irreversibly inactivated by covalent modification of the reactive serine 195 using a non-specific inhibitor (DFP, Fig. 1).

Benzamidine (Fig. 1) is a competitive, specific and reversible inhibitor of trypsin, which binds with a $K_1 = 1{\cdot}8 \times 10^{-5}$ M (Mares-Guia & Shaw, 1965). Its amidinium

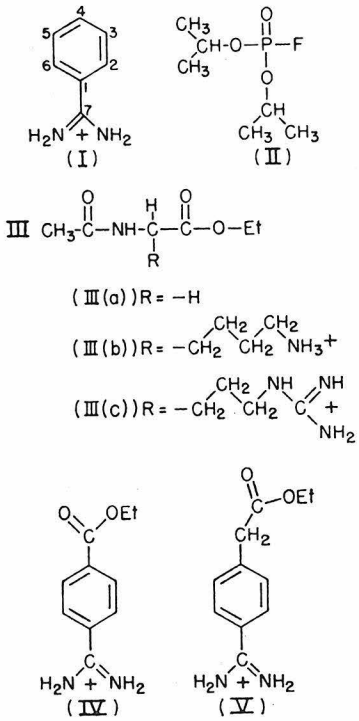


FIG. 1. Chemical structures: (I) benzamidine; (II) diisopropyl fluorophosphate; (III) *N*-acetyl amino acid ethyl esters, (a) acetyl glycine ethyl ester, (b) acetyl lysine ethyl ester, (c) acetyl arginine ethyl ester; (IV) ethyl *p*-amidino benzoate; (V) ethyl *p*-amidino phenyl acetate.

cation interacts specifically with the binding pocket of trypsin in a manner that is probably similar to that of an arginine side chain on a specific substrate.

In studying BA-trypsin, we hoped to learn more about the highly specific binding pocket of trypsin, how it functions in co-operation with the active site, and how certain small molecules shut down tryptic catalysis, while others serve to enhance the catalytic rate for substrates with small, non-polar side chains.

2. Experimental

(a) *Preparation of crystals of inhibited trypsin*

(i) *Benzamidine-inhibited trypsin*

Bovine trypsin (150 mg) (Mann Fine Chemical Company) and benzamidine hydrochloride (5·7 mg) (Aldrich Chemical Company) were dissolved in 5·0 ml cold 0·05 M-Tris buffer (pH 8·17) and stored at 5°C for 30 min. 500 mg of anhydrous $MgSO_4$ were added slowly with constant stirring at 5°C. The solution was clarified by centrifugation and sterilized by passing it through a Millipore filter into 3 sterile vials. The vials were stoppered and stored at 5°C.

Tiny needles were observed in the solution after about 1 month; these crystals grew to small, but usable size (about 0·4 mm $\times$0·2 mm $\times$0·2 mm) in about 2½ months.

(ii) *Diisopropyl phosphoryl trypsin*

Bovine trypsin was inhibited with DFP according to the procedure of Cunningham (1954).

150 mg of DIP-trypsin was dissolved in 3·75 ml distilled water at room temperature and the pH was adjusted to 7·5 with 0·05 N-NaOH. 1·25 ml of a solution of $MgSO_4$ (15 g anhydrous salt dissolved in 50 ml water) was added slowly with stirring. The solution was clarified by centrifugation and sterilized by passing it through a Millipore filter into 2 or 3 sterile vials. The vials were stoppered and stored at room temperature. Rod-shaped crystals of usable size usually took at least 2 months, and often much longer, to grow.

(iii) *State of the trypsin in the crystals*

Preparations of DIP-trypsin and BA-trypsin crystals are quite similar. In order to retard autolysis of the reversibly inhibited trypsin, BA-trypsin solutions were refrigerated. It was then necessary to raise the precipitant ($MgSO_4$) concentration by about 33% above that used in the DIP-trypsin crystallizations. In both cases, the pH was between 7·5 and 8·2, in the region where the uninhibited enzyme would be most active. Furthermore, evidence summarized by Stroud *et al.* (1971) leads us to believe that the structure of DIP-trypsin at pH 7·5 is very similar to native trypsin at pH 7·5 in gross conformation and state of protonation.

Chromatography of solutions prepared from crystals of DIP-trypsin and BA-trypsin indicates that both contain mixtures of inhibited α and β-trypsins and small amounts of other partially cleaved molecules. Although the ratios vary, they tend toward a 1:1 ratio of α to β-trypsin, a mixture richer in α-trypsin than the commercial trypsin or the DIP-trypsin from which the crystals were prepared. β-Trypsin contains a single polypeptide chain of 223 amino acids, while α-trypsin has a single tryptic scission between Lys145 and Ser146, and shows almost identical activity to α-trypsin (Schroeder & Shaw, 1968).

(b) *X-ray data collection*

Complete 3-dimensional X-ray diffraction data from BA-trypsin and DIP-trypsin were collected to a resolution of 2·7 Å. Data for DIP-trypsin were collected on an extensively modified Supper diffractometer using Ni-filtered CuKα X-radiation. A graphite monochromator designed by Dr Sten Samson was built and incorporated into the system for collection of BA-trypsin data (Stroud *et al.*, 1974). In all, 32 crystals were used for DIP-trypsin and 6 for BA-trypsin. The difference in number is due to the reduced rate of crystal decay in the monochromatized beam.

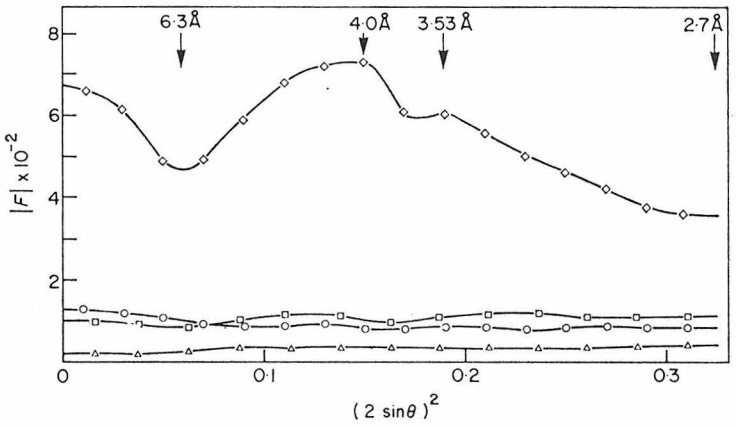


FIG. 2. The $(2\sin\theta)^2$ dependence of the average values of (◇) $|F_{DIPT}|$, (○) $|F_{DIPT-TL} - F_{DIPT}|$, (□) $|F_{BAT} - F_{DIPT}|$, and the estimated errors ($|\Delta F|$) for DIP-trypsin data collected from 2 different crystals, one with a monochromator and one without (△) (±2·5% from the mean value.).

It was surprising to find that the mean difference in diffracted amplitudes $\langle|\Delta F|\rangle$ between BA-trypsin and DIP-trypsin ($\Delta F = F_{BAT} - F_{DIPT}$) was greater than the mean difference between DIP-trypsin and a thallium derivative previously used in phase analysis. The value of $\langle|\Delta F|\rangle/\langle|F_P|\rangle$ was 21·6%, at least 6 to 7 times the expected errors from all sources, and was roughly constant out to 2·7 Å, indicating that quite specific changes were occurring in the enzyme between highly isomorphous crystals (Fig. 2).

An estimate of the expected error in a single measurement, arising from all sources, was made by comparing data sets from 2 different crystals of DIP-trypsin, one set taken with Ni-filtered CuKα radiation (F_{Pn}) and the other with monochromatized CuKα X-rays (F_{Pm}). The mean value of these differences, shown in Fig. 2, gives a mean residual,

$$R = \sum_{hkl} (|F_{Pm} - F_{Pn}|)/\sum_{hkl}|F_{Pm}|,$$

of 5·1% after corrections and scaling (including all zero intensities) or ±2·5% from the mean value of F_P.

Data were corrected for absorption, time-dependent decay, geometry, monochromator polarization, etc., and scaled together to obtain F_{BAT} and F_{DIPT} according to the procedure of Stroud *et al.* (1974). Final scaling of the two data sets was accomplished with a 2-parameter (K', B') exponential function derived by least-squares minimization of

$$E'(s) = [K' \exp\{B's^2 \langle|F_{BAT}|\rangle_s - \langle|F_{DIPT}|\rangle_s\}]^2,$$

computed from 20 zones of $s = 2\sin\theta$. The cell dimensions of BA-trypsin and DIP-trypsin crystals were identical within experimental error. The space group and cell dimensions are:

space group: $P2_12_12_1$ $a = 54{\cdot}8 \pm 0{\cdot}05$ Å
$b = 58{\cdot}7 \pm 0{\cdot}05$ Å
$c = 67{\cdot}6 \pm 0{\cdot}05$ Å.

(c) *Electron density calculations*

A difference electron density synthesis was calculated using phases ϕ_P and figures of merit, m, previously determined for DIP-trypsin (Stroud *et al.*, 1974) using the expression

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

where $\Delta F_s = (|F_{BAT}| - |F_{DIPT}|)_s$. Reflections where F_{BAT} or F_{DIPT} were less than 2σ were left out of the synthesis, and the map was calculated at intervals of 0·01 in y and z, and 0·02 in x, making the finest interval 0·6 Å.

The difference map was contoured at intervals of 0·05 electrons/Å^3, omitting the zero

TABLE 1

Analysis of Fourier maps

Map	$\langle\|F_{obs}\|\rangle$ (e)	σ (e)	Calculated† $\langle\Delta\rho^2\rangle^{\frac{1}{2}}$ (e Å^{-3})	Observed‡ r.m.s. error (e Å^{-3})	Observed highest noise (e Å^{-3})	s.d.§	Observed highest peak (e Å^{-3})	s.d.
BA-trypsin—DIP-trypsin	84·7	2·3	0·069	0·059	0·17	2·5	0·75	11
DIP-trypsin	573·0	21·0	0·38	—	—	—	—	—

† ΔF: $\langle\Delta\rho^2\rangle = \frac{1}{2V^2}\sum_{hkl}\Delta F^2\,(2-m^2)$,

F_{DIPT}: $\langle\Delta\rho^2\rangle = \frac{1}{V^2}\sum_{hkl}F^2{}_{DIPT}\,(1-m^2)$,

(after Henderson & Moffat, 1971).

‡ The observed root mean-square density error is based on a relatively featureless region of the map.
§ s.d., the electron density given as a multiple of the calculated r.m.s. error.

and first contour in both the negative and positive density, using an approximate (statistically determined) absolute scale factor (Stroud *et al.*, 1974). Five sections of the map in the region of benzamidine binding are shown in Plate I. The root mean-squared error $\langle \Delta\rho^2 \rangle^{\frac{1}{2}}$ in this map is 0·069 electrons/Å^3, and a comparison of the difference and DIP-trypsin electron density maps may be found in Table 1. The highest uninterpreted "noise" density peak corresponds to 2·5 times the calculated value for the standard deviation of difference density in the map, indicating that all the observed electron density changes above the noise have been accounted for.

Molecular movements that were less than the resolution (2·7 Å) were calculated from the difference map using vector moments evaluated for pairs of equal positive and negative features, and by analysis of the difference between the density map, computed using terms

$$(2|F_{\mathrm{BAT}}| - |F_{\mathrm{DIPT}}|)\, m \exp\{i\phi_{\mathrm{DIPT}}\},$$

and the density map for DIP-trypsin obtained previously (Stroud *et al.*, 1974). Peaks in the difference map were integrated and their electron equivalents evaluated assuming that the benzamidine peak in the difference map contained its full complement of 64 electrons (Table 2). Actually, this peak represents (benzamidine—solvent) density; however, an approximation of this kind is probably more accurate than any estimate of the absolute scale of the protein. It also accounts for the factor $\langle m^2 \rangle/2$, which reduces apparent electron densities in difference maps (Henderson & Moffat, 1971).

TABLE 2

Intensities of some difference map peaks scaled to the integrated benzamidine density with its full complement of 64 electrons

Peak region	Positive peak electrons†	Negative peak electrons
His57	5·9	8·9
Cystine bridge (191–220)	8·0	9·6
Gln192	22·6	18·0
Benzamidine	64·0	—
DIP	—	90·8

† Relative peak contents were evaluated by integration and scaled so that the benzamidine peak corresponds to 64 electrons.

An estimate of the error in a shift calculated from vector moments was made by computing an "error vector moment." The error vector moment for a pair of positive and negative difference peaks is equivalent to the vector moment for the pair, except that the root mean-squared electron density error rather than the observed electron density is integrated over the peak volumes. The expected error in a shift was evaluated by dividing the error vector moment by the number of electrons, N, in the group which moves:

$$\text{shift error} = \frac{\text{"error vector moment"}}{N} = \frac{\int_{v_1} + [\sigma(\rho_r)]r\, \mathrm{d}r + \int_{v_2} - [\sigma(\rho_r)]r\, \mathrm{d}r}{N}.$$

Plate I shows an analysis of part of the difference density in a view showing the relevant region of the molecular model of DIP-trypsin to which a skeletal model of benzamidine has been added simultaneously. Superposition was accomplished with a half-silvered mirror (Richards, 1968). Newly ordered solvent molecules whose densities lie only partially in these sections are indicated by W_1 and W_2.

3. Results

(a) *Interpretation of the map*

The whole difference map, which is remarkably clean, allows a detailed interpretation of peaks with an integrated electron content down to about three electrons (corresponding to a spherical peak of diameter 2·7 Å, with density of $3\times\langle\Delta\rho^2\rangle^{\frac{1}{2}}$; Table 1).

The large positive peak (BA) shows how benzamidine binds to trypsin in the specific binding pocket. Evidently, the uniquely high specificity of trypsin and the binding energy of benzamidine, about 6·2 kcal/mol (Mares-Guia & Shaw, 1967), lead to one-site binding in contrast to other pancreatic serine proteases. α-Chymotrypsin, for example, often binds small substrate analogues at multiple sites (A. Tulinsky, personal communication). The benzamidine density reaches a maximum of about $+0{\cdot}45$ e/Å^3 in the centre of the phenyl ring, and shows quite distinct necking in the region of the $C_{(1)}$–$C_{(7)}$ bond. Solvent molecules are displaced from the binding pocket when benzamidine is bound, so that this density represents benzamidine minus solvent. The broad negative peak beneath benzamidine is due to a displacement of partially ordered solvent from the binding pocket. The amidinium cation

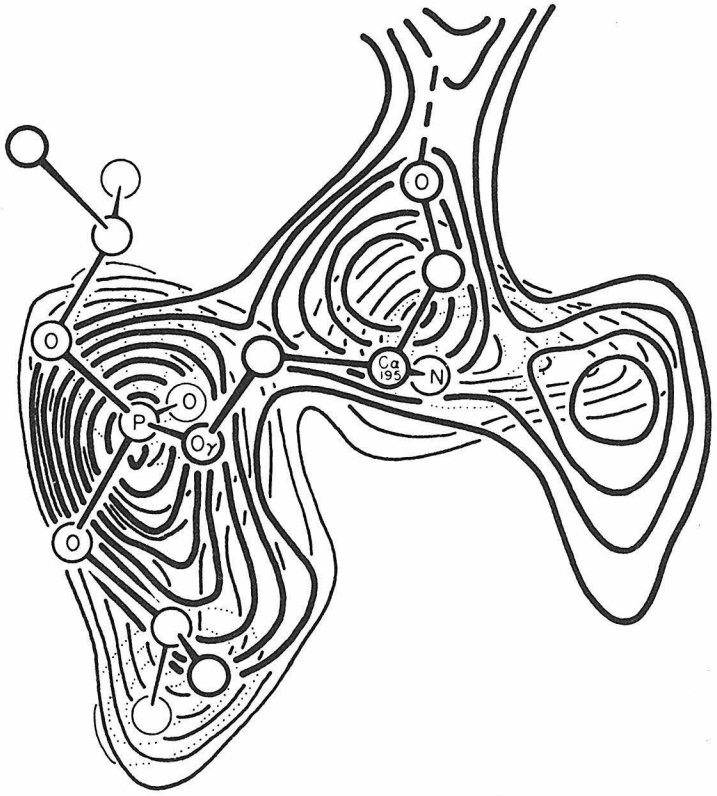


FIG. 3. Part of the DIP-trypsin electron density map showing the position of the DIP inhibitor bound to the γ-oxygen of Ser195. The first contour represents 0·25 e/Å^3. The calculated noise is between 1 and 2 contour levels, and the mean figure of merit was 0·71 out to 2·7 Å resolution. Atomic positions are labeled. The lower isopropyl group enters the binding pocket, while the upper one is not visible, presumably because of thermal motion.

of benzamidine interacts with the negatively charged carboxyl group of Asp189 at the back of the binding pocket. $N_{(2)}$ of benzamidine lies 2·9 Å from one of the carboxyl oxygens. The large negative region in the map is due to the subtraction of DIP density. The ring atom $C_{(4)}$ in benzamidine lies on a region of zero electron density, and there is a scarp at the top of the ring around $C_{(5)}$–$C_{(4)}$. One of the isopropyl groups on DIP lies very close to the region of $C_{(5)}$–$C_{(4)}$–$C_{(3)}$ (Fig. 3), so

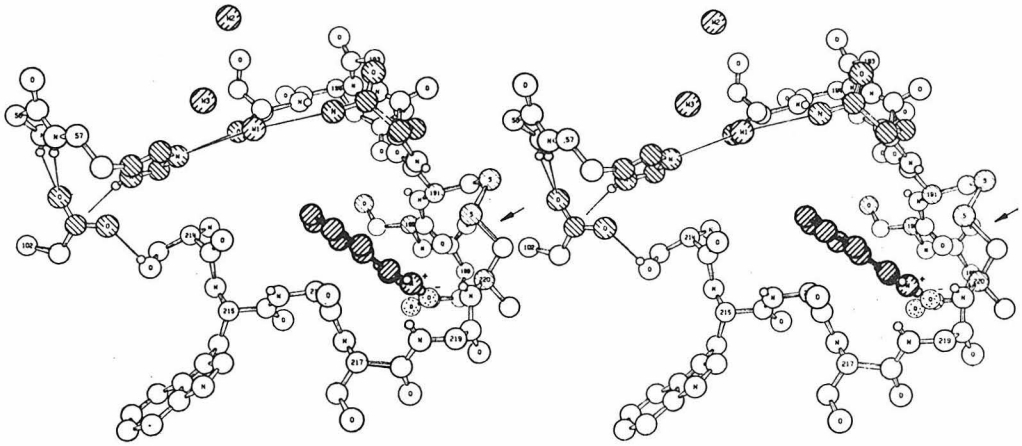


FIG. 4. The structure around the active site and binding pocket in BA-trypsin. Benzamidine and solvent molecules are indicated by right-handed shading. Parts of the trypsin molecule in this region which adopt different conformations in BA-trypsin and DIP-trypsin are shown in left handed shading. The general "sandwiching" effect of pocket closure is not shown. In the crystal structure this movement is due to repositioning of the chain between residues 190 and 191 and the cystine bridge 191–220. Other changes occur around the left-hand side of Asp102. This view is approximately 90° from Plate I and Fig. 3, about an almost vertical axis. This view has been adopted as "standard" and rotated to match the structure of α-chymotrypsin (Birktoft & Blow, 1972) for ease of comparison. Asp189 is dotted and forms a salt bridge with benzamidine. This and the other computer-plotted Figures (5 and 6) were drawn by a modified version of ORTEP (Johnson, 1965).

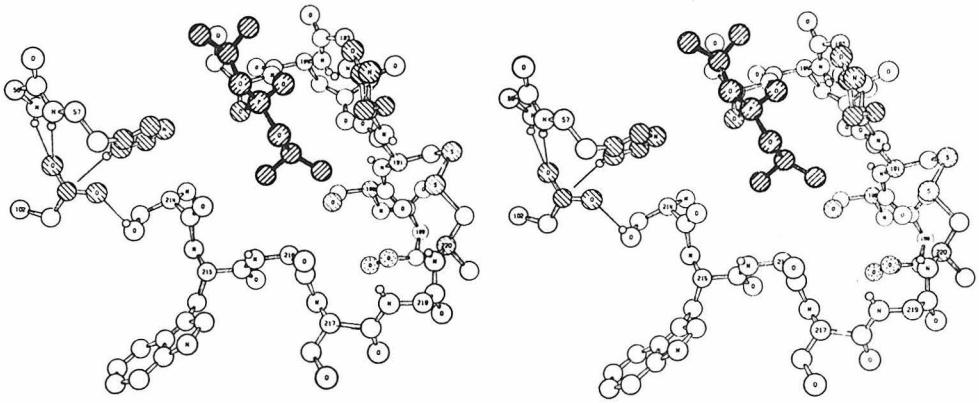


FIG. 5. The structure around the active site in DIP-trypsin. Right-handed shading picks out the DIP group attached to Ser195. One isopropyl group partly fills the neck of the binding pocket approximately in the same position as $C_{(3)}$–$C_{(4)}$–$C_{(5)}$ in BA-trypsin (Fig. 4).

that this region in the difference map (Plate I) corresponds to the difference between part of the firmly bound benzamidine and an isopropyl group vibrating about the same general location. The sharp concave scarp between benzamidine and DIP corresponds to the region of maximal difference density gradients between them.

The positive region associated with Gln192 is matched by a negative region behind the sections shown in Plate I. These peaks define the complete reorientation of

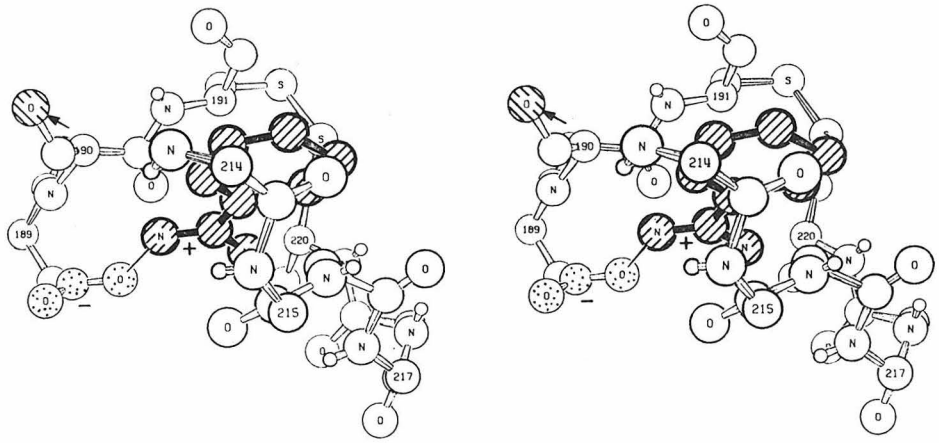


FIG. 6. A view of benzamidine in the specific binding pocket oriented 90° about a vertical axis from Fig. 4.

TABLE 3

Benzamidine-inhibited trypsin: co-ordinates for benzamidine, serine 190, serine 195, glutamine 192, and associated solvent molecules

	x	*y*	*z*	Atom
Benzamidine	0·537	0·260	0·764	BC1
	0·538	0·233	0·756	BC2
	0·523	0·224	0·740	BC3
	0·511	0·245	0·729	BC4
	0·510	0·268	0·736	BC5
	0·523	0·275	0·753	BC6
	0·551	0·268	0·781	BC7
	0·563	0·253	0·792	BN1
	0·550	0·290	0·786	BN2
Active-site solvent	0·498	0·188	0·667	BW1
	0·510	0·174	0·605	BW2
	0·476	0·154	0·643	BW3
Ser190	0·580	0·327	0·765	190CA
	0·554	0·336	0·760	190CB
	0·551	0·349	0·743	190OG
Gln192	0·604	0·208	0·699	192CB
	0·592	0·184	0·696	192CG
	0·578	0·190	0·674	192CD
	0·554	0·184	0·673	192NE2
	0·590	0·197	0·660	192OE1
Ser195	0·482	0·250	0·666	195OG

TABLE 4

Co-ordinates for the active centre in diisopropyl fluorophosphate-inhibited trypsin

	x	y	z	Atom
DIP	0·514	0·230	0·672	DIPP
	0·537	0·241	0·667	DIPO1
	0·505	0·211	0·656	DIPOU
	0·520	0·216	0·695	DIPOD
	0·504	0·224	0·636	DIPCU1
	0·526	0·226	0·625	DIPCU2
	0·489	0·205	0·619	DIPCU3
	0·524	0·238	0·711	DIPCD1
	0·542	0·234	0·728	DIPCD2
	0·499	0·246	0·720	DIPCD3
Asp189	0·613	0·358	0·811	189CA
	0·602	0·339	0·827	189CB
	0·576	0·334	0·824	189CG
	0·560	0·349	0·827	189OD2
	0·569	0·316	0·819	189OD1
Ser195	0·497	0·249	0·675	1950G
Gln192	0·603	0·208	0·699	192CB
	0·586	0·189	0·695	192CG
	0·592	0·181	0·673	192CD
	0·596	0·159	0·672	192NE1
	0·593	0·194	0·659	192OE2
Ser190	0·580	0·327	0·765	190CA
	0·552	0·326	0·760	190CB
	0·544	0·302	0·761	190OG

Gln192 in BA-trypsin, as shown in Figures 4 and 5. Gln192 acts as a polar flap covering the entrance to the binding pocket and must necessarily undergo a displacement on binding substrate side chains which interact with Asp189.

There are several other changes in the binding pocket when benzamidine is bound. There is a significant "squeeze" of the pocket, which leads to a close sandwiching of the inhibitor between the peptide planes of 190–191, and 215–216 (indicated in Figs 4 and 6). The movement is only visible in the right-hand side of the pocket seen in Figure 4. No movement is seen around Ser214 (which is hydrogen-bonded to Asp102). The disulfide bridge 191–220 moves in toward benzamidine by $0{\cdot}7 \pm 0{\cdot}3$ Å. The side chain of Ser190 rotates by 120° about its α–β bond and the γ-oxygen forms a hydrogen bond to the hydroxyl group of Tyr228, and possibly to the carboxyl of Asp194. The position of Ser190 in DIP-trypsin is sterically incompatible with binding of benzamidine or an arginine side chain, so it presumably undergoes this type of burial on binding specific substrates or their analogues. The positions of inhibitor groups, and of groups in significantly different positions in BA-trypsin and DIP-trypsin are listed in Tables 3 and 4.

(b) *The active site*

Readjustments in the active site due to the BA-trypsin–DIP-trypsin difference include a shift in the imidazole ring of His57, binding of additional solvent molecules

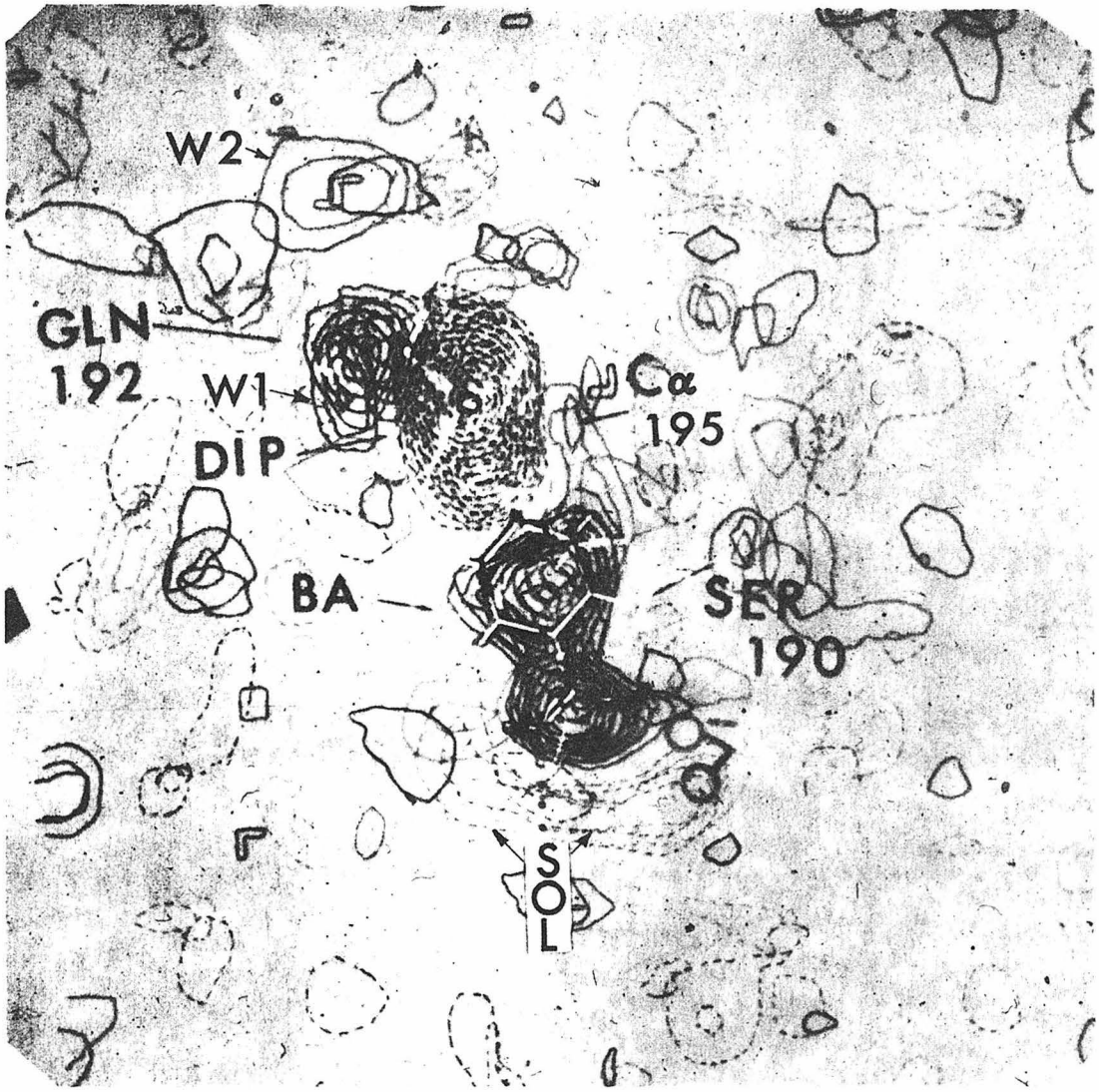


PLATE I. Five sections of the difference density map showing a block 25·2 Å (horizontal Δy = 0·04 to 0·5) by 22·0 Å (vertical Δz = 0·54 to 0·92) by 5·5 Å (thick Δx = 0·41 to 0·51). Contours are at $\pm$ [0·10 (0·05)] e/Å^3, and negative density is shown with dotted contours; P corresponds to the phosphorus atom position in DIP-trypsin. Similarly, DIP, Gln192 (in BA-trypsin), and BA (benzamidine) indicate their associated densities. Superimposed on the sections is a part of the skeletal model of trypsin showing the DIP group attached to Ser195. Benzamidine (BA) and solvent molecules (W1, W2, etc.) have been added to the model. The conformations of Gln192 and Ser190 are for DIP-trypsin, and the carboxyl of Asp189 is indicated. The negative region (SOL) corresponds to the displaced solvent in BA-trypsin.

[*facing p. 218*

and some rearrangements in the side chains of Leu99, Val 227, Ser195 and Asp102. Two complementary positive and negative peaks in the difference map (see Table 2) indicate a rotation about the α–β bond of His57, a reorientation of the carboxyl group of Asp102, and the binding of a solvent molecule in the active site (W3). It is difficult to estimate the extent of movement of His57 from the difference peaks,

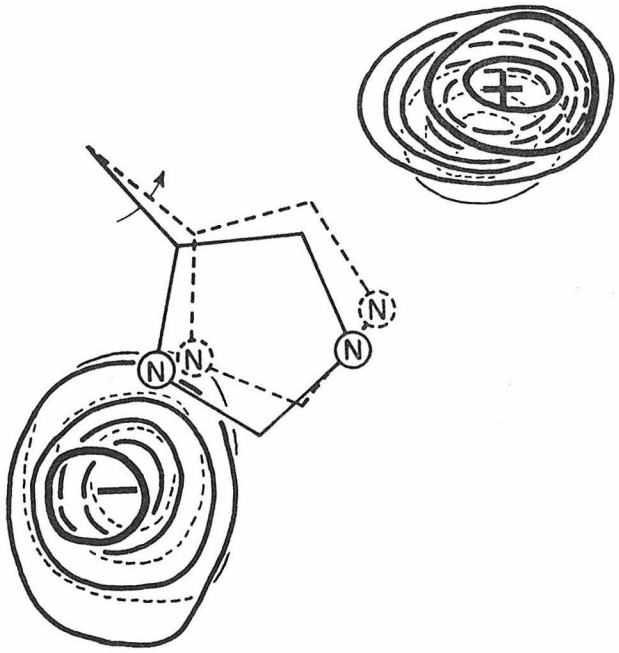


FIG. 7. The peaks associated with His57 on the difference map. The lower peak is negative density (−) while the other one is positive (+). The latter peak is a composite with a solvent molecule density (see text).

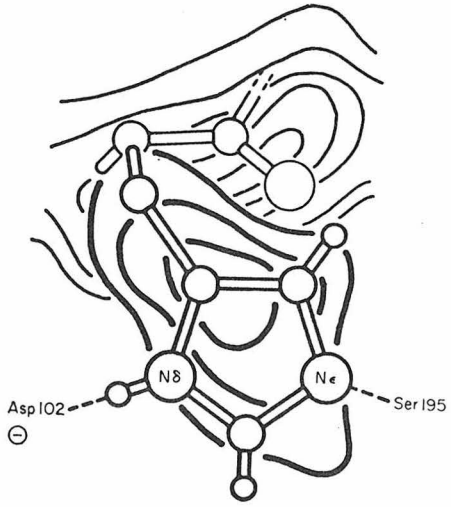


FIG. 8. Electron density for His57 in the DIP-trypsin Fourier map, computed for the plane parallel to the imidazole ring.

because both the movement of His57 and the solvent molecule W3 contribute to the positive peak and this solvent molecule shifts the peak away from the imidazole ring. The negative peak is also a hybrid due to rearrangements of the Asp102 and His57 side chains.

As a result, the vector moment of these two peaks comes to 36·4 electron Å, which would indicate a falsely large movement (1·04 Å) of the imidazole containing 35 electrons. The two peaks are shown in Figure 7, where the imidazole ring as it appears in DIP-trypsin is indicated by a solid line drawing. (It is clear from this Figure that the positive peak is too far from the imidazole to be due to histidine movement alone.) The movement of the imidazole was therefore measured by comparing a density map $\rho'(\mathbf{r})$ computed using terms $(2|F_{BAT}|-|F_{DIPT}|)$, ϕ_P and the original $\rho(\mathbf{r})$ map for DIP-trypsin. The well-defined electron density for the imidazole in the latter map is shown in Figure 8 (and also in Plate IV of Stroud *et al.*, 1974). The shift was evaluated from the concerted movements of the side-chain density between the two maps calculated from

$$\Delta x = \frac{\int_x \rho'(\mathbf{r})\cdot\mathbf{r}\,d\mathbf{x}}{\int_x \rho'(\mathbf{r})\cdot d\mathbf{x}} - \frac{\int_x \rho(\mathbf{r})\cdot\mathbf{r}\,d\mathbf{x}}{\int_x \rho(\mathbf{r})\cdot d\mathbf{x}},$$

where $\mathbf{x}$ is the line joining the two difference peaks in the direction of rotation. This expression is essentially a $\Delta\mathbf{x}$ difference in centers of gravity of the peaks between $\rho'(\mathbf{r})$ and $\rho(\mathbf{r})$ and accounts for changes in thermal motion parameters. The imidazole ring of His57 has a somewhat larger thermal motion in BA-trypsin than in DIP-trypsin. In BA-trypsin the imidazole ring has moved $0{\cdot}20\pm0{\cdot}07$ Å away from Asp102 by rotation about the C_α–C_β bond.

This movement is insufficient to account for all of the negative peak between Asp102 and the ring. The remainder of this peak indicates that Asp102 has also moved away from the imidazole slightly in BA-trypsin. A corresponding positive region is not seen in the map and has presumably been relayed to the surface of the molecule by compensatory movements around Asp102. The side chains of Leu99 and Val227, for example, rotate about their C_α–C_β bonds, such that their hydrophobic side chains move closer to the indole ring of Trp215. A solvent molecule not previously observed in DIP-trypsin is hydrogen-bonded to the nitrogen of the indole ring in BA-trypsin.

In BA-trypsin, as in DIP-trypsin (Stroud *et al.*, 1974), one of the oxygen atoms on the carboxyl group of Asp102 is hydrogen-bonded to two backbone N–H groups (56 and 57), while the other oxygen forms a hydrogen bond to the hydroxyl of Ser214. The carboxyl group is shielded from solvent by the side chains of Ile99, Tyr94 and Leu57, such that there is no direct access to solvent. Although there is room for one, or perhaps two, water molecules to bind in a cavity behind Ile99, it would be impossible for a water molecule in this cavity to hydrogen-bond to Asp102 without disrupting the hydrogen bond to Ser214.

The γ-oxygen of Ser195, no longer tethered to the DIP group, is free to move; however, the large negative DIP density obscures any sign of the γ-oxygen position in the difference map. There are two almost equal and separate peaks for the hydroxyl group in the (2 BA-trypsin–DIP) electron density map. One of these lies close to the position found in stable acyl enzymes, indole acryloyl and carbamyl chymotrypsins (Henderson, 1970; Robillard *et al.*, 1972). The second site involves an upward rotation of 120° about the α–β bond of the serine, making it close to the position found

for Ser195 in native α-chymotrypsin at pH 4·5 (Steitz *et al.*, 1969). Based on our assigned position for His57, the hydroxyl group of Ser195 would form a more favorable hydrogen bond to the imidazole in the first (acyl enzyme) position, as shown in Figure 4. In spite of the clarity of the imidazole plane in density maps (Fig. 8 of this paper, and Plate IV of Stroud *et al.*, 1974), exact assignment of the tilt of a ring within five or ten degrees from a 2·7 Å map is always open to some small unpredictable error, which should be kept in mind. However, this structure differs significantly from what we might expect to find for native trypsin, based on the structures of other native serine proteases. The uniquely high autolysis rate in trypsin has delayed preparation of native trypsin crystals; however, the crystallographic analysis of native trypsin at pH 7·5 is now under way.

It seems most probable that the side chains of Asp102, His57 and Ser195 in native trypsin are arranged as they are in native chymotrypsin. If transition to the acyl-like position is a direct consequence of binding the side-chain analogue benzamidine, it seems reasonable that it could also be a consequence of binding substrate side chains during formation of the Michaelis complex at pH 8·0.

(c) *Other differences*

There seems to be an overall increase in the number of solvent molecules specifically bound to the surface of BA-trypsin compared with DIP-trypsin, which may be due in part to slightly higher salt concentrations in BA-trypsin crystals. On the surface of the protein the side chains of Lys145, Asn25 and Val200 move slightly outward, and there are small shifts along the backbone chain between Tyr94 and Leu99. Solvent molecules that were bound between Lys204 and Cys202, between Lys109 and Ser84, and in a small crevice between Tyr172 and the cystine bridge 182–168 are no longer bound in BA-trypsin, while newly bound solvent appears near Ser96 and Ser110. Within the globular structure there is a small inward shift of the chain Ile47–Ser45, with the hydroxyl group of Ser45 rotating approximately 180° and Pro198 tilting slightly to accommodate the hydroxyl groups' new orientation.

There is a crevice in the surface of the molecule behind the specificity pocket and between Gly188 and Lys145 that is normally filled by a solvent molecule hydrogen-bonded to the carbonyl group of Cys220 (in one $HgBr_2$ derivative of trypsin, a heavy-atom group fills this crevice, Stroud *et al.*, 1974). When the lining of the specificity pocket moves in towards benzamidine, the molecule in the surface crevice is lost.

A solvent molecule W1 is hydrogen-bonded to the ε-nitrogen of His57, the amide moiety of Gln 192 and is at hydrogen bond distance from the hydroxyl of Ser195 in BA-trypsin (Fig. 5). Another solvent molecule, W2, lies 4·4 Å above W1.

4. Discussion

Benzamidine binds in the specificity pocket of trypsin, which normally binds the positively charged side chains of lysine or arginine during proteolysis. How much can we now infer about the binding of real substrates by trypsin? If BA-trypsin is a good model for side-chain binding, why is BA-trypsin an inhibitor even for non-specific substrates?

(a) *The binding pocket and specificity of trypsin: a model for substrate binding*

The entrance to the binding pocket on the enzyme is a slit that provides enough room for the entry of a lysine or arginine side chain. Binding of benzamidine leads to induced changes in the specificity pocket that involve displacement of the γ-oxygen of Ser190, and the closure of the pocket to sandwich the benzamidine molecule between the peptide planes of 190–191 and 215–216. These close contacts are presumably made between the methylene groups of lysine and arginine and the enzyme in normal substrate binding. Benzamidine contains an amidinium cation, coplanar with the phenyl ring, which closely resembles the guanidinium cation of an arginine side chain (Fig. 1). Therefore, we built a model of an arginine amide bound to trypsin, such that the guanidinium cation superposed exactly onto the position of the amidinium moiety in BA-trypsin. The carbonyl carbon of arginine, which forms an acyl bond to the enzyme *via* Ser195 during catalysis, is brought to within 0·3 Å of the phosphorus position of DIP. The phosphorus is covalently bound to Ser195 in DIP-trypsin. This model allows the carbonyl group of arginine to adopt a position very close to that postulated for a "good" substrate (Henderson, 1970).

In a second model experiment, the model substrate was built from the serine-bound end, using bond lengths and angles compatible with an acyl enzyme in which the serine γ-oxygen is bonded to the carboxyl carbon of arginine. The arginine side-chain was constricted to lie in the benzamidine plane. In this case, the guanidinium cation ends up rotated slightly with respect to the amidinium group of benzamidine (see Fig. 9). Slight readjustments of the arginine side-chain confor-

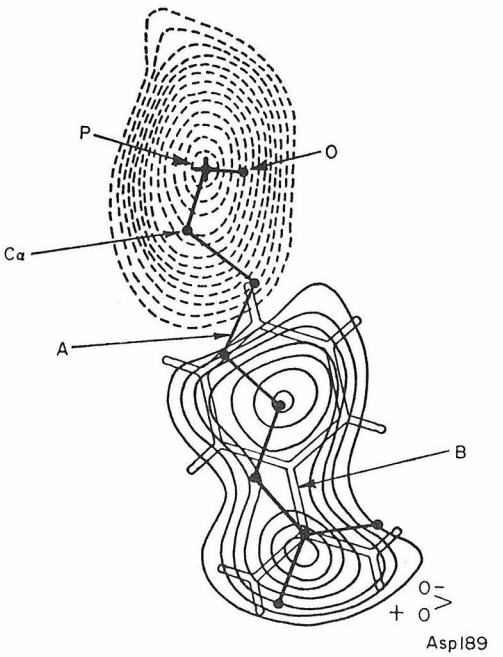


FIG. 9. A section through the difference map (BA-trypsin–DIP-trypsin). The position of benzamidine is shown in open line. An arginine side chain (A) of an acyl enzyme is placed so that the carbonyl carbon is at the same position as the phosphorus (P) of DIP-trypsin. The side chain is assumed to be planar, bringing the cation very close to the benzamidine cation. The scarp between benzamidine and DIP is clearly visible.

mation to preserve cation–anion interactions lead to a model in which $N_{(2)}$ and the carboxyl of Asp189 are about 2·9 Å apart.

In a third model-building experiment, the structure of the bovine pancreatic trypsin inhibitor (Huber *et al.*, 1970) was used in conjunction with the BA-trypsin structure to generate a model for substrate binding. An arginine side-chain was

TABLE 5

Model for substrate binding fractional co-ordinates of residues 13 to 17 of the "arginine-15" inhibitor and glutamine acid 192 shown in Figure 10

	x	*y*	*z*	Atom
Pro13	0·467	0·137	0·745	13CA
	0·469	0·165	0·739	13C
	0·468	0·180	0·753	13O
Cys14	0·469	0·167	0·720	14N
	0·470	0·158	0·714	14NH
	0·469	0·191	0·711	14CA
	0·456	0·189	0·691	14CB
	0·496	0·200	0·708	14C
	0·513	0·186	0·706	14O
"Arg"15	0·495	0·223	0·709	15N
	0·487	0·231	0·710	15NH
	0·519	0·235	0·704	15CA
	0·520	0·258	0·715	15CB
	0·518	0·255	0·739	15CG
	0·542	0·264	0·751	15CD
	0·545	0·252	0·770	15NE
	0·552	0·261	0·786	15CW
	0·553	0·285	0·787	15N2
	0·556	0·248	0·801	15N1
	0·520	0·242	0·684	15C
	0·531	0·256	0·678	15O
Ala16	0·509	0·227	0·672	16N
	0·503	0·214	0·675	16NH
	0·508	0·229	0·650	16CA
	0·483	0·222	0·643	16CB
	0·530	0·215	0·639	16C
	0·547	0·201	0·648	16O
Arg17	0·529	0·215	0·618	17N
	0·521	0·222	0·614	17NH
Gln192	0·601	0·204	0·699	192CB
	0·612	0·194	0·680	192CG
	0·590	0·170	0·682	192CD
	0·569	0·170	0·671	192NE1
	0·597	0·155	0·692	192OE2

built into the inhibitor structure replacing the side chain of Lys15. Models of trypsin and the modified inhibitor were brought together such that the "Arg 15" side-chain was superposed onto the benzamidine density in the difference map (Fig. 9). The side chain of Gln192 was moved to accommodate the inhibitor. Otherwise, the trypsin active-site configuration was that of BA-trypsin. The resulting model complex

was essentially the same as that proposed earlier by Stroud *et al.* (1971). Bond lengths of some of the previously predicted hydrogen bonds are listed in Table 6, and the structure of the complex around the active center of trypsin is shown in Figure 10.

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 binding of specific substrate side-chains. Benzamidine binding is, therefore, a good model for substrate side-chain binding, showing

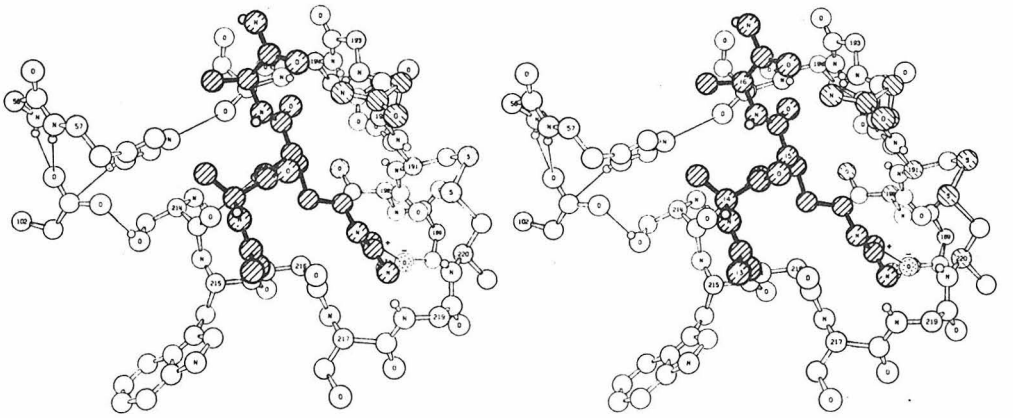


FIG. 10. A model for the binding of the bovine pancreatic trypsin inhibitor (right-handed shading) to trypsin (Stroud *et al.*, 1971), in which the side chain of Lys15 has been replaced by an arginine side-chain, matching the electron density for benzamidine. Co-ordinates for this model are listed in Table 5.

a number of interactions that we expect to be present during the specific interaction between real substrates and trypsin.

This important correspondence helps to explain why the length of the basic side-chain, as well as its charge, is important in catalysis. When the side chains of arginine or lysine esters are shortened or lengthened by a methylene group (Baines *et al.*, 1964; Baird *et al.*, 1965), the rate of tryptic hydrolysis drops markedly. The BA-trypsin structure explains this sharp specificity. The β-carboxyl group of Asp189 is tied to a fixed position by the orientation of the peptide chain and by hydrogen bonding. There is no movement of this carboxyl group on binding benzamidine; thus, even though there are changes in the binding pocket region, the position of the critical charge on Asp189 is stabilized in the specificity pocket and the orientation of this carboxyl group is unperturbed on binding. If an ester or amide of a basic aliphatic amino acid with a side chain one methylene group shorter, or one longer, than normal arginine or lysine side-chains is attacked by trypsin, the coulombic attraction between the positive side-chain and the negative carboxyl group of Asp189 will either pull or push the susceptible bond of the substrate away from the position and orientation required for optimal catalysis.

Data for the hydrolysis of various ester analogues of benzamidine and phenyl guanidine (Mares-Guia *et al.*, 1967) shed more light on the relation between the length of the side chain and catalytic activity, and lead to the questions of how benzamidine inhibits trypsin, and how the structure of BA-trypsin relates to the

TABLE 6

Some of the hydrogen bonds between the "arginine-15" trypsin inhibitor and trypsin

Trypsin	Inhibitor	Distance (Å)
Ser214 C=O	N—H Arg15	3·0
Gly216 N—H	C=O Pro13	2·8
Ser195 N—H	C=O Arg15	2·5
Gly193 N—H	C=O Arg15	3·2

The Table gives the distances between amide nitrogens and the carbonyl oxygens.

free enzyme and the enzyme substrate complex with specific substrates. Mares-Guia *et al.* (1967) report that trypsin catalyzed the hydrolysis of ethyl and methyl esters of *p*-guanidino and *p*-amidinobenzoic and phenyl acetic acids. The distance between the charged groups and the ester moiety in true substrates is more closely approximated in the phenyl acetates than in the benzoates (see Fig. 1), and the phenyl acetates are in fact hydrolyzed more rapidly. From the model of BA-trypsin (Fig. 5) the ester group of trypsin-ethyl *m*-amidinobenzoate would be forced to project away from the enzyme outward into the solvent, and this explains the observation that even though this ester binds more tightly than its *para* analogue, it is not hydrolyzed by trypsin. If trypsin can still hydrolyze esters of compounds resembling benzamidine, does benzamidine competitively inhibit trypsin by merely filling up the binding pocket and thereby sterically interfere with substrate binding, or do induced structural changes in the enzyme itself contribute to inhibition?

(b) *Benzamidine inhibitor action*

Model building shows that the top of the phenyl ring in benzamidine ($C_{(3)}$–$C_{(4)}$–$C_{(5)}$) lies too close to the γ-oxygen of Ser195 to permit substrates to approach and form an acyl-serine. For example, a prohibitively close contact (<2 Å) would occur between $C_{(5)}$ benzamidine and the α-carbon of a glycyl substrate bound in the manner shown in Figure 10.

While large alkyl amines, guanidines and amidines competitively inhibit trypsin†, smaller amines (methyl, ethyl and *n*-propyl) and guanidines (methyl and ethyl), which are also competitive for specific substrates, promote the slow hydrolysis of the non-specific substrate acetyl glycine ethyl ester up to 11 times of that without these side-chain analogues. Far from contributing to inhibition, as is the case with some of the larger cations, these small positively charged molecules enhance the apparent reactivity of trypsin (Inagami & Murachi, 1964; Inagami & York, 1968; Erlanger & Castleman, 1964; Mares-Guia & Shaw, 1965). Thus, the BA-trypsin structure and these data indicate that benzamidine and other large amines, amidines

† Erlanger & Castleman (1964) have observed the tryptic hydrolysis of acetyl glycine ethyl ester activated by 2-aminoheptane and other large alkyl amines. It is possible that these large amines no longer bind in the specificity pocket, but rather at some second site (Sanborn & Hein, 1968), and they might affect catalysis in the same way that excesses of specific substrates act (substrate activation) (Trowbridge *et al.*, 1963).

and guanidines competitively inhibit trypsin by sterically hindering substrate accessibility, not only to the specificity pocket but also to the active site itself.

(c) *Histidine 57–aspartic acid 102*

It is clear from the density maps that the δ-nitrogen of His57 in BA-trypsin or DIP-trypsin points directly towards the center of the carboxyl group of Asp102 (see Plate IV of Stroud *et al.*, 1974) rather than to the lower carboxyl oxygen (as found in the α-chymotrypsin structure of Birktoft & Blow (1972) at pH 4·5). Our analysis was done at pH 8·0, where the native enzyme would be active. M. Hunkapiller, *et al.* (1973) have shown that the pK_a of the equivalent residue to Asp102 in α-lytic protease (a bacterial serine protease related to the trypsin family, showing similar kinetics to elastase) is 6·7; therefore, the enzyme is inactive until Asp102 becomes charged (above pH 6·7). Consequently, the interaction we observe between the neutral His57 and the ionized Asp102 is one in which the protonated δ-nitrogen interacts with the delocalized negative charge on the carboxylic acid.

Structure analysis carried out in the acid pH range (where the enzyme is inactive) show interaction between a neutral Asp102 and neutral imidazole. (The imidazole is neutral above pH 4·5, Hunkapiller *et al.* (1973).) Thus Asp102 is the base during catalysis and its pK_a, rather than that of His57, is responsible for the low limb of the pH activity profile in the serine proteases. The imidazole presumably acts as a proton relay between Asp102 and the substrate or solvent, and remains neutral throughout catalysis. Therefore, the effective pK_a of the Asp102–His57 system as seen in kinetics is the pK_a of the better base, Asp102 ($pK_a = 6{\cdot}7$). The mechanistic importance of this assignment is that no unfavorable charge separation is required during catalysis. In a proposed mechanism incorporating this assignment of pK_a values (Stroud & Krieger, *9th Int. Congr. Biochem.*, 1973), negative charge is developed on the tetrahedral intermediate as negative charge is neutralized on the carboxyl group of Asp102. At no stage is the imidazole required to be positively charged, so smoothing the barriers between intermediate states during hydrolysis.

If, on the other hand, the pK_a of Asp102 were low, or even normal for aspartic acid in solution, protonation of Asp102 would be unlikely even during the hydrolysis reaction. Then the role of Asp102 in enhancing enzymatic rate would be difficult to explain in other than purely structural terms. For example, its role might be to maintain and stabilize the orientation of His57 during catalysis. However, most other assignments of pK_a values from kinetic data are inherently incapable of distinguishing between the pK_a values of Asp102 and His57. Likewise, studies of the net proton release upon denaturation (Ferscht & Sperling, 1973) or competitive labeling techniques (Cruickshank & Kaplan, 1973; Beeley & Neurath, 1968) can only conclude that the Asp–His *system* has an apparent pK_a near neutrality and, therefore, that one of these groups loses a proton as the pH is raised through neutrality.

Thus the presence of a carboxyl group of high pK_a and a neutral side-chain of His57 with a low pK_a suggests two compelling new evolutionary reasons why the Asp–His–Ser arrangement should be universal to serine proteases. First, by neutralizing a negative charge on Asp102, rather than generating a positive charge on His57 during formation of the tetrahedral intermediate, there is no unfavorable charge separation. This contributes to reducing transition-state internal energies, and so to rate enhancement. Second, if the charged Asp102 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 His57 is ideally suited to both insulate Asp102 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 Asp102 and His57 and the structure which shows a symmetric interaction between the charge on Asp102 and His57, are unlike the situation in aqueous solution and more like interactions and proton affinities expected for groups in the gas phase or in hydrophobic solvents. The protein environment of Asp102 is responsible for these effects.

(d) *The role of glutamine 192*

The side chain of Gln192 covers the entrance to the specific binding pocket in BA-trypsin and forms a hydrogen bond to W1 (Fig. 4), which in turn is hydrogen-bonded to the ϵ-nitrogen of His57 or to the γ-oxygen of Ser195 in the "down" position. This orientation is completely different from that found in DIP-trypsin. If Gln192 adopts the BA-trypsin conformation in native trypsin (a lid to the binding pocket) then it must move away during substrate binding. It is also possible that Gln192 might "pack" against a substrate once it has formed an enzyme–substrate complex. This residue (192) is glutamine in five different trypsin species sequenced so far (listed by Stroud *et al.*, 1971), whereas in bovine chymotrypsin it is a methionine residue. Perhaps this substitution serves to offer a polar environment to exchange with polar side-chains in trypsin substrates, and a non-polar, hydrophobic environment for side chains of chymotrypsin substrates. A reorientation similar to that found for Gln192 has been observed in chymotrypsin (Met192) by Steitz *et al.* (1969).

The side chain of residue 192 might also have an important role in stabilizing the conformation of the specificity pocket region in the zymogen precursors of trypsin and chymotrypsin. In the crystal structure of chymotrypsinogen (Freer *et al.*, 1970), the polypeptide chain lining the binding pocket is rotated away from the chymotrypsin orientation, and the side chain of Met192 fills the pocket. In the zymogen Met192 could be stabilizing the zymogen pocket through "hydrophobic" interactions. Gln192 could play a similar role in trypsinogen by rotating downward, and possibly hydrogen-bonding to Asp189. (The structure analysis of bovine trypsinogen is nearing completion in our laboratory.)

(e) *Side-chain activation of trypsin*

The rate of hydrolysis of specific substrates by trypsin is much faster than the rate for non-specific amides or esters (Inagami & Mitsuda, 1964); however, a marked increase in the rate of hydrolysis of acetyl glycine ethyl ester, a non-specific substrate, accompanies the binding of small alkyl amines or guanidines to the trypsin specificity pocket. Furthermore, this effect is primarily an effect on k_{cat}, with a negligible effect on K_m (Inagami & Murachi, 1964). Inagami & York (1968) also showed that the binding of these small cations shifts the lower limb of the pH profile of acetyl glycine ethyl ester hydrolysis to lower pH values. For example, on binding methyl guanidine, the observed pK_a shifts by 0·4 unit. On the basis of the recent determination of pK_a values by Hunkapiller *et al.* (1973), we assume that this group with a pK_a around neutrality is Asp102. The pK_a change contributes to a rise in V_{max}; however, even after correction for this effect, methyl guanidine or ethyl amine still raise V_{max} by factors of approximately 4 and 11, respectively, clearly implicating

other induced effects. Ester substrates of chymotrypsin show a similar relation between binding affinity and the pK_a of Asp102, indicating that charge interactions are insufficient to explain the pK_a change. If the binding of benzamidine to trypsin induces a similar activation state, which is masked by its action as a steric inhibitor, the BA-trypsin structure should resemble the activated state, and might suggest how the alkyl cations activate hydrolysis of non-specific substrates. When the structure of native trypsin (work in progress) is completed, the changes induced by benzamidine (or the small cations) can be assessed directly; however, there is a good deal of evidence indicating that induced structural changes are involved. The masking of tyrosine and tryptophan residues in BA-trypsin, or *n*-butylamine-trypsin are examples (Villanueva & Herskovits, 1971; D'Albis & Béchet, 1967). Lysine residues are also affected by benzamidine binding (Beaven & Gratzer, personal communication). Following this evidence, it may be that one function of side-chain binding on specific substrates is to reposition parts of the enzyme structure so as to enhance the catalytic efficiency. In any event "induced fit" (Koshland, 1958) is the most reasonable explanation of the small alkyl amine or guanidine effect on hydrolysis of non-specific substrates by trypsin.

The rate-limiting step for hydrolysis of esters by trypsin is deacylation. Therefore, the primary effect detected by Inagami & York (1968) is one on the deacylation rate-constant. In order to assess the effect of small cations on the acylation rate-constant, Chambers & Stroud (unpublished results) have studied the effect of alkyl cations on hydrolysis of other non-specific substrates. We have shown that methyl guanidine also activates hydrolysis of *p*-nitrophenyl acetate by trypsin, and that a major part of this activation is due to an increase in the deacylation rate-constant. The preliminary indication is that changes in internal energy of the acyl enzyme could explain the major contribution to side-chain activation.

Other explanations are less likely. For example, small alkyl amines or guanidines might stabilize an active form of the enzyme *versus* an inactive one, thus apparently raising the catalytic rate-constant. An equilibrium such as this has been demonstrated for chymotrypsin (Ferscht & Requena, 1971), where the transition from inactive to active enzyme is correlated directly with the formation of a buried salt bridge between the α-amino group of Ile16 and the carboxyl group of Asp194. At neutral pH, about 10% of the enzyme exists in the inactive form, which does not bind specific substrates and in which the amino-terminal is assumed to be outside of the molecule (Ferscht, 1972). Trypsin has a similar internal salt bridge between Asp194 and the α-amino group of Ile16, although a similar conformational equilibrium has not yet been established. However, we expect the salt bridge to be much stronger in trypsin than it is in δ or α-chymotrypsin (thus leading to a much smaller proportion of "inactive" enzyme), because the apparent pK_a defining the high pH limb of the pH-activity profile associated with the α-amino group of Ile16 is more than 1·5 pK_a units above that found for chymotrypsin. Typical values are about pK_{app} 8·5 to 8·9 for chymotrypsin, and pK_{app} of 10·1 for hydrolysis of α-*N*-benzoyl-L-argininamide by trypsin (Spomer & Wootton, 1971). The indication is that the salt bridge is much stronger in trypsin, and, therefore, that the proportion of inactive enzyme, where the α-amino terminus is no longer constrained in the salt bridge, is much smaller.

The binding of indole and DIP to chymotrypsin (Hess *et al.*, 1970) and the binding of *n*-butylamine (D'Albis & Béchet, 1967) or DIP (Ghelis *et al.*, 1967) to trypsin appear

to stabilize the enzymes in a conformation similar to the active, or "in" form. By pushing the conformational equilibrium toward the active enzyme, the specific side chain analogue activators could produce an increase in V_{max} for non-specific substrate reactions, but could not in any case explain a differential effect on deacylation or acylation rate-constants as we observe for hydrolysis of nitrophenyl acetate. Furthermore, this effect cannot be responsible for most of the observed increase in V_{max}, for even if the active–inactive trypsin equilibrium were similar to that of chymotrypsin, only 10% of the enzyme would be affected, clearly less than required for a factor of 11 in k_{cat}. Also, one would expect those alkyl cation effectors that bind most strongly to trypsin's specificity pocket to have the greatest effect on the conformational equilibrium; however, alkyl amines do not bind as tightly to trypsin as do their corresponding guanidines, yet the amines of corresponding size are better activators.

In summary, it seems inescapable that the small alkyl amines or guanidines induce small changes in the catalytic site of trypsin, and that these changes are responsible for activating the enzyme and affecting k_{cat}. The specific side chains of real substrates may be responsible for similar rearrangements. It seems likely that the structure of BA-trypsin resembles an "activated" species, although the enzyme is inhibited by steric hindrance of productive substrate binding. If this is indeed true, we might expect the side chains of real specific substrates to induce similar effects in the enzyme.

We thank Robert Huber for sending us co-ordinates of the trypsin inhibitor, and a preprint of some of his results. We are grateful to Lalli Samson for preparation of the ORTEP drawings in Figs 5, 6 and 7, and to Michael Hunkapiller, Steve Smallcombe and John Richards for permission to cite some of their results. We thank Sten Samson for his help and advice in rebuilding our diffractometer system.

This work is contribution no. 4553 from the Norman W. Church Laboratory of Chemical Biology of the California Institute of Technology. This work has been carried out with the support of the United States Public Health Service (grant nos GM12121 and GM19984), whose help is gratefully acknowledged. One of us (M. K.) is the recipient of a Danforth Graduate Fellowship, and another (R. M. S.) is the recipient of a National Institutes of Health Career Award (GM70469).

REFERENCES

Baines, N. J., Baird, J. B. & Elmore, D. T. (1964). *Biochem. J.* **90**, 470–476.
Baird, J. B., Curragh, E. G. & Elmore, D. T. (1965). *Biochem. J.* **96**, 733–738.
Beeley, J. G. & Neurath, H. (1968). *Biochemistry*, **7**, 1239–1251.
Birktoft, J. J. & Blow, D. M. (1972). *J. Mol. Biol.* **68**, 187–240.
Cruikshank, W. H. & Kaplan, H. (1973). *Biochem. J.* **130**, 1125–1131.
Cunningham, L. (1954). *J. Biol. Chem.* **211**, 13–19.
D'Albis, A. & Béchet, J.-J. (1967). *Biochim. Biophys. Acta*, **140**, 435–458.
Erlanger, B. F. & Castleman, H. (1964). *Biochim. Biophys. Acta*, **85**, 507–509.
Ferscht, A. R. (1972). *J. Mol. Biol.* **64**, 497–509.
Ferscht, A. R. & Requena, Y. (1971). *J. Mol. Biol.* **60**, 279–290.
Ferscht, A. R. & Sperling, J. (1973). *J. Mol. Biol.* **74**, 137–149.
Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T. & Xuong, N. H. (1970). *Biochemistry*, **9**, 1997–2009.
Ghelis, C., Labouesse, J. & Labouesse, B. (1967). *Biochem. Biophys. Res. Commun.* **29**, 101–106.
Henderson, R. (1970). *J. Mol. Biol.* **54**, 341–354.
Henderson, R. & Moffat, J. K. (1971). *Acta Crystallogr. sect.* B**27**, 1414–1420.

Hess, G., McConn, J., Ku, E. & McConkey, C. (1970). *Phil. Trans. Roy. Soc. Ser. B*, **257**, 89–104.

Huber, R., Kukla, D., Rühlmann, A., Epp. O. & Fomanek, H. (1970). *Naturwissenshaften*, **57**, 389–392.

Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. R. & Richards, J. H. (1973). *Biochemistry*, **12**, 4732–4742.

Inagami, T. & Mitsuda, H. (1964). *J. Biol. Chem.* **239**, 1388–1394.

Inagami, T. & Murachi, T. (1964). *J. Biol. Chem.* **239**, 1395–1401.

Inagami, T. & York, S. S. (1968). *Biochemistry*, **7**, 4045–4052.

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

Koshland, D. E., Jr (1958). *Proc. Nat. Acad. Sci., U.S.A.* **44**, 98–104.

Lazdunski, M. & Delaage, M. (1967). *Biochim. Biophys. Acta*, **140**, 417–434.

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

Mares-Guia, M., Shaw, E. & Cohen, W. (1967). *J. Biol. Chem.* **242**, 5777–5781.

Northrop, J. H. & Kunitz, M. (1932). *J. Gen. Physiol*, **16**, 295–311.

Richards, F. M. (1968). *J. Mol. Biol.* **37**, 225.

Robillard, G. T., Powers, J. C. & Wilcox, P. E. (1972). *Biochemistry*, **11**, 1773–1784.

Sanborn, B. M. & Hein, E. G. (1968). *Biochemistry*, **7**, 3616–3624.

Schroeder, D. D. & Shaw, E. (1968). *J. Biol. Chem.* **243**, 2943–2949.

Spomer, W. E. & Wootton, J. F. (1971). *Biochim. Biophys. Acta*, **235**, 164–171.

Steitz, T. A., Henderson, R. & Blow, D. M. (1969). *J. Mol. Biol.* **46**, 337–348.

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

Stroud, R. M., Kay, L. M. & Dickerson, R. E. (1974). *J. Mol. Biol.* **83**, 185–208.

Trowbridge, C. G., Krehbiel, A. & Laskowski, M., Jr (1963). *Biochemistry*, **2**, 843–850.

Vandlen, R. & Tulinsky, A. (1973). *Biochemistry*, **12**, 4193–4200.

Villanueva, G. B. & Herskovits, T. T. (1971). *Biochemistry*, **10**, 4589–4594.

Weinstein, J. J. & Doolittle, R. F. (1972). *Biochim. Biophys. Acta*, **258**, 577–580.

APPENDIX II

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

R. E. Koeppe II, M. Krieger, and R. M. Stroud

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-incubation 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 x 10^{-4} M (K_m = 1 x 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. (1976), 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.

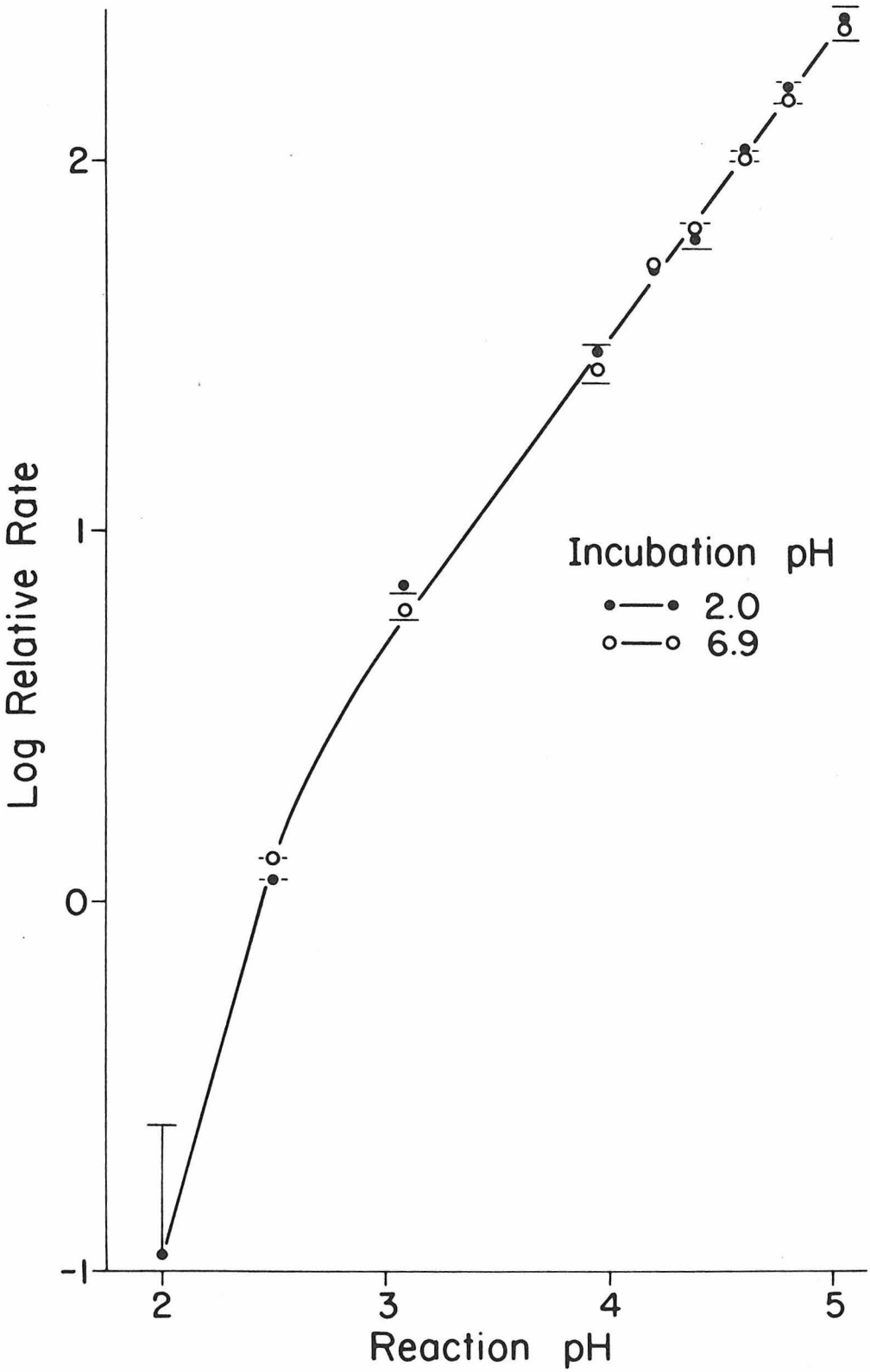
[b] $$\sigma = \left(\frac{\sum_{i=1}^{N} (k_r \text{ (overall)} - k_{r_i})^2}{N - 1} \right)^{\frac{1}{2}}$$

[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 III

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

M. Krieger, R. E. Koeppe II, and R. M. Stroud

Biochemistry--in press, 1976.

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 ion (k_b) (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, 1976).

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 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).

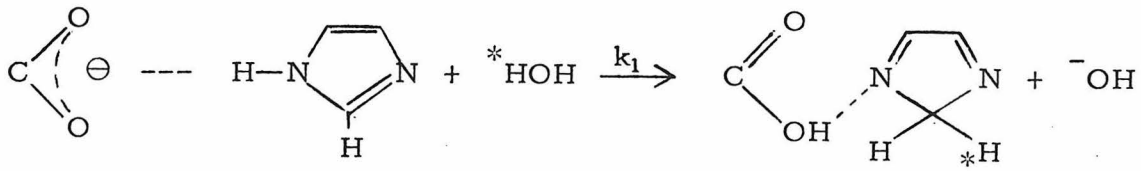
If the side chain of His 57 is involved in a conformational equilibrium between the active "in" position and an inactive "out" position, it might only exchange when it is in the "out" conformation. Exchange into His 57 in the "in" position seems unlikely because of the structure of the active site and 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 (2%) 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,

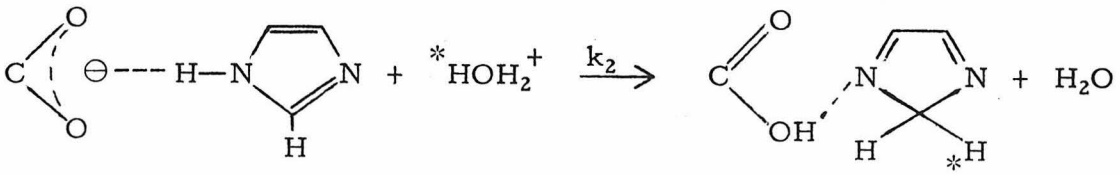
and is an especially serious problem for the case of His 57 of trypsin, which exchanges at a very slow rate.

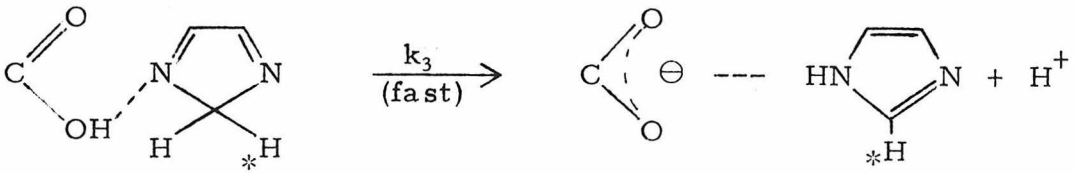
In addition to the ylide mechanism, there are other possible explanations for the tritium exchange into His 57. Of particular interest are mechanisms in which the ionization of a neighboring group affects the rate of tritium exchange of the imidazole. These mechanisms could be of two types: 1) those 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., Biochemistry--in press, 1976.

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 nonlinear 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). $\underline{k}$ was calculated using Eq. 2, a pH of 7.14 and the kinetic parameters in Table III. The specific activity ($\underline{sa}$) was determined from Eq. 1 assuming $\underline{t}$ = 14 days and $\underline{e}$ = 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; △ - 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

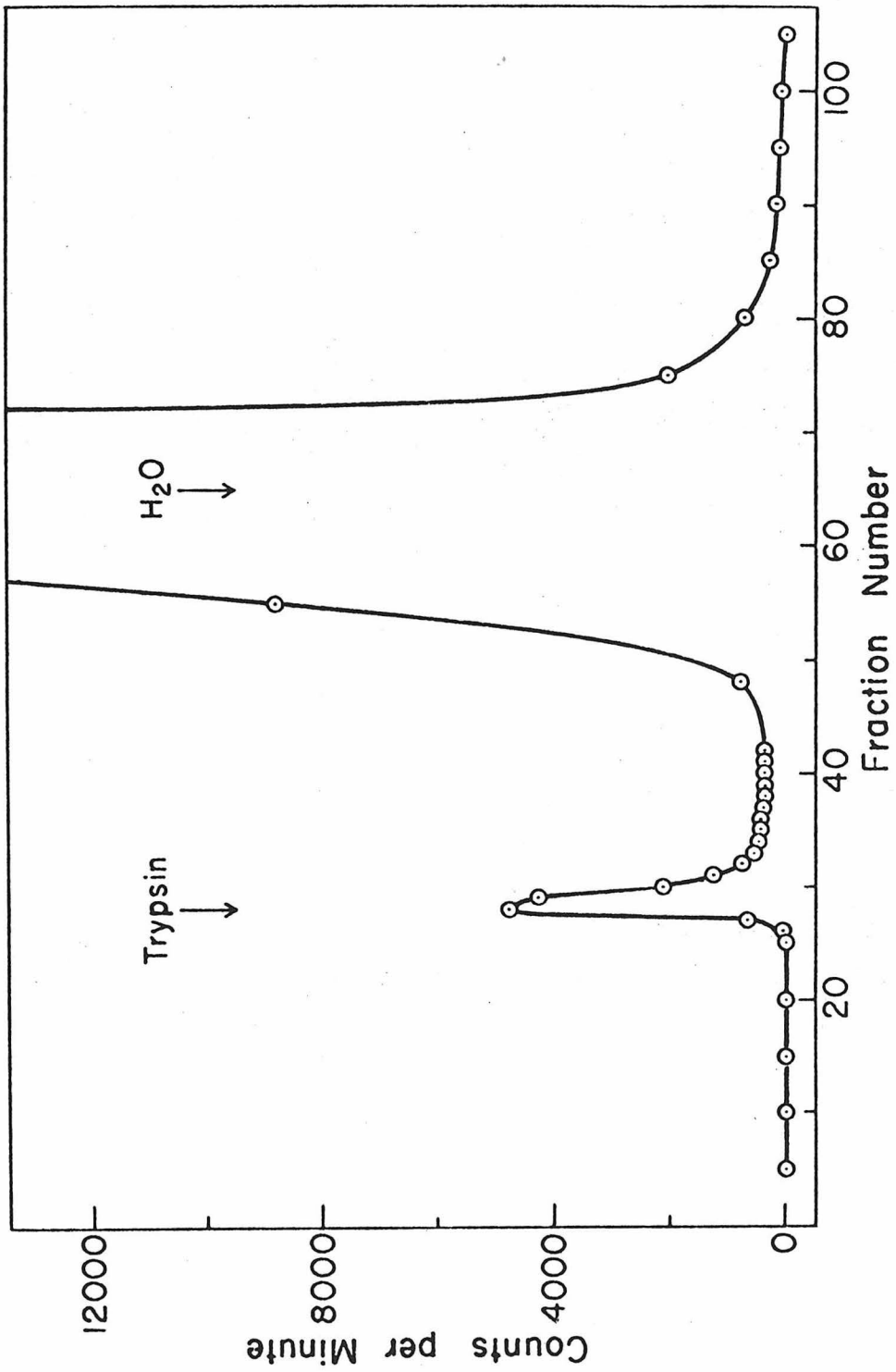
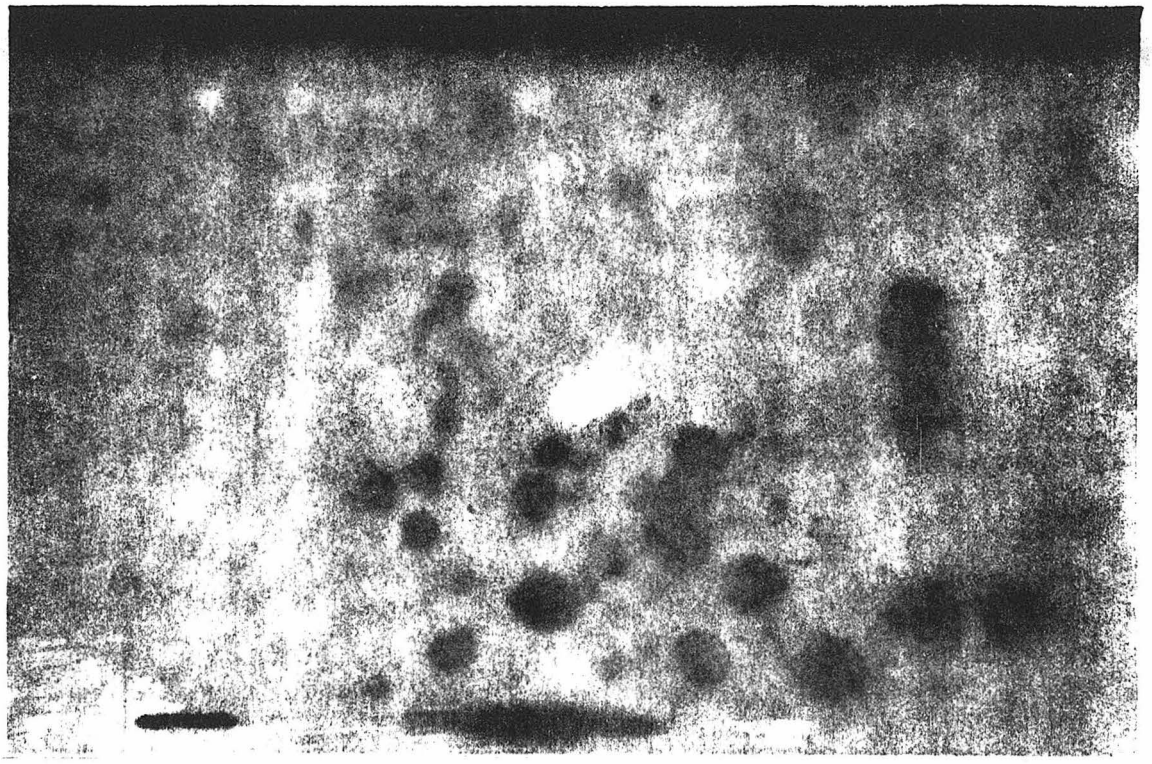

Counts per Minute
12000
8000
4000
0
20
40
60
80
100
Fraction Number
Trypsin
H_2O

Figure 2



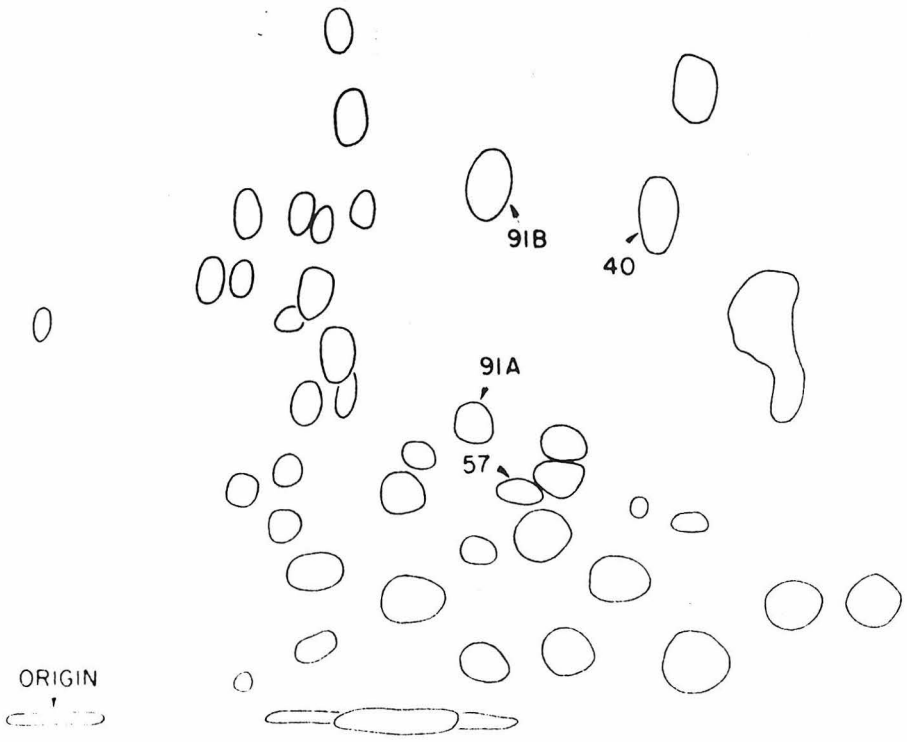


Figure 3

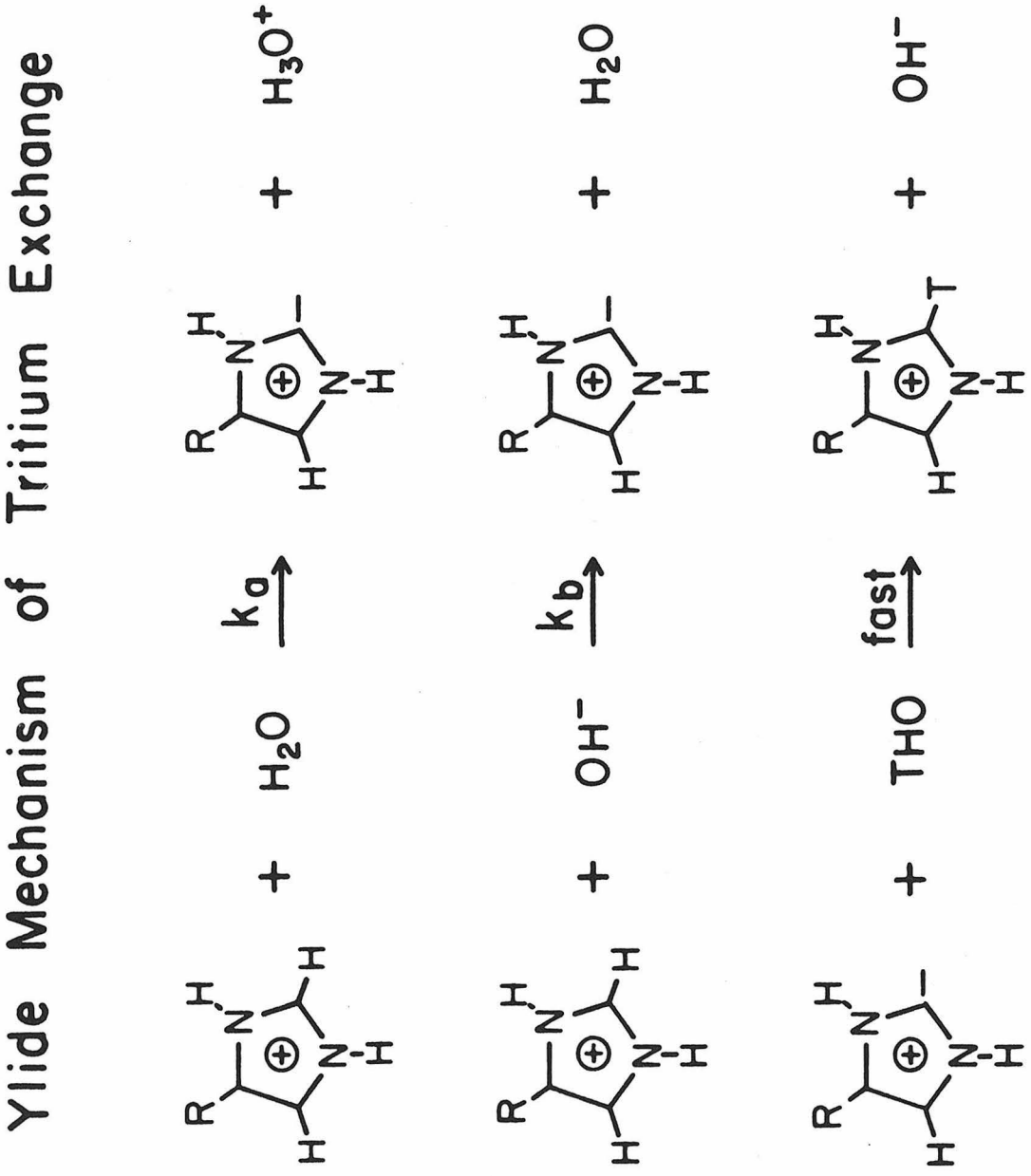

Ylide Mechanism of Tritium Exchange
+ H2O
ka
+ H3O+
+ OH-
kb
+ H2O
+ THO
fast
+ OH-

Figure 4

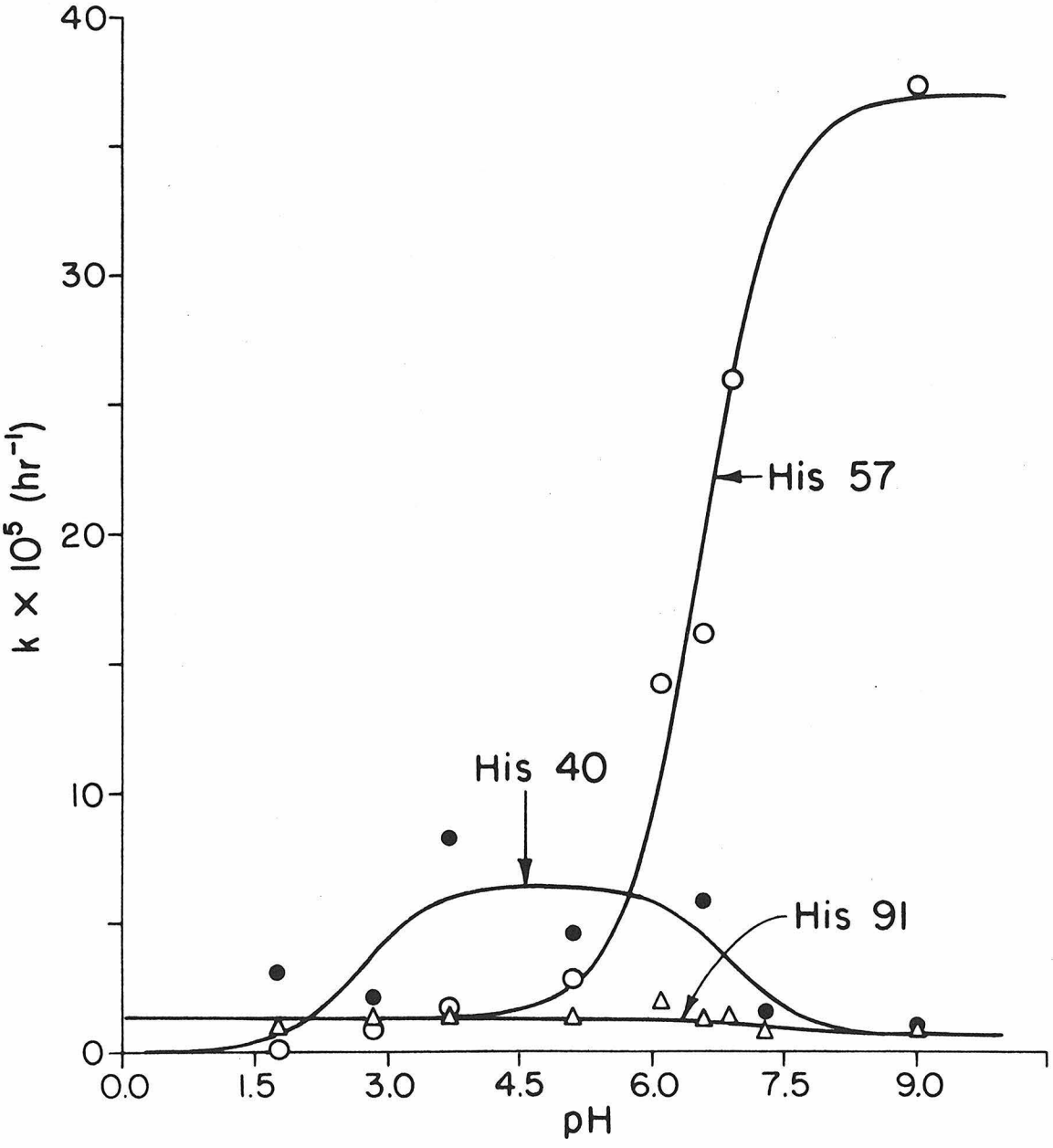


Figure 5

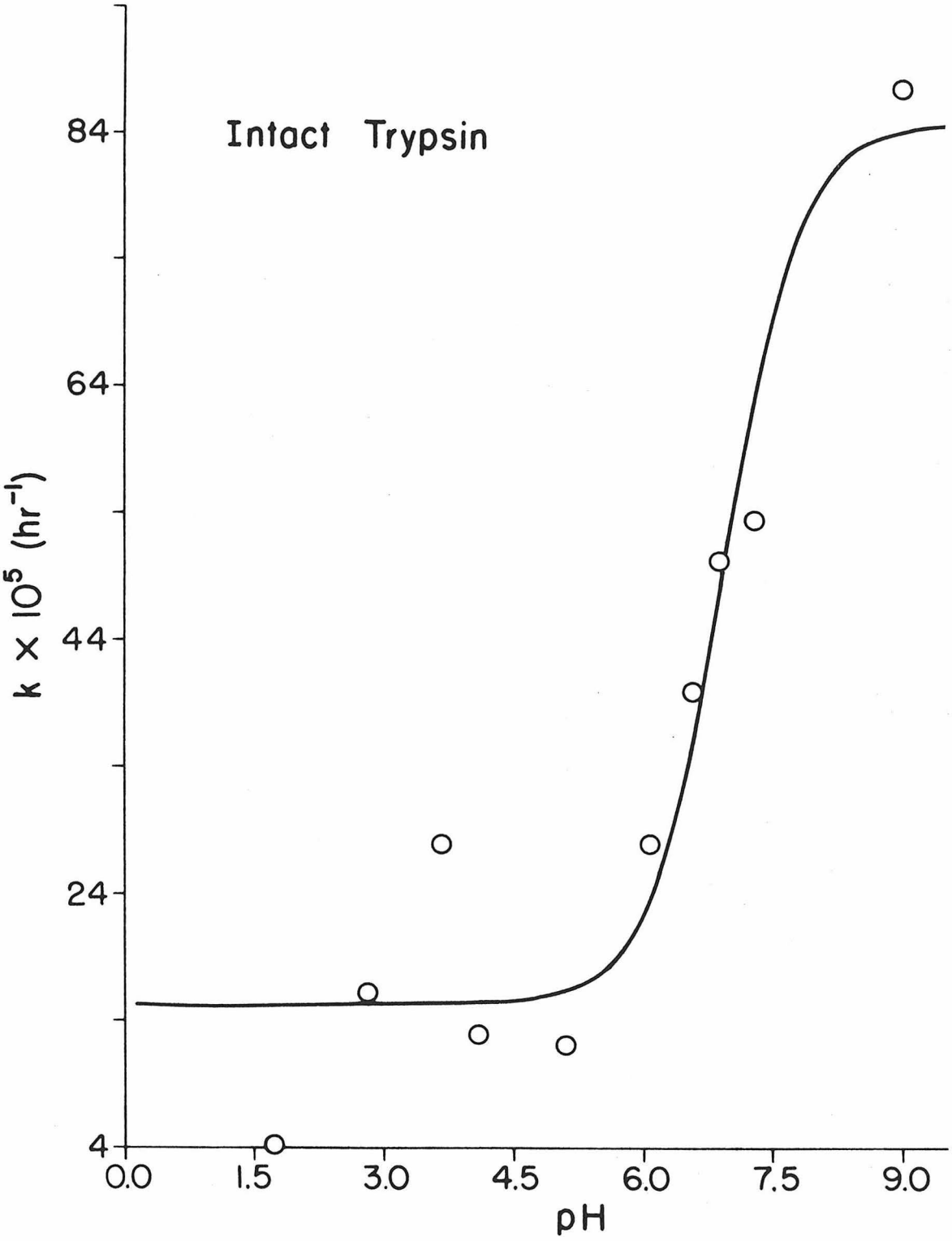


Figure 6

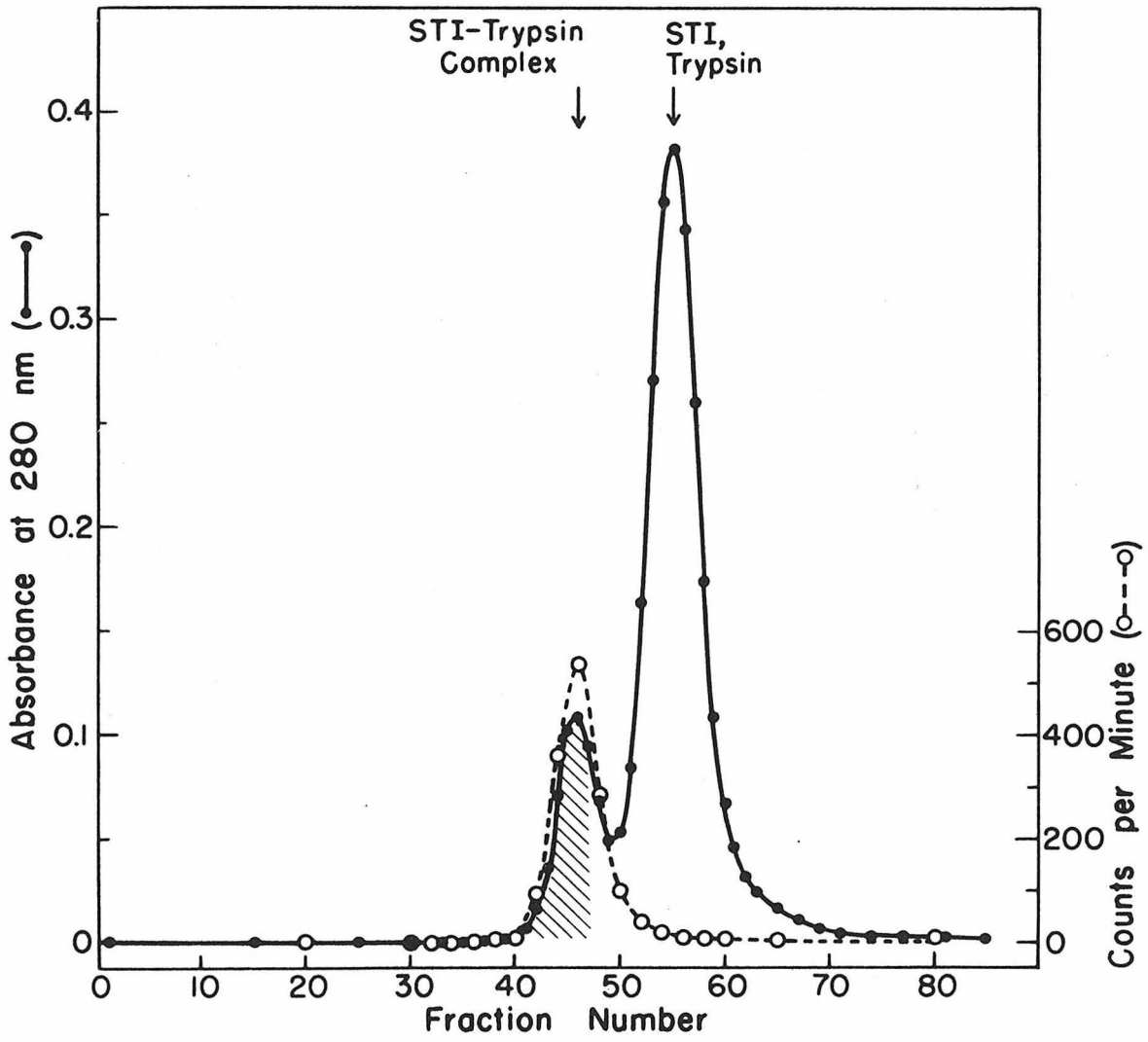

STI-Trypsin Complex
STI, Trypsin
Absorbance at 280 nm (●—●)
Counts per Minute (○--○)
Fraction Number
0
0.1
0.2
0.3
0.4
0
200
400
600
0
10
20
30
40
50
60
70
80

Figure 7

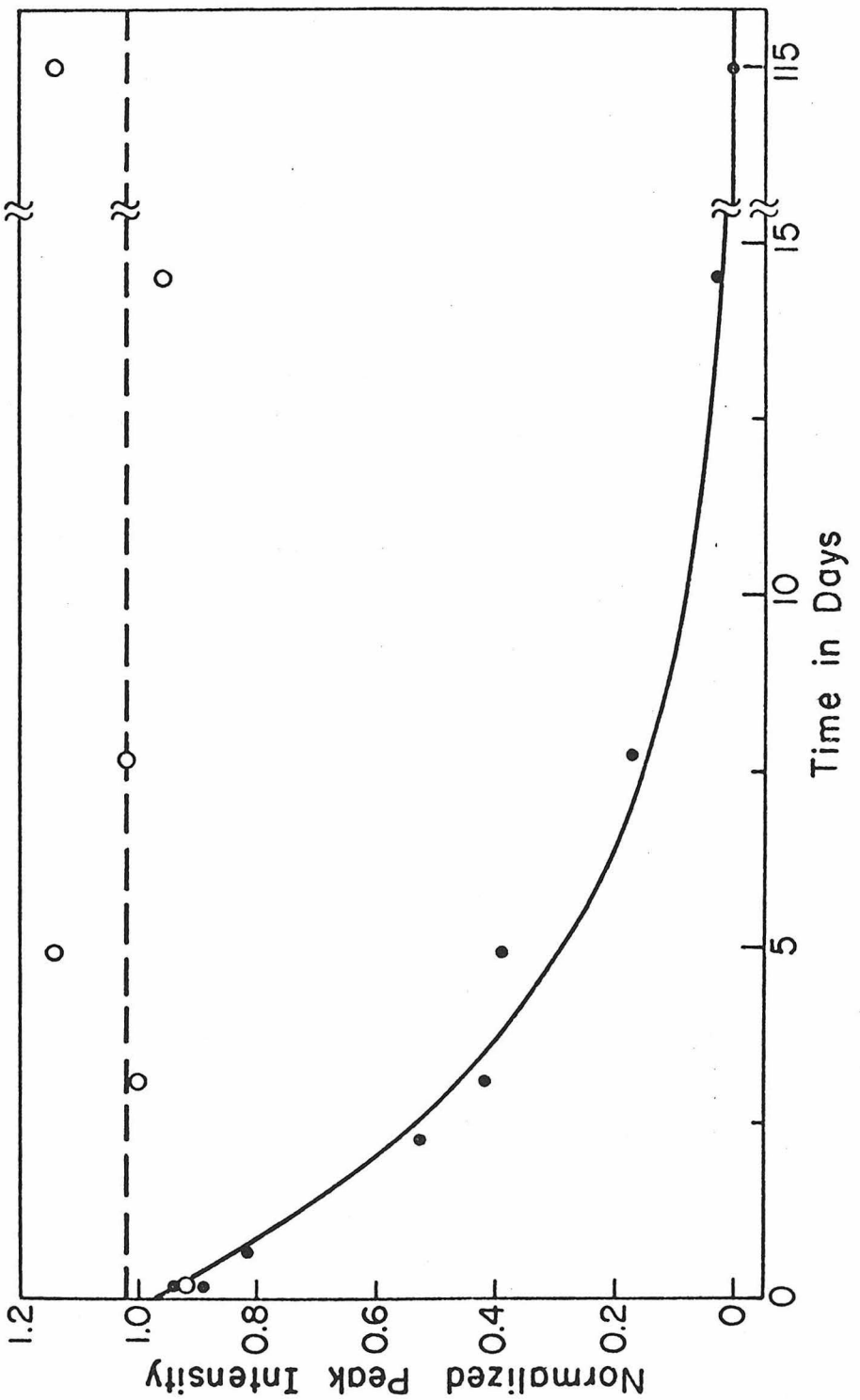
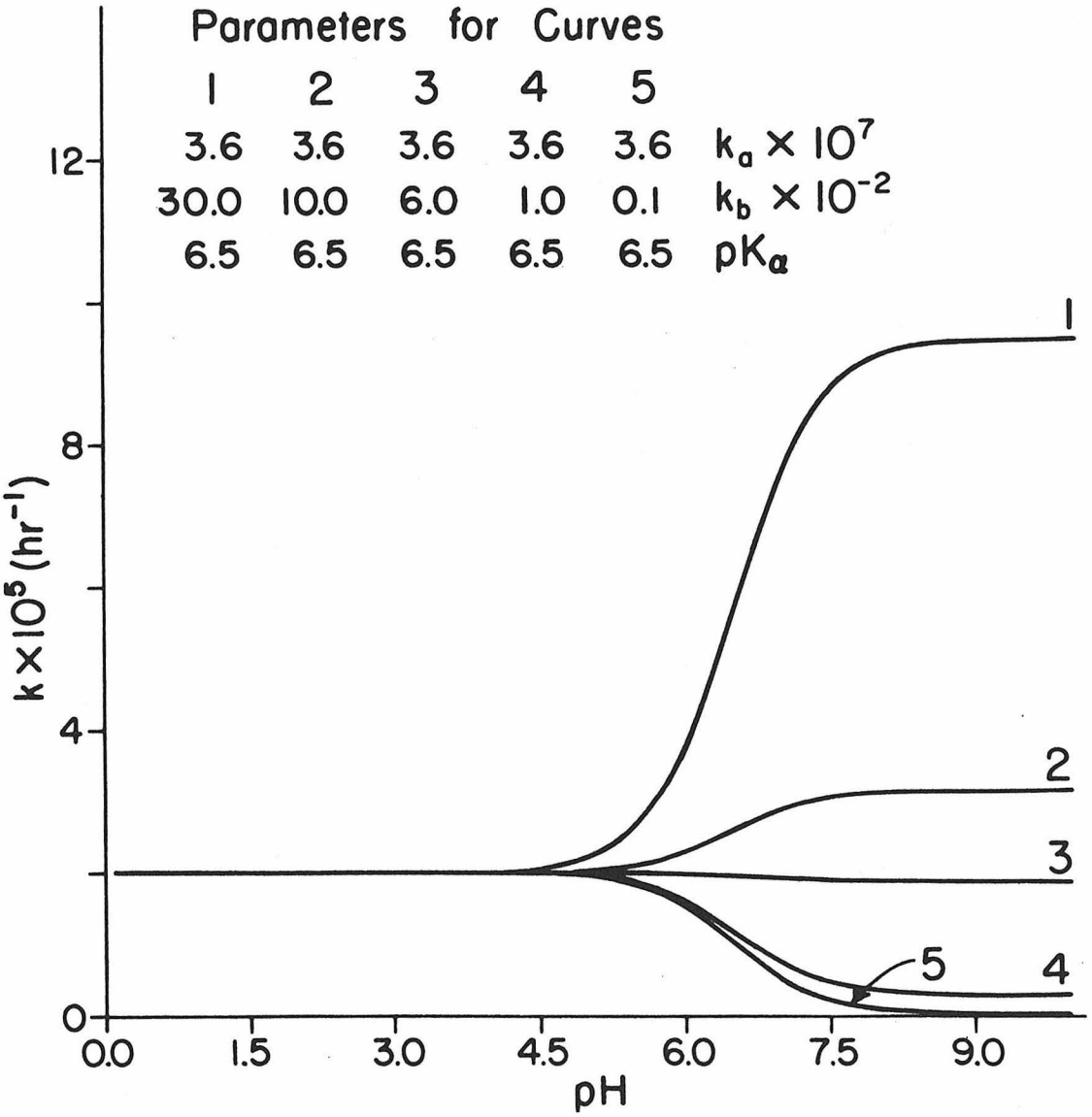

Normalized Peak Intensity
1.2
1.0
0.8
0.6
0.4
0.2
0
0
5
10
15
115
Time in Days

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

APPENDIX IV

Silver Ion Inhibition of Serine Proteases: Crystallographic Study of Silver-Trypsin

J. L. Chambers, G. G. Christoph, M. Krieger, L. Kay, and R. M. Stroud

Biochem. Biophys. Res. Commun. 59, 70-74 (1974).

Vol. 59, No. 1, 1974 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS

SILVER ION INHIBITION OF SERINE PROTEASES:

CRYSTALLOGRAPHIC STUDY OF SILVER-TRYPSIN

J. L. Chambers, G. G. Christoph, M. Krieger,
L. Kay, and R. M. Stroud

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

Received May 3,1974

Summary

Silver ion is a potent inhibitor of trypsin and chymotrypsin, with K_I's of 4×10^{-5} M. and 3×10^{-5}, respectively. A crystallographic study shows that the primary silver ion binding site on trypsin is at the active center between the carboxyl group of Asp 102 and the δ-nitrogen of His 57. This result is correlated with the fact that Ag^+ interferes primarily with the acylation rate constant, k_2, and does not significantly affect the binding constant, K_s. The location of this site explains the potent inhibitory effect of silver (I) ions on trypsin activity: The imidazole ring of His 57 is repositioned 1.8 Å further out into the solvent to accommodate the silver ion, preventing its normal interaction with the hydroxyl group of Ser 195. Consequently, His 57 cannot directly assist the proton transfer in the catalyzed reaction.

Since silver ion binds to the catalytic site in this highly specific manner, silver may be used as a specific probe of the active site of serine proteases.

This communication reports the 2.7 Å resolution structure of the isomorphous silver derivative of bovine trypsin inhibited by diisopropyl-fluorophosphate (DIP). Martinek et al.[1,2] found that silver ion is a potent trypsin (or chymotrypsin) inhibitor with a $K_I = 4 \times 10^{-5}$ M. They concluded, first, that silver ion prevents the acylation of the enzyme while not appreciably interfering with substrate binding. Second, silver ions compete with protons for the binding site and the silver binding depends on a group with an apparent pK_a of 7.1, which they suggested was the imidazole of His 57. In light of more recent evidence, however, the pK_a reflected in those experiments is more likely that of Asp 102[3,4].

EXPERIMENTAL

To prepare the silver (I) derivative crystals of DIP-trypsin[5] were soaked in solutions containing 0.012 M $AgNO_3$ for periods of four to eight days. Three-dimensional, 2.7 Å data sets for both the native DIP-trypsin

and the silver (I) derivative were collected using a Syntex $P\bar{1}$ automated diffractometer. Data reduction and scaling were accomplished using standard techniques.[5]

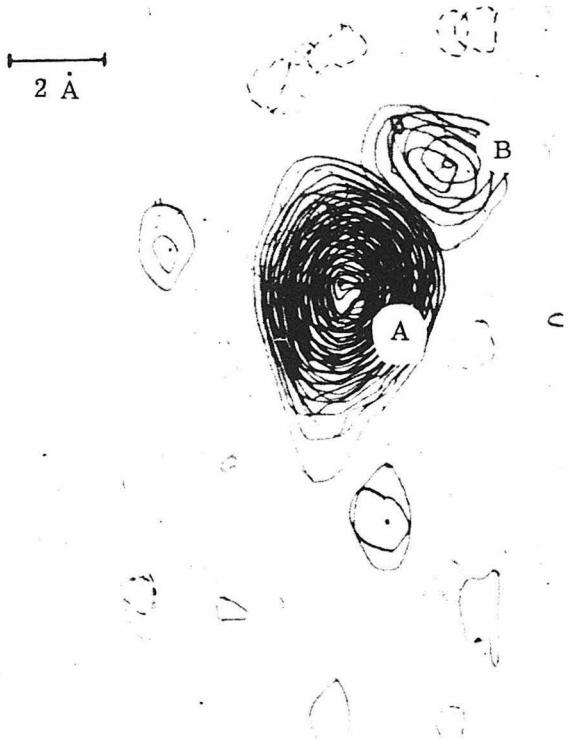
RESULTS AND DISCUSSION

From our data and the phases previously determined for DIP-trypsin,[5] a difference Fourier map was obtained. The region of this map in the area of the active site is shown in Fig. 1. The large peak "A" corresponds to the position of the fully occupied, primary silver ion binding site. In addition, there is a smaller region of positive density, "B", above and behind the primary site. This peak results from a movement of the imidazole ring of the catalytic site residue His 57 into the solvent by approximately 1.8 ±(≤ 0.2) Å. A secondary silver ion binding site of 35% refined occupancy was found elsewhere on the surface of the molecule, in the vicinity of His 40.

Fig. 2 shows an ORTEP[6] representation of the catalytic site in the silver (I) trypsin derivative. The silver ion is coordinated in an approximately linear fashion between the lower carboxyl oxygen of Asp 102 and the δ-nitrogen of His 57. Bond distances are: Oδ1 (Asp 102)-Ag^+, 2.3 ±0.2 Å; Ag^+-Nδ (His 57), 2.3 ±0.2 Å. This suggests that the silver ion is in a two coordinate sigma-bonded complex characteristic of silver (I)[7]. This configuration is also structurally very similar to complexes of silver (I) with free amino acids.[8]

The silver DIP-trypsin structure provides a model for the mechanism of the silver (I) inhibition of trypsin. The distance between the ε-nitrogen of His 57 and the position of the γ-oxygen of Ser 195 found previously for benzamidine-trypsin (where the serine oxygen was hydrogen bonded to the ε N of His 57) is 4.2 ±0.2 Å in the silver derivative. This long distance, coupled with the unfavorable directionality between these two atoms, prevents proton transfer from the hydroxyl group of Ser 195 to the His 57 imidazole. The effect of Ag^+ on k_2, the acylation rate constant, thus gives an indication of the contribution of His 57 and Asp 102 to enhancement of the catalytic rate of serine proteases.

As well as defining the mechanism of the silver ion inhibition, the silver derivative data have been incorporated into the phase refinement



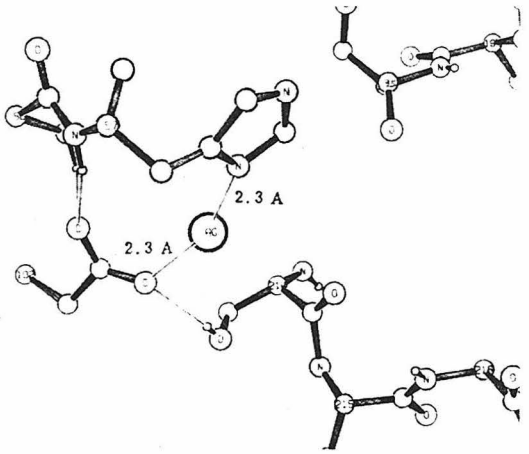


Figure 2. An ORTEP drawing of the catalytic site of silver DIP-trypsin, viewed approximately down the crystallographic b-axis. The DIP group has been omitted for clarity. The γ-oxygen of Ser 195 is in the position found for DIP-trypsin, close to that found in tetrahedral intermediates or in acyl enzymes.

Figure 1. A view parallel to the crystallographic a-axis of the active site region of a difference Fourier map calculated using coefficients $|F_{silver}| - |F_{native}|$. The view shown is a projection through 12 sections, each 1.1 Å apart. Solid contours enclose regions of positive density; the broken lines, regions of negative density. Peak "A" corresponds to the location of the main silver ion binding site. Peak "B" results from the movement of the His 57 imidazole ring. Contours begin at ± 2.5 standard deviations and the contour interval is one standard deviation in the difference electron density.

of the 2.7 Å structure of DIP-trypsin. The overall mean figure of merit of the 2.7 Å phases increased to 0.83 during this refinement. The details of this refinement will be discussed elsewhere.

One of the key problems in the assignment of microscopic pK_a's to residues at the active center of serine proteases lies in the requirement that one should be examining the "native" enzyme rather than a modified derivative of it. Two familiar problems are thus apparent: 1) presence of a covalent label near the active site perturbs the system under study in an unknown fashion; 2) spectroscopic techniques which study the native enzyme often have problems of assignment to particular residues.

Silver ion can be used to assign peaks to the active center residues by perturbation of the native enzyme spectrum, primarily affecting peaks due to Asp 102 and His 57. We are currently applying this technique to assign peaks associated with carboxylic acid groups in the difference infrared titration spectra of native serine proteases.[9]

ACKNOWLEDGEMENT

This work is contribution no. 4850 from the Norman W. Church Laboratory of Chemical Biology of the California Institute of Technology. This work has been carried out with the support of the United States Public Health Service Grants GM-19984 and GM-12121, whose help is gratefully acknowledged. One of us (MK) is the recipient of a Danforth Graduate Fellowship, another (RMS) is the recipient of a National Institutes of Health Career Award (GM-70469), and another (GGC) is a National Institutes of Health Postdoctoral Fellow.

REFERENCES

1. Martinek, K., Savin, Y.V., and Berezin, I.V., Biokhimiya 36, 806-815 (1971).
2. Martinek, K., Vill, K., Strel'tsova, Z.A., and Berezin, I.V., Molekulyarnaya Biol. 3, 554-565 (1969).
3. Krieger, M., Kay, L.M., and Stroud, R.M., J. Mol. Biol. 83, 209-230 (1974).

4. Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. R., and Richards, J. H., Biochemistry 12, 4732-4743 (1973).
5. Stroud, R. M., Kay, L. H., and Dickerson, R. E., J. Mol. Biol. 83, 185-208 (1974).
6. Johnson, C. K., "ORTEP: A Fortran Thermal Ellipsoid Plot Program for Crystal Structure Illustrations", Oak Ridge National Laboratory, Oak Ridge, Tenn., Publ. No. ORNL-3794, revised 1965.
7. Cotton, F. A. and Wilkinson, G. "Advanced Inorganic Chemistry", Interscience, New York, 2nd ed., 1966, p. 1040.
8. Acland, C. B. and Freeman, H. C., J. Chem. Soc. D1971, 1016-1017.
9. Koeppe, R. and Stroud, R. M. --in preparation.

APPENDIX V

Data Collection in Protein Crystallography: Capillary Effects and Background Corrections

M. Krieger, J. L. Chambers, G. G. Christoph, R. M. Stroud and B. L. Trus

Acta Crystallogr. A30, 740-748 (1974).

Reprinted from *Acta Crystallographica*, Vol. A 30, Part 6, November 1974

PRINTED IN DENAMRK

Acta Cryst. (1974). A30, 740

Data Collection in Protein Crystallography: Capillary Effects and Background Corrections*

BY M. KRIEGER†, J. L. CHAMBERS, G. G. CHRISTOPH‡, R. M. STROUD§

Crellin Laboratory of Chemistry

AND B. L. TRUS¶

Norman W. Church Laboratory of Chemical Biology, California Institute of Technology, Pasadena, California 91109, *U.S.A.*

(*Received* 16 *January* 1974; *accepted* 24 *May* 1974)

In protein crystallography, observed diffraction intensities must be corrected for background radiation due to scatter from air and scatter and absorption by capillary, crystal and mother liquor. A systematic study shows that a major contribution to background intensity is air scatter arising from the air intercepted by the direct X-ray beam as 'seen' by the receiving-counter aperture. As a result there is a first-order dependence of background on the 2θ angle. The second-order variations in this function are principally due to absorption of the direct beam or air-scattered radiation by the capillary and to diffraction by the glass in the direct beam. To reduce data collection time and crystal exposure, individual background measurements may be approximated by interpolation from empirical background curves or, alternatively, by collecting background intensities for short times and fitting these data with a multidimensional function. If isotropic interpolation is used, *i.e.*, if background is considered to be a function of 2θ alone, systematic errors of up to about 30% can be introduced into the interpolated backgrounds. Methods of accounting for the anisotropy in the background are derived and shown to reduce this error to 1–2%.

Introduction

Because protein crystals are quite susceptible to radiation damage, crystallographers have been exploring different methods for accurate data collection which minimize the X-ray exposure of the crystal. With automated diffractometers, one widely used method of data collection involves counting the background on one or both sides of each measured reflection. There are several methods of reducing the time of data collection and thereby increasing the number of reflections collected per crystal (*e.g.*, Wyckoff, Tsernoglou, Hanson, Know, Lee & Richards, 1970; Watson, Shotton, Cox & Muirhead, 1970). In one method, the observation of individual backgrounds is omitted and background corrections for the measured intensities are calculated from an empirical curve of background *versus* $\sin\theta/\lambda$ measured for each crystal.

If the backgrounds do not vary significantly during the time of collection, almost all of the crystal exposure time can be devoted to intensity data collection, and empirical background curves can be measured after the intensity data have been collected. Empirical background curves have generally been obtained from background values measured with long counting times for points along one lattice row, and applied using the approximation that the background depends only on the Bragg angle and not on the other setting angles (Matthews, Levine & Argos, 1972; Jensen, 1972; Salemme, Freer, Xuong, Alden & Kraut, 1973; Wyckoff, Doscher, Tsernoglou, Inagami, Johnson, Hardman, Allewell, Kelly & Richards, 1967). This approximation can give rise systematic errors of about 30% between interpolated and observed backgrounds.

Hill & Banaszak (1973) have reported observing an additional, 2θ-invariant φ dependence. Our experiments show that the background radiation can vary systematically with φ and χ in addition to 2θ, and furthermore that such φ and χ dependences are functions of 2θ. These dependences can be accounted for by additional components in a simple interpolative procedure. In contrast to Hill & Banaszak, we conclude that the variation with φ arises primarily from capillary absorption and scattering rather than absorption of background radiation by the protein crystal and its mounting to the capillary. There is substantial improvement in interpolated backgrounds if the φ and χ-dependent variations are accounted for.

A second approach to streamlining data collection involves observing very short backgrounds with each reflection, and fitting, by least squares, a function of the diffractometer setting angles to all of the data. Such a function must include cross terms between χ and 2θ, and φ and 2θ, and can vary in complexity depending upon the range of 2θ and the experimental conditions. By pooling the data in this manner, the resulting backgrounds are more accurate than the individual short measurements from which the background function was constructed.

* Contribution No. 4726. Supported by U.S. Public Health Service Grants GM-19984 and GM-12121.

† Danforth Foundation Fellow.

‡ National Institutes of Health Postdoctoral Fellow.

§ National Institutes of Health Career Development Awardee, U.S. Public Health Service Grant No. GM-70469.

¶ Jane Coffin Childs Memorial Fellow.

It should be noted that systematic variations between real and interpolated values for the background can give rise to systematic errors in phase determination, in ΔF terms used in calculating difference maps, and in scaling. Phases and ΔF's free from such errors are crucial to accurate descriptions of molecular shifts from difference maps (Dickerson, Kopka, Varnum & Weinzierl, 1967; Henderson & Moffat, 1971; Krieger, Kay & Stroud, 1973). Furthermore, with protein crystals whose reflection intensities are small, a reflection which is only twice as intense as the background may be ten standard deviations above background. Therefore, well determined reflection intensities ($\gtrsim 3\sigma$) may be in error by as much as 30% if anisotropy in the background is not accounted for.

As the background intensity is independent of crystal absorption, the backgrounds, estimated or measured, should be subtracted from the observed diffraction intensities before applying the standard absorption and Lp corrections.

In this paper we describe two methods for obtaining background corrections of similar quality to individually measured backgrounds, but at significant savings in the time spent on observing backgrounds. Although it is desirable to minimize the background present (*i.e.*, improve the signal-to-noise characteristics of the experiments), we are primarily concerned here with developing background correction methods whose use would be applicable to a wide variety of existing four-axis diffractometers, without modification of the experimental hardware. As a result of this analysis, and the consideration of the origins of components in the background intensity, experimental methods of reducing the background, and thereby increasing the signal-to-noise ratio, are suggested.

Experimental

Measurements were made on several different diffractometers both with and without a monochromator. Details of each experimental arrangement used are described in Table 1. Each machine will henceforth

Table 1. *Characteristic data for diffractometers used in this study*

In all cases the basic four-circle machines have been rebuilt and modified; in all cases except *A* and *D* modifications have been extensive to the point of total remachining.

	(*A*) Syntex *P*$\bar{1}$	(*B*) GE XRD-490	(*C*) Datex/GE hybrid	(*D*) Hilger Watts*	(*E*) Buerger Supper†	(*F*) S. Samson‡ hybrid
Source						
Target size (mm)	10 × 1	12·5 × 0·8	12·5 × 0·8	10 × 1	8 × 0·4	12·5 × 0·8
Voltage (kV)	40	45	45	46	33	45
Current (mA)	20	18·5	15	16	20	18
Radiation	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$	Mo $K\alpha$
Geometry						
Take-off-angle	6°		3°	3	3	
Monochromator type	Graphite	Graphite	None	None	Graphite	Graphite
Monochromator mounting (ϱ)§	90°	90°				0
Monochromator dispersion	0·3	0·3				0·3 0·4
Source to crystal						
Distance from source to:						
1st aperture(cm) (size in mm)	5·5 (2·0)	4·5 (1·0)	-		8·0 (2·0)	4·0 (1·0)
Monochromator crystal	6·8	5·0	None	None	9·0	5·5
Aperture (size)	19·8 (1·0 1·5)	8·0 (1·0)	4·5 (1·5)	~ 4·0 (1·0)	11·5 (0·6)	12·0 (1·5)
Aperture (size)	21·3 (1·5 2·0)	14·5 (1·8)	14·3 (1·8)	~ 9·5 (1·0)	15·9 (0·6)	20·8 (1·8)
Crystal	27·3	15·5	15·0	23·5	19·4	22·5
Crystal to counter ¶						
Distances from crystal to:						
Aperture (size)	32·5 (2·5)	4·0 (1·0 1·5)	9·2 (1·0)	17·0 (3·0)	6·0 (2·2)	8·0 (2·11)
Aperture (size)	38·5 (1·0 2·0)	12·0 (2·0 2·5)		22·5 (3·0)	11·8 (1·5 2·0)	11·4 (5·04)
Counter	40·5	16·0	22·2	25·0	13·5	21·6
Noise count rate with X-rays off (per 100 s)	35	55	25	38	30	13

* Standard four-circle machine with no monochromator.

† This machine was essentially redesigned by, and rebuilt under the direction of, R. M. Stroud, and uses Datex automation for operation in a φ-step scan mode (two-circle Weissenberg geometry).

‡ Four-circle diffractometer assembled locally under the direction of S. Samson. Electronics and Alloys goniostat, Varian computer, Datex resoverdynes, and Ortec counter chain.

§ ϱ is the angle between the normals to the planes of incidence at the monochromator.

¶ Counter type: all diffractometers use scintillation counter. Hermetically sealed Tl-drifted NaI crystals are individually selected and checked regularly for decay and damage.

be referred to by a letter (*A–F*) which in turn corresponds to one arrangement in Table 1. Each of the figures refers to a specific case, although the discussion considers results obtained from many different experiments. It should be noted that these different arrangements include diffractometers both with and without monochromators, with both of the generally used monochromator settings, with copper and molybdenum $K\alpha$ radiation, and with or without helium-filled pathways to the counter.

Glass capillaries were 0·3–1·0 mm in diameter with walls approximately 0·01 mm thick. Crystals of trypsin, trypsinogen, cytochrome *c* and their derivatives with dimensions ranging from 0·1 × 0·1 × 0·35 mm to 0·4 × 0·4 × 0·7 mm were used in the crystal–capillary experiments. Trypsin crystals grow with acicular habit, cytochrome crystals are rectangular prisms, and trypsinogen crystals grow as approximately equidimensional trigonal bipyramids. Thus crystals of the most commonly encountered morphological forms have been used.

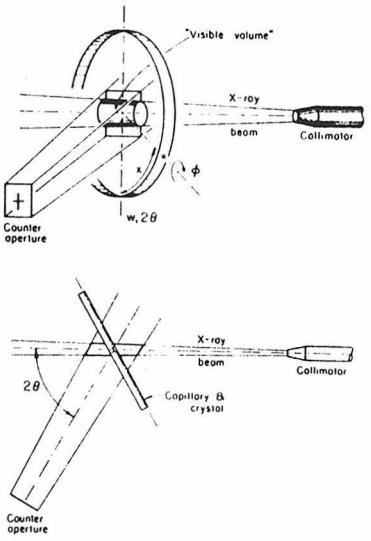


Fig. 1. The usual diffractometer data-collection arrangement. The 'visible volume' is the volume of air irradiated by the direct beam that is seen by the receiving-counter aperture. This volume changes only with 2θ, and governs the 2θ dependence of the air-scattered part of the background. Portions of the 'visible volume' are at times hidden from the X-ray beam and at times from the counter by the interposition of the capillary and crystal. The overall background will be affected by changes in the capillary–crystal setting angles, as well as the amount of capillary glass in the direct beam.

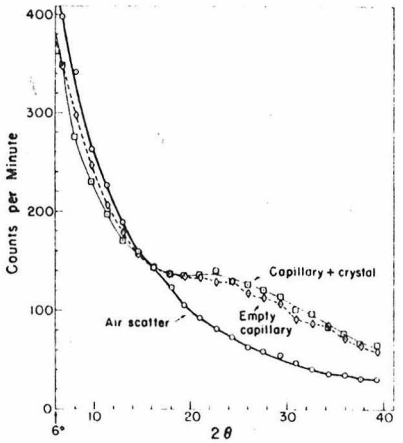


Fig. 2. The background 2θ dependence for: (*a*) air scatter only (○, 5 min/point), (*b*) an empty capillary (◇, 3 min/point), and (*c*) for a trypsinogen crystal, mounted in the same 0·5 mm diameter capillary (□, 3 min/point). All measurements were made at $\chi = 90°$, using diffractometer *A* as described in the text.

Backgrounds were measured under the following conditions:

(1) without crystal or capillary in the beam;

(2) with a well centered capillary (mounted coaxially with φ) in the beam;

(3) with a capillary miscentered, but mounted parallel to the φ axis; and

(4) with a crystal in a capillary such that the crystal, but not the capillary, is centered in the beam (the standard data-collection condition).

Background counting times were usually 3 to 5 min per point for the empirical background curves, and 10 to 120 s for individual backgrounds used for comparison. Reflection intensities were measured in both the ω scan and the Wyckoff step-scan (Wyckoff *et al.*, 1967) modes. Measurements were made over the ranges $0 \le 2\theta \le 46°$; $0 \le \varphi \le 360°$; and $0 \ge \chi \le 360°$. Attenuators were used to obtain data for small 2θ angles down to 0°.

Results

Fig. 1 shows the usual diffractometer data collection arrangement and illustrates the origins of the different variations in the background intensity.

Primary effects: 2θ dependence

Fig. 2 shows the overall background as a function of 2θ: (*a*) with no crystal, no capillary; (*b*) for a well centered capillary; and (*c*) for a capillary containing a trypsinogen crystal (dimensions approximately 0·3 ×

0·25 × 0·35 mm). The intrinsic absorption of a glass capillary was measured and found to be generally less than 25%. The absorptive component due to the glass which affects the curves of Fig. 2 must consequently be less than about 25%. It is much less where 'visible volume' dimensions (see Fig. 1) are greater than the diameter of the capillary – which is often the case. It is therefore clear that scattering by air is a major contributing factor to the background. In the presence of a capillary the background is reduced by absorption by the glass capillary. The background intensity increases at higher 2θ angles where scatter from the glass predominates over its absorption. Variations in the orientation of the capillary lead to φ- and χ-dependent secondary perturbations in the background.

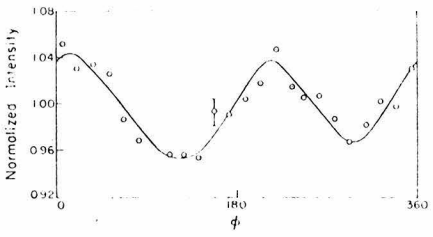


Fig. 3. The background φ dependence observed at $2\theta = 22{\cdot}6°$ and $\chi = 90°$ for a DIP-trypsin crystal. The average intensity is normalized to 1·0 (Diffractometer A).

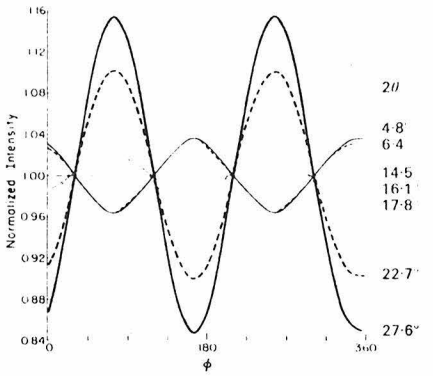


Fig. 4. The φ dependence of the background for an eccentrically mounted empty capillary at various 2θ values ($\chi = 90°$ in all cases). While the amplitude and sign of the variations change with 2θ, the phase is constant. The curves were measured at $2\theta = 27{\cdot}6°$, 22·7°, 17·8°, 16·1°, 14·5°, 6·4°, and 4·8°. The traces, except for $2\theta = 14{\cdot}5°$, represent the best least-squares fit of 24 points measured at 15° intervals to the function: Normalized intensity $= 1{\cdot}0 + a\cos 2(\varphi - \varphi_0)$. The average intensity at each 2θ is normalized to 1·0. The curve for $2\theta = 14{\cdot}5°$ is the best straight line through the data at this 2θ value. Measurements were made on diffractometer A.

Secondary effects: φ dependence

When it is necessary to collect data with a crystal mounted in a capillary whose *diameter* is greater than the radius of the X-ray beam, there can be a significant systematic variation of the background radiation with the diffractometer φ setting. Because there is little difference between the angular distribution of the background of an empty capillary and that of a capillary containing a crystal and its mother liquor, the φ dependence must be almost wholly due to the eccentricity of the capillary about the φ axis.

Fig. 3 shows typical φ dependence of background intensity for a capillary containing a protein crystal, and Fig. 4 shows this dependence for an eccentrically mounted empty capillary at different 2θ values. A well centered capillary shows no φ dependence, while with eccentrically mounted capillaries the background variation is an approximately sinusoidal function of 2φ as the capillary rotates in and out of the beam. At low 2θ values the background, primarily due to air scatter, is modulated by absorption of the eccentric capillary; therefore, the background is greatest when the capillary blocks the direct beam least ($\varphi = 0$ and 180° in Fig. 5). At higher 2θ angles the capillary scatters more radiation than it absorbs, and backgrounds are highest at $\varphi = 90$ and 270° in Fig. 5. At intermediate 2θ values, near the crossing point of Fig. 2, the absorption and scattering of the capillary are nearly equal and the backgrounds are essentially φ-independent. The extent of the background variation with φ will depend on the eccentricity and diameter of the capillary and the value of 2θ. The curves in Fig. 6 depict the background as a function of 2θ measured at two values of φ; one when the low-angle background is a minimum [curve $A_{\varphi_0,\varphi_{90}}(2\theta)$] and the other when the low-angle background is a maximum [curve $B_{\varphi_{90},\varphi_{90}}(2\theta)$]. The magnitude of the differences between the two curves is the amplitude of the φ-dependent variation. There may be circumstances under which these curves would not cross, although we have not observed this. In such a case the amplitude of the φ-dependent variation with 2θ would not change sign.

For a DIP-trypsin crystal in a 0·6 mm capillary and a 1 mm beam, the background at 23° in 2θ varied 11% in intensity with rotation about φ. The average reflection intensity in the shell $22 < 2\theta < 26°$ for this crystal was seven times the background level. For a reflection with this 'average' intensity, using interpolated backgrounds uncorrected for the φ dependence would produce a systematic error up to 1·6% in its net intensity. Weaker reflections, which are still important for heavy-atom refinement and difference maps, are subject to substantially larger errors: a reflection with raw intensity twice background, *although at least* 10 σ (*standard deviations*) *above background*, would be in error by as much as 11%. These errors are systematic and must be corrected for, if accurate phases are to be calculated.

The following scheme was tested for generating interpolated backgrounds corrected for φ dependence:

(1) Measure the φ dependence at a low 2θ value ($\sim 6°$) with a φ scan at $\chi = 90°$.

(2) Measure the 2θ dependence, $A_{\varphi_0,\chi_{90}}(2\theta)$, at the φ value (φ_0) for which the background is a minimum on the φ scan.

(3) Measure the 2θ dependence, $B_{\varphi_1,\chi_{90}}(2\theta)$, at the φ value ($\varphi_1 = \varphi_0 + 90°$) for which the background is a maximum on the φ scan.

(4) Tabulate the difference, $D_\varphi(2\theta) = B_{\varphi_1,\chi_{90}}(2\theta) - A_{\varphi_0,\chi_{90}}(2\theta)$. [Normally, it is convenient to interpolate the values between the observed backgrounds on the $A_{\varphi_0,\chi_{90}}(2\theta)$ and $B_{\varphi_1,\chi_{90}}(2\theta)$ curves at intervals in $\sin(\theta)/\lambda$.]

(5) Calculate the backgrounds, $BG(2\theta,\varphi)$. For background corrections (just as for absorption corrections) the φ setting angle, φ_s must be modified to account for geometrical contributions to rotation with respect to the incident beam from χ and ω.*

$$BG(2\theta,\varphi_s) = A_{\varphi_0,\chi_{90}}(2\theta) + D_\varphi(2\theta)\sin^2(\varphi_c - \varphi_0) \quad (I)$$

where

$$\varphi_c = \varphi_s - \tan^{-1}(\cos\chi \tan\omega)\,. \quad (II)$$

For 1500 reflections observed over a wide range of 2θ, φ, and χ for a DIP-trypsin crystal ($0{\cdot}25 \times 0{\cdot}3 \times 0{\cdot}4$ mm) mounted in a 1 mm capillary on diffractometer *A*, the agreement between observed and interpolated backgrounds typically improved by 40% when φ-dependent anisotropic, rather than isotropic, interpolation was used. This scheme assumes that the φ dependence is sinusoidal† and that the amplitude of the variation is the difference $[D_\varphi(2\theta)]$ between the $A_{\varphi_0,\chi_{90}}(2\theta)$ and $B_{\varphi_1,\chi_{90}}(2\theta)$ curves of Fig. 6. If possible, crystals should be mounted only in capillaries whose diameters are smaller than the radius of the X-ray beam. Under these conditions there is no observable φ dependence.

* Any additional correction for a 2θ component in φ_c (due to rotation with respect to the counter) is much less significant in general where the visible volume (Fig. 1) is larger than the irradiated portion of the capillary. The correction for χ and ω deals with the source of scattering directly, while the 2θ correction deals with only a small component of the scattered radiation. In any case an additional 2θ correction would be asymmetric and complex and is usually unnecessary.

† We have always observed a certain asymmetry in the φ scans which we presume to be due to the difference in position of the eccentrically mounted capillary with respect to the counter at φ and $\varphi + 180°$. In order to account for this asymmetry, we have calculated backgrounds using the expression:

$$BG(2\theta,\varphi) = A_{\varphi_0,\chi_{90}}(2\theta) + D_\varphi(2\theta)\left[\frac{F(\varphi_c) - F(\varphi_0)}{F(\varphi_1) - F(\varphi_0)}\right]$$

where, for example,

$F(\varphi) = a(x - x_0)^2 + b(y - y_0)^2 + c(x - x_0)(y - y_0)$; $x = \sin\varphi$, $y = \cos\varphi$, and x_0, y_0, a, b, and c are refinable parameters. Alternatively, one could use an interpolated φ curve for $F(\varphi)$. However, we have obtained the best results using the sine function in equation (I).

χ dependence

The background χ dependence is analogous to the φ dependence, but a rigorous definition of the problem is more difficult as a result of more complex geometry. However, the χ dependence may be understood qualitatively. As χ varies from 0 to 90°, the amount of glass in the equatorial plane increases approximately as the secant of χ. This, in turn, increases the absorption of air-scattered radiation. The exact nature of this angular background variation with χ depends strongly on the experimental conditions.‡

‡ We have observed that each diffractometer has a characteristic χ 2θ dependence, as shown in Figs. 7 and 8. Subtle differences in the geometry of the diffractometers (*e. g.*, collimator-to-crystal distance) are apparently responsible for the difference in these patterns. The correction scheme outlined below has proven effective for data collected on different diffractometers with different χ 2θ patterns.

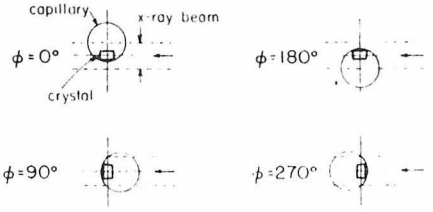


Fig. 5. View parallel to the φ axis of a capillary and crystal, showing the amount of direct-beam radiation they intercept at various φ settings. The capillary diameter is greater than the radius of the X-ray beam.

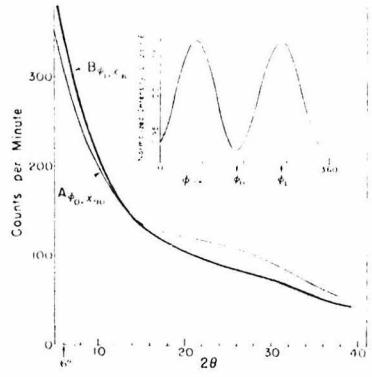


Fig. 6. Curves $A_{\varphi_0,\chi_{90}}$ and $B_{\varphi_1,\chi_{90}}$ the background 2θ dependence observed at two φ settings, 90° apart, for a capillary-mounted DIP-trypsin crystal. Inset: a φ scan for this crystal taken at $2\theta = 6°$, showing the positions of the φ_0 and φ_1 settings (Diffractometer 4).

Figs. 7 and 8 show the 2θ dependence of the background radiation for DIP-trypsin and cytochrome c crystals at $\chi = 0°$ [$C_{\varphi_0,\chi_0}(2\theta)$] and $\chi = 90°$ [$A_{\varphi_0,\chi_{90}}(2\theta)$], and the variation of the χ dependence for several values of 2θ. Just as with the φ dependence, the amplitude of the χ variation depends upon the difference between the backgrounds at the extremes, $\chi = 0$ and $\chi = 90°$. Depending upon the experimental conditions, it is possible to observe a crossover of the $A_{\varphi_0,\chi_{90}}(2\theta)$ and $C_{\varphi_0,\chi_0}(2\theta)$ curves analogous to that for the φ curves. In such cases, the high-angle background increases as χ goes from 0 to 90°, and the shape of the curve can vary from that observed at low 2θ (Fig. 8).

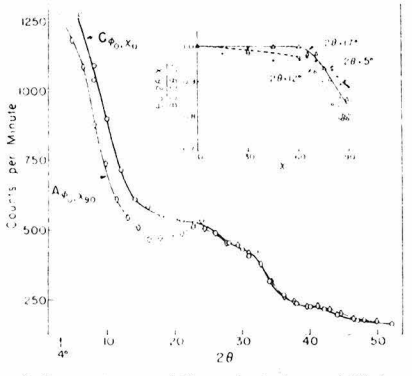


Fig. 7. Curves $A_{\varphi_0,\chi_{90}}$ and C_{φ_0,χ_0}: the background 2θ dependence observed at $\chi = 90$ and $\chi = 0$ on diffractometer A for a DIP-trypsin crystal mounted in a 1 mm capillary. Each data point was counted for 4 min. Inset: the χ dependence observed at several 2θ values. Each point was measured for 5 min.

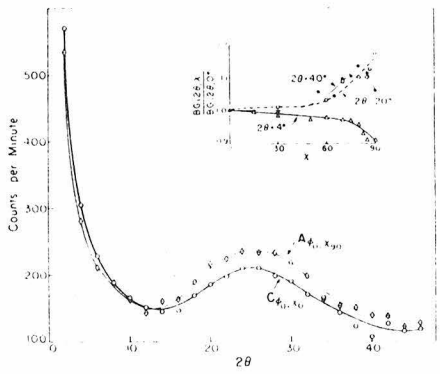


Fig. 8. Curves $A_{\varphi_0,\chi_{90}}$ and C_{φ_0,χ_0}: the background 2θ dependence observed at $\chi = 90$ and $\chi = 0°$ for a cytochrome c crystal using diffractometer B. Inset: background χ dependence at several 2θ values. All points were measured for 4 min.

The χ-dependent background variation can represent a significant fraction of the total background. We have routinely observed 15% differences in backgrounds between $\chi = 0$ and $\chi = 90$, and differences as high as 25% are not uncommon. The following interpolation scheme, analogous to that used for φ, was used to correct for the χ dependence:

(1) Measure the 2θ dependence, $C_{\varphi_0,\chi_0}(2\theta)$, at $\chi = 0$.

(2) Measure the 2θ dependence, $A_{\varphi_0,\chi_{90}}(2\theta)$, at $\chi = 90$.

(3) Tabulate the differences, $D_\chi(2\theta) = C_{\varphi_0,\chi_0}(2\theta) - A_{\varphi_0,\chi_{90}}(2\theta)$.

(4) Measure the χ dependence at a $2\theta = 2\theta_m$ for which the difference, $D_\chi(2\theta)$, is large. If possible, $2\theta_m$ should be representative of the 2θ's in the data set.

(5) Calculate the backgrounds from:

$$\mathrm{BG}(2\theta,\chi) = A_{\varphi_0,\chi_{90}}(2\theta) + D_\chi(2\theta)\left[\frac{F(\chi) - F(90)}{F(0) - F(90)}\right]_{2\theta_m} \quad \text{(III)}$$

where $F(\chi)$ is either an empirical function representing the χ dependence or an interpolated χ curve at $2\theta_m$. Our best results have been obtained with:

$$F(\chi) = a\exp\{-b[(\cos\chi + c)]\} \quad \text{(IV)}$$

where the coefficients a,b,c are determined by a least-squares fit. If the background shown in Fig. 7, for example, is not corrected for χ dependence, *i.e.*, if isotropic interpolated backgrounds (depending only on 2θ) are used, systematic errors of up to 30% of the background can result. χ-dependent (anisotropic) interpolation can be improved upon by subdividing the data into smaller 2θ ranges and applying appropriate χ curves to the data in each range.

When both φ and χ dependences are present, backgrounds may be approximated by combining the φ and χ corrections:

$$\mathrm{BG}(2\theta,\varphi,\chi) = A_{\varphi_0,\chi_{90}}(2\theta) + D_\varphi(2\theta)\sin^2(\varphi_i - \varphi_0) + D_\chi(2\theta)\left[\frac{F(\chi) - F(90)}{F(0°) - F(90°)}\right]_{2\theta_m}. \quad \text{(V)}$$

Non-linear least squares

A different approach to the estimation of backgrounds is the use of non-linear least squares to approximate backgrounds as some function of 2θ, φ, and χ. This function is determined by fitting backgrounds measured for short times in the vicinity of each reflection. The reliability of such a procedure depends on the choice of a well behaved function with a relatively small number of parameters. Its accuracy also depends on the number of data and time spent on measuring the individual backgrounds.

The dependence of background intensity on 2θ may be satisfactorily accounted for by a third-order function of 2θ:

$$\mathrm{BG}(2\theta) = a + bT + cT^2 + dT^3, \quad \text{(VI)}$$

where $T = 2\theta$ and the lower case letters a,b,c, etc., are parameters to be determined. Some improvement can be achieved when the data are separated into low, middle, and high-angle regions in 2θ. Other terms are added to BG(2θ) to account for the variation of BG (2θ) as functions of φ and χ.

In any region of reciprocal space (not necessarily over all 2θ ranges), it is sufficient to approximate the changing magnitude and sign of the φ correction with a third-order polynomial:

$$\Delta B_1(\varphi,2\theta) = \sin^2(\varphi-\varphi_0)(e+fT+gT^2+hT^3) \quad \text{(VII)}$$

where φ_0 is a parameter corresponding to the φ value for minimum or maximum background. If the background does not have a φ dependence, $\Delta B_1(\varphi,2\theta)$ is zero.

Since the magnitude, curvature, and sometimes the sign of the χ correction are also 2θ-dependent, we have tried various combinations of terms of the form $T^i\chi^j$; $i = 1$ to 4; $j = 1$ to 4 to determine which terms are useful to best correct for the χ dependence. An expression of the form

$$\Delta B_2(\chi,2\theta) = iT\chi + jT\chi^2 + kT\chi^3 + lT^2\chi^2 + mT^3\chi \quad \text{(VIII)}$$

provides the best compromise between the number of refined parameters and the overall quality of fit.

The overall expression for least-squares refinement is:

$$\mathrm{BG}(2\theta,\varphi,\chi) = \mathrm{BG}(2\theta) + \Delta B_1(\varphi,2\theta) + \Delta B_2(\chi,2\theta). \quad \text{(IX)}$$

This function can be used in any 2θ range. However, BG(2θ) and $\Delta B_2(\chi,2\theta)$ can often be simplified for middle and high 2θ ranges with essentially no loss of accuracy: for middle or high 2θ, $\mathrm{BG}(2\theta) = a + bT$ suffices; for the middle range, $\Delta B_2(\chi,2\theta) = iT\chi + jT\chi^2 + kT\chi^3 + lT^2\chi^2$; and for the high range $\Delta B_2(\chi,2\theta) = iT\chi + jT\chi^2 + kT\chi^3$.

Discussion

The anisotropic interpolation (AI) and non-linear least-squares (LS) techniques have been used routinely to compare and measure background intensities for crystals of several proteins on different diffractometers. Data for the comparison to be discussed first were taken from a DIP-trypsin crystal using diffractometer *A*. The data shown in Table 2 and Fig. 9 compare the isotropic interpolation (II), AI and LS methods of estimating background intensity with individual backgrounds, BG_{obs}, each measured for 40 s.

Fig. 9 shows the distribution of errors, $\Delta\mathrm{BG}/\sigma_{obs}$, where

$$\Delta\mathrm{BG} = (\mathrm{BG}_{obs} - \mathrm{BG}_{calc}) \text{ and } \sigma_{obs} = \sqrt{\mathrm{BG}_{obs}},$$

for each of the three methods and for a perfectly normal distribution. For the II method the distribution of error is far from normal because of the systematic error introduced by ignoring the anisotropy in the background. The error distribution is skewed and distorted regardless of the setting angles used for observing the 2θ curve (see Fig. 10).

The error distributions for the AI and LS methods, on the other hand, are very nearly Gaussian,* indi-

Table 2. *Error analysis of background-approximation techniques*

Estimated backgrounds compared to backgrounds measured for 40 s per point.*

Method	$\langle\Delta\mathrm{BG}^2\rangle^{1/2}$†	s‡	R_{BG} (%)§	GOF¶	R_I (%)**
II	61·7	55·3	10·0	2·22	1·76
AI	41·2	30·8	6·0	1·38	0·88
LS	32·3	17·1	5·3	1·19	0·87

* Results are for 1956 reflections measured on diffractometer (*A*) as described in the text. The average observed background was 501 counts per min, and the rms σ (BG_{obs}) was 27·4 counts per min. The comparisons are all tabulated in counts per min.

† $\Delta\mathrm{BG} = \mathrm{BG}_{obs} - \mathrm{BG}_{calc}$.

‡ s is the error component due to errors arising from the II, AI or LS methods, $s^2 = \langle\Delta\mathrm{BG}^2\rangle - \langle\sigma^2_{obs}\rangle$.

§ $R_{\mathrm{BG}} = \sum|\Delta\mathrm{BG}| / \sum \mathrm{BG}_{obs}$.

¶ $\mathrm{GOF} = \left[\frac{\sum(\Delta\mathrm{BG}/\sigma_{obs})^2}{1956 \text{ pts.}}\right]^{1/2}$, the 'goodness-of-fit'.

** $R_I = \sum|I_{\mathrm{BG},obs} - I_{\mathrm{BG},calc}| / \sum|I_{\mathrm{BG},obs}|$.

* We have shown this in two ways: least-squares fitting a Gaussian to the error frequency curves in Fig. 9, and by direct comparison to perfectly normal distributions by means of normal probability plots (Abrahams & Keve, 1968).

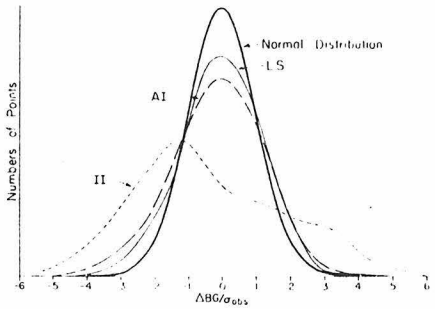


Fig. 9. The distribution of errors between individually observed (40 s) backgrounds and backgrounds calculated using the three techniques described in the text. The statistics plotted are the numbers of points within intervals of $\Delta\mathrm{BG}/\sigma_{obs}$, where $\Delta\mathrm{BG} = \mathrm{BG}_{obs} - \mathrm{BG}_{calc}$ and $\sigma_{obs} = \sqrt{\mathrm{BG}_{obs}}$. The total area under each curve is equivalent to 1956 data points. The curves are frequency distributions for the errors of the II, or isotropic interpolated backgrounds, the AI, or anisotropic interpolated backgrounds, and the LS, or least-squares backgrounds. The heavy curve represents an ideal Gaussian distribution of error for the same number of points. (DIP-trypsin crystal on diffractometer *A*).

cating that the systematic errors resulting from the II method have been largely eliminated. The slightly increased breadth of the AI and LS error distributions relative to the perfectly normal distribution is a consequence of the errors introduced by these methods. We can estimate the extent of these errors, s, using the approximation:

$$s^2 \simeq \langle \Delta BG^2 \rangle - \langle \sigma^2_{obs} \rangle \qquad (X)$$

where $\langle \Delta BG^2 \rangle$ is the mean squared ΔBG and $\langle \sigma^2_{obs} \rangle$ is the mean squared error in the individually observed backgrounds. For the data illustrated in Fig. 9 the estimated average error for backgrounds calculated using the AI technique, with empirical background curves observed for 4 min per point, is roughly equivalent to that which would be obtained by measuring each background for 32 s. For LS, s is equivalent to error expected for backgrounds measured for 91 s each. The value of s for AI depends upon the time spent measuring each point for the interpolated curves, while that for LS depends on the counting times for individually measured backgrounds. In this example, individual 40 s backgrounds were used to calculate the LS function. In another case results were essentially identical when 10 s backgrounds were used, and this method is clearly most useful when still shorter times are used.

We have compared the effects of the II, AI and LS methods on the structure factors by calculating the R values (R_F) between the data sets obtained using observed and calculated backgrounds (see Table 1). The R_I value for the II corrected data is approximately twice as large as that calculated for the AI and LS corrected data. Applying absorption corrections to the scan intensities *before* subtracting II backgrounds, as recommended by Hill & Banaszak, increases the R_F value to 1·95%. Clearly the AI and LS methods are substantial improvements over the II method.

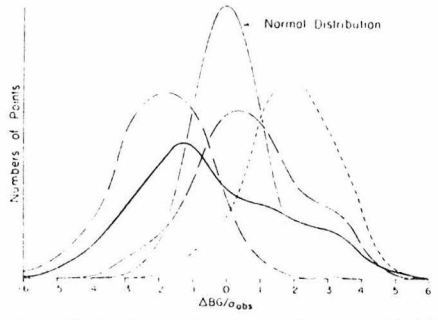


Fig. 10. The distribution of errors when different empirical 2θ curves are used in the II method. The frequency distributions are for backgrounds calculated using: a 2θ curve measured at $\varphi = \varphi_0$, $\chi = 0°$ (– · –), a 2θ curve measured at $\varphi = \varphi_1$, χ 90 (– – –), a 2θ curve measured at $\varphi = \varphi_0$, $\chi = 90°$ (——), and a 2θ curve calculated from the φ_1, χ_{90} and φ_0, χ_{90} curves, representing a 2θ curve which would be obtained if $\varphi \sim \frac{1}{2}(\varphi_0 + \varphi_1)$ (– – –). The light curve (– – –) represents an ideal Gaussian distribution of error. As in Fig. 9, the area under each curve represents 1956 data points distributed over the range $4\theta - 2\theta - 4\theta$.

Conclusion

The most significant conclusions to be drawn from these studies of background intensity are:

(1) The principal 2θ-dependent component of background intensity arises from X-rays scattered by the volume of air illuminated by the direct beam and 'seen' by the receiving-counter aperture. Modulation of this air scatter by capillary absorption and scatter gives rise to φ and χ-dependent variations in the background.

(2) The background intensity is essentially uncorrelated with the crystal absorption. Under certain experimental conditions the two phenomena may appear to be correlated; for example, where a flat crystal rests on the side of the capillary such that its long axis lies parallel to the φ axis of the diffractometer.

(3) As a consequence of (2), the background intensity should always be subtracted from the scan intensity before the absorption corrections are applied, contrary to the suggestion of Hill & Banaszak (1973).

(4) The anisotropy of background scatter as a function of 2θ, φ, and χ can be satisfactorily accounted for by using a simple interpolative procedure, or by fitting a suitable function to many backgrounds measured for short times near each reflection.

(5) The use of an isotropic background interpolation is usually unsatisfactory for accurate data reduction. This method can introduce systematic errors of up to 30% in the estimated background intensities. Errors of this sort lead to systematic errors in phases and increase the noise levels in difference maps.

Both the AI and LS methods substantially reduce the systematic errors inherent in the II method. As a result, the differences between observed backgrounds and estimated backgrounds closely follow a normal distribution. Both methods represent significant savings in time and are improvements over the II method. The overall accuracy of either method is limited by the extent and duration of the background sampling, and is generally commensurate with that obtained for individual backgrounds. When deciding between the two methods, one must choose between the greater savings in time provided by the AI method, or the somewhat better accuracy of the LS technique at the cost of increased crystal exposure and decay.

Our results identify the principal sources of background intensity. Consequently, there are several experimental steps which can be taken to reduce background intensity: (1) Capillaries should ideally be chosen to be smaller in diameter than the radius of the incident X-ray beam. This renders φ dependence insignificant, diminishes χ dependence of the background and reduces the high-angle background due to capillary

scatter. (2) Helium-filled tubes are often used to reduce absorption by air in the input or output beam directions. However, as the main problem is generally one of peak-to-background ratio, these measures only slightly affect this ratio, and in any case do not reduce background intensity significantly. It would seem highly advantageous to fill the 'visible volume' with helium, so affecting the background intensity directly. This is obviously difficult to do as it implies either a helium-filled enclosure over the entire diffractometer, or a helium-filled chamber mounted around the φ axis which surrounds the capillary and crystal completely. There are obvious mechanical difficulties in building such a device. First, it must be almost X-ray transparent over the angular ranges used. Second, it must be moderately well sealed to minimize leakage of helium if it encloses the φ drive shaft bearing. A mylar cylinder with solid supports mounted onto the top of the goniometer head would seem to be a good compromise allowing for rigid support of the top and bottom in the X-ray shadow.

(3) Any means of restricting the visible volume will reduce the background in almost direct proportion to the volume change. This can be achieved by placing the final restricting aperture and the scatter cap on the input collimator as close to the crystal as possible. Similarly, there should be a defining aperture as close to the crystal as possible in the crystal counter pathway, and a second one close to the counter.

We thank Drs A. Kossiakoff and R. Swanson for generously supplying background data, Mr R. Almassy for helpful assistance, and Dr R. E. Marsh and Mr J. Greif for valuable discussions. We are grateful to Dr Sten Samson for allowing us to collect data on diffractometer *F*, a machine designed in the most part by himself, and for his supervision in the redesign of diffractometer *B*. We also recognize his constant attention in the redesign and improvements made to diffractometer *C*.

References

ABRAHAMS, S. C. & KEVE, E. T. (1971). *Acta Cryst.* A**27**, 157–165.

DICKERSON, R. E., KOPKA, M. L., VARNUM, J. & WEINZIERL, J. E. (1967). *Acta Cryst.* **23**, 511–522.

HENDERSON, R. & MOFFATT, J. K. (1971). *Acta Cryst.* B**27**, 1414–1420.

HILL, E. J. & BANASZAK, L. J. (1973). *Acta Cryst.* B**29**, 372.

JENSEN, L. (1972). Personal communication.

KRIEGER, M., KAY, L. M. & STROUD, R. M. (1974). *J. Mol. Biol.* **83**, 209–230.

MATHEWS, F. S., LEVINE, M. & ARGOS, P. (1972). *J. Mol. Biol.* **64**, 449–464.

SALEMME, F. R., FREER, S. T., XUONG, Ng. H., ALDEN, R. & KRAUT, J. (1973). *J. Biol. Chem.* **248**, 3910–3921.

WATSON, H. C., SHOTTON, D. M., COX, J. M. & MUIRHEAD, H. (1970). *Nature, Lond.* **225**, 806–811.

WYCKOFF, H. W., DOSCHER, M., TSERNOGLOU, D., INAGAMI, T., JOHNSON, L. N., HARDMAN, K. D., ALLEWELL, N. M., KELLY, D. M. & RICHARDS, F. M. (1967). *J. Mol. Biol.* **27**, 563–578.

WYCKOFF, H. W., TSERNOGLOU, D., HANSON, A. W., KNOX, J. R., LEE, B. & RICHARDS, F. M. (1970). *J. Biol. Chem.* **245**, 305–328.

APPENDIX VI

Data Collection in Protein Crystallography: Experimental Methods for Reducing Background Radiation

M. Krieger and R.M. Stroud

Acta Crystallogr. A32--in press, 1976.

Data Collection in Protein Crystallography: Experimental Methods for Reducing Background Radiation*

BY

MONTY KRIEGER** and ROBERT M. STROUD***

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

---

*Contribution No. 5226. Supported by U.S. Public Health Service Grant No. GM-19984.

**Danforth Foundation Predoctoral Fellow.

***National Institutes of Health Career Development Awardee, U.S. Public Health Service Grant No. GM-70469.

Abstract

The primary source of background intensity when measuring diffraction intensities by diffractometry is air scatter from the volume of air irradiated by the direct x-ray beam which scatters incident radiation into the photon detection system. In protein crystallography where crystals are generally enclosed in glass capillaries, scattering of the direct beam from the glass is the second most significant source of background intensity. Glass scattering leads to a broad diffuse ring centered at spacings of 3.0 Å ($2\theta = 20° - 32°$), whereas air scattering continuously diminishes in intensity between $2\theta = 0°$ and 45° (Krieger, Chambers, Christoph, Stroud, and Trus, 1974). Compared to these sources of background radiation, other sources are usually insignificant.

Two experimental methods are described which reduce the air scattered component of backgrounds by as much as a factor of 100. Use of these methods increased the peak to background ratio by a factor of 3-10 times in the case of typical diffractometry of protein crystals. This can be translated into time saving during data collection, and into an increase in the range of measurable intensities.

Introduction

In x-ray crystallography, the speed and accuracy of data collection is ultimately limited by the signal-to-noise ratio. Rapid data collection is especially important in protein crystallography because the useful crystal exposure time is limited by crystal decay. If the signal-to-noise ratio can be maximized, more and better data can be collected from a single protein crystal.

Our first investigation (Krieger et al., 1974) of background intensity in protein crystallography served to identify the primary sources: first, background radiation is produced by scatter from the volume (the "visible volume") of air irradiated by the direct incident beam, and seen by the photon detector. This effect generally overwhelms all other sources in the range of 2θ between 0° and 15° (i.e., within the ∞ - 6 Å resolution range). A second source is scatter from the glass capillary in which protein crystals are usually mounted. Capillary scatter becomes most significant in the range of 2θ between 20° and 32°, and generally overwhelms any contribution from solvent around the protein crystal in the capillary, or from inelastically scattered photons from the protein crystal.

Based upon this understanding of the sources of background intensity, two accurate interpolative methods for obtaining

approximate background intensities were devised. Both methods allowed for significant savings of crystal x-ray exposure time, and were capable of producing background estimates of the same quality as measured backgrounds. The conclusion that air scatter was a major contributor to the background led to proposed ways of reducing or eliminating the air scattered component (Krieger et al., 1974).

Since any reduction of the background intensity is reflected directly in the peak*-to-background ratio of the measurements of diffraction maxima, the gain in reducing background scatter is significant--particularly for small crystals. If, for example, the background is reduced tenfold, then equivalent intensity data may be collected from ten times smaller crystals, although the measuring time is ten times longer. This is important when crystal size is limiting.

The first of two experimental techniques for reducing background is a simple precaution which should be taken to reduce the "visible volume" of air which may scatter x-rays into the detector. In the second method, air in the visible volume is replaced by helium contained in a small mylar-walled cylinder around the

---

* "Peak" is defined to be the intensity of a diffraction maximum after correction for background.

crystal. This device may also be conveniently used to control the temperature of the crystal.

Experimental

Measurements were made on a Syntex $P\bar{1}$ diffractometer using copper Kα radiation selected by reflection from a graphite monochromator. The location of restricting apertures and the running conditions of the diffractometer are summarized in Table 1.

A controlled-environment chamber was assembled and mounted on top of the goniometer head. The chamber was constructed in two parts, which are illustrated in Fig. 1. The first component, the base plate A (illustrated in section in Fig. 1), is rigidly mounted on top of a eucentric goniometer head. A crystal can be mounted in the usual manner, with the mounting pin locked rigidly into the center of the base plate. The second part, B, which fits on top of the base plate, carries a cylindrical mylar window and provides an environmental seal. Gas flow through the chamber is achieved through ports, P, mounted into the base plate. The temperature inside the chamber is measured using a thermistor, T, mounted into the top of the device, opposite the φ drive shaft. In order to prevent condensation onto the outside of the mylar windows, a laminar flow of dried

air can be maintained on the outside of the mylar window by leakage of air through holes in the copper tube, C. The laminar flow is in the direction away from the goniometer head. Using this device, the environment around the crystal may be carefully controlled and monitored. Leakage from the cell is low so that a closed circuit of cooled helium gas may be recycled through the cell. A slightly greater-than-air pressure is maintained in the helium system to ensure that helium is not displaced by air.

Background intensity measurements were made both with and without the mylar chamber; with and without a capillary, or a capillary and crystal; and with air or helium in the chamber. The effects of varying the size of counter aperture 1, located 9.0 cm from the crystal and 31.5 cm from the counter, were assessed.

Crystals of di-isopropyl phosphoryl-trypsin (DIPT) with dimensions 0.5 x 0.15 x 0.1 mm mounted in a 0.7 mm diameter capillary, or 0.15 x 0.15 x 0.5 mm mounted in a 0.2 mm capillary, were used in the crystal-capillary experiments. Background counting times varied from 15 seconds to 10 minutes per point. Reflection intensities were measured in the ω-scan mode. All data were collected at $\chi = 0.0°$ and over the 2θ range $2.53° - 44.93°$.

Results

The diffractometer is normally equipped with only one aperture located on the detector arm immediately in front of the detector. A second (variable-sized) aperture system was fixed 9.0 cm from the crystal on the detector arm, and can be seen in Fig. 2. This aperture limits the visible volume of air in the direct beam which can radiate into the detector.

The effect of including this aperture is illustrated in Fig. 3. The upper curve represents the background intensities measured as a function of 2θ for a trypsin crystal mounted in a capillary tube in the normal way. Aperture 1 is present, but is at its maximum opening of 4.0 mm across. The lower curve shows the effect of closing down the aperture to about 2.5 mm across. This represents an approximately 60% reduction in the visible volume. As a result, the background intensity at both high and low 2θ angles is diminished as predicted (although the effect is clearly greatest at low angles). The dotted line shows the effect of removing the crystal from the capillary, and confirms our previous observation that background intensity is essentially independent of the crystal under normal circumstances. If there were no aperture 1, the background intensity would be even higher than that represented by the upper curve.

Therefore, the addition of a second aperture on the counter arm as close as possible to the crystal significantly improves the data. This would not be true if most of the background intensity was generated by capillary and crystal. The width of aperture 1 must be large enough so it will not interfere with the diffracted beam.

In order to effectively reduce the high angle background due primarily to scatter from glass, the amount of glass in the direct beam must be reduced. The differences in the background between crystals mounted in 0.7 mm and 0.2 mm diameter capillaries are shown in Fig. 4. As expected, the backgrounds are similar at low 2θ angles, and the smaller capillary scatters less at higher angles.

In the second series of studies, the effect of substituting helium for air in the visible volume around the capillary was assessed using the environmental chamber. Fig. 5 shows the effect of replacing the air in the visible volume with helium. In the absence of a capillary or crystal, there is an approximately 26-fold reduction in the background at $2\theta = 12°$ which diminishes to a 13-fold reduction at $2\theta = 17.7°$ (corresponding to interplanar spacings of 5 Å). As far out as 37° in 2θ, the air-only background is still three times greater than the background with the helium-filled chamber. The backgrounds measured without the chamber show that the mylar wall is essentially x-ray transparent. With a crystal and capillary

in the beam, the effect of the helium-filled volume is still substantial, although capillary scattering dominates the background at high 2θ angles as expected. The helium had no effect on the magnitude of the integrated reflection intensities after correcting for the background.

The overall effect of adding the helium chamber, using a small capillary and minimizing the width of aperture 1, is depicted in Fig. 6. In concert, these precautions reduce the background (and therefore increase the peak-to-background ratio) by a factor of between three and ten throughout the 2θ range, $2\theta = 0°$ to $2\theta = 45°$.

Discussion

The atomic absorption coefficients of nitrogen and oxygen for CuKα radiation are 17.5 and 30.5, respectively (International Tables of X-ray Crystallography, 1968). The absorption coefficient for helium is .255; thus, helium absorbs less, so scatters less radiation. For helium to effectively reduce backgrounds, it must replace the air that is contributing significantly to the background. The common practice of reducing air absorption of incident or diffracted beam intensity by using helium-filled tubes which enclose the beams, only slightly affects the peak-to-background ratio since the source of

background intensity is not affected in any way. The main value of a helium pathway placed between the crystal and the detector is that it reduces air absorption of the diffracted beam. A 25 cm helium pathway typically increases peak intensities by about 20%. Data collection time can be reduced in direct proportion to the increase in diffracted intensity without sacrificing accuracy. By using a gas flow chamber around the crystal, helium replaces air in the visible volume and thus directly reduces the background. For the case illustrated in Fig. 5, the air scattered component of the background was reduced 26-fold at $2\theta = 12°$. The use of smaller capillaries or of restricting apertures close to the crystal also reduces the background because these precautions directly affect the content or the size of the visible volume. In concert, these techniques can produce a three- to tenfold reduction in the background intensity over the entire 2θ range normally used for protein crystallography (out to a resolution of 2.0 Å).

When backgrounds are reduced by three to ten times, the peak-to-background ratio increases by the same factor. While changes of this magnitude have essentially no impact on the accuracy or data acquisition times for intense reflections, they can affect the collection of weak diffraction data.

The principal advantage of using background reduction techniques is that the minimum peak intensity required for statistically significant measurements is lowered as the background intensity is diminished. If the background is reduced by a factor of f, a small crystal can provide data equivalent to that from a crystal f-times larger, although the data collection time would be f-times longer. Alternatively, a decrease in the background will permit a reduction in the data collection time for weak reflections without loss of accuracy. Background reduction methods may also be useful in extending the resolutions of structures when the diffraction intensities rapidly become small compared to the background; for example, tRNA crystals (Kim, Quigley, Suddath, McPherson, Sneden, Kim, Weinzierl, Blattmann, and Rich, 1972).

In addition to its role as a helium chamber for reducing backgrounds, the mylar gas flow cell provides an economical method of cooling crystals during data collection. The gas flowing through the cell can be cooled by passage through copper heat exchange coils immersed in a cooling bath. Several other ways of cooling crystals have been described (Coppens, Ross, Blessing, Cooper, Larsen, Leipoldt, and Rees, 1974; Marsh and Petsko, 1973; Cucra, Singman, Lovell, and Low, 1970; Streib and Lipscomb, 1962; and Woodley, Hine, and Richards, 1971). One difference between the gas flow cell

and cooling methods which rely on an open gas flow over the capillary and crystal is that there is essentially no condensation of atmospheric water on the moving parts of the diffractometer when the flow cell is used because the cooled gas does not flow over the instrument.

We wish to thank Mr. Steven Spencer for help in the design and assembly of the gas flow cell.

References

Coppens, P., Ross, F.K., Blessing, R.H., Cooper, W.F., Larsen, F.K., Leipoldt, J.G., and Rees, B. (1974). J. Appl. Crystallogr. 7, 315-319.

Cucra, P., Singman, L., Lovell, F.M., and Low, B.W. (1970). Acta. Crystallo. B26, 1756-1760.

International Tables for X-ray Crystallography, Vol. III (C.H. Macgillaury, D.G. Rieck, eds.), p. 166. Kynoch Press, Birmingham, England, 1968.

Kim, S.H., Quigley, G., Suddath, F.L., McPherson, A., Sneden, D., Kim, J.J., Weinzierl, J., Blattmann, P., and Rich, A. (1972). Proc. Nat. Acad. Sci. USA 69, 3746-3750.

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

Marsh, D. J. and Petsko, G. A. (1973). J. Appl. Crystallogr. 6, 76-80.

Streib, W. E. and Lipscomb, W. N. (1962). Proc. Nat. Acad. Sci. USA 48, 911-916.

Woodley, M. J A., Hine, R., and Richards, J. P. G. (1971). J. Appl. Crystallogr. 4, 9-12.

Table 1. Characteristic data for the diffractometer used in this study.

Model	Syntex $P\bar{1}$
Source	
Target size (mm)	10 x 1
Voltage (kV)	40
Current (mA)	20
Radiation	Cu Kα
Geometry	
Take-off-angle	6°
Monochromator type	Graphite
Monochromator mounting (ρ)*	90°
Monochromator dispersion	0.3°
Source to crystal	
Distance from source to:	
1st aperture(cm) (size in mm)	5.5 (2.0)
Monochromator crystal	6.8
Aperture (size)	19.8 (1.0-1.5)
Aperture (size)	21.3 (1.5-2.0)
Crystal	27.3
Crystal to counter	
Distances from crystal to:	
Aperture 1 (size)	9.0 (0-4)
Aperture 2 (size)†	35.0-40.0 (5)
Counter	40.5
Counter	Tl-drifted NaI
Noise count rate with x-rays off (per 100s)	35

* ρ is the angle between the normals to the planes of incidence at the monochromator.

† Aperture 2 is a 5 mm diameter tube which is 5 cm long.

Fig. 1. Gas-flow chamber for the control of crystal environment. The base plate A viewed in section is attached to a goniometer head via a central pin and contains a coaxial mounting slot for the crystal mounting pin normally attached to a capillary tube. Gas flows through the cell via ports, P. A laminar flow of dried air can be maintained around the cell by leakage from the tube, C. The body of the cell, B, is constructed from aluminum tubing. The outer wall is 0.001 inch thick mylar. A thermistor, T, is mounted in the top of the cell for temperature measurement. The chamber dimensions determine the maximum 2θ angle at which data can be collected. For this device data can be collected to $2\theta_{max} = 45°$. The aluminum connection between top and bottom of the cell, D, is adjusted so that it lies on the opposite side of the φ drive shaft to the quadrants of the reciprocal lattice to be measured.

Fig. 2. The gas-flow chamber is shown mounted on the diffractometer. The variable aperture 1 is seen on the left, mounted on the detector arm.

Fig. 3. The effects of changing the width of counter aperture 1 (see Table 1) on the 2θ dependence of background intensity.

Backgrounds were observed with air in the mylar chamber and with a DIPT crystal in a 0.7 mm capillary. The data were measured using a wide setting of counter aperture 1 (4 mm, 15 seconds per point, O), and using a smaller aperture 1 size (~2.5 mm, 5 minutes per point, □). Background intensity was also measured with the same capillary, but without the crystal in the beam (~2.5 mm aperture, 50 seconds per point, ◇). The vertical axis is corrected to counts per minute.

Fig. 4. The effects of capillary size on the 2θ dependence of the backgrounds. Backgrounds were observed with air in the mylar chamber, and with DIPT crystals mounted in a large (0.7 mm diameter, 5 minutes per point, O), or a small (0.2 mm diameter, 1 minute per point, □) glass capillary. The counter aperture width was ~2.5 mm.

Fig. 5. The 2θ dependence of backgrounds with air or helium in the mylar chamber. a) Backgrounds observed without a capillary or crystal in the beam. All points were measured for 8.33 minutes each either with air (O) or helium (●) in the chamber, or without the chamber (◇). b) Backgrounds observed with a DIPT crystal in a 0.7 mm diameter capillary in

the beam. All points were measured for 15 seconds each, either with air (△) or helium (▲) in the chamber. The counter aperture 1 width was 4.0 mm.

Fig. 6. The cumulative effects of the background reduction procedures on the 2θ dependence of the background. a) Backgrounds observed without a crystal or capillary in the beam were measured for 8.33 minutes per point with a 4.0 mm aperture and air in the gas-flow chamber (O) or for 5 minutes per point with a ~2.5 mm aperture and helium in the chamber (●). b) Backgrounds observed with a DIPT-crystal in the beam were measured with the crystal mounted in a large capillary (0.7 mm diameter) with a large aperture (4.0 mm) and air in the chamber (△ - 15 seconds per point) or in a small capillary (0.2 mm diameter) with a smaller aperture (~2.5 mm) and helium in the chamber (▲ - 5 minutes per point).

Figure 1

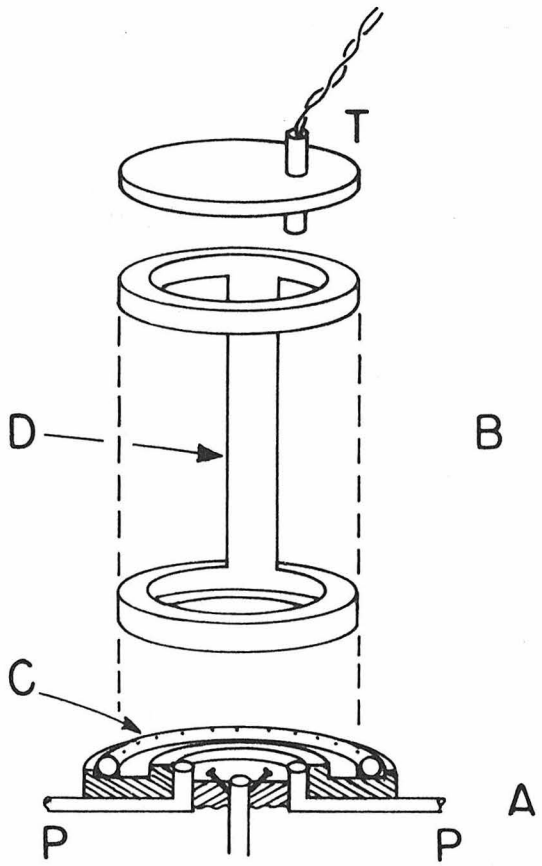


Figure 2

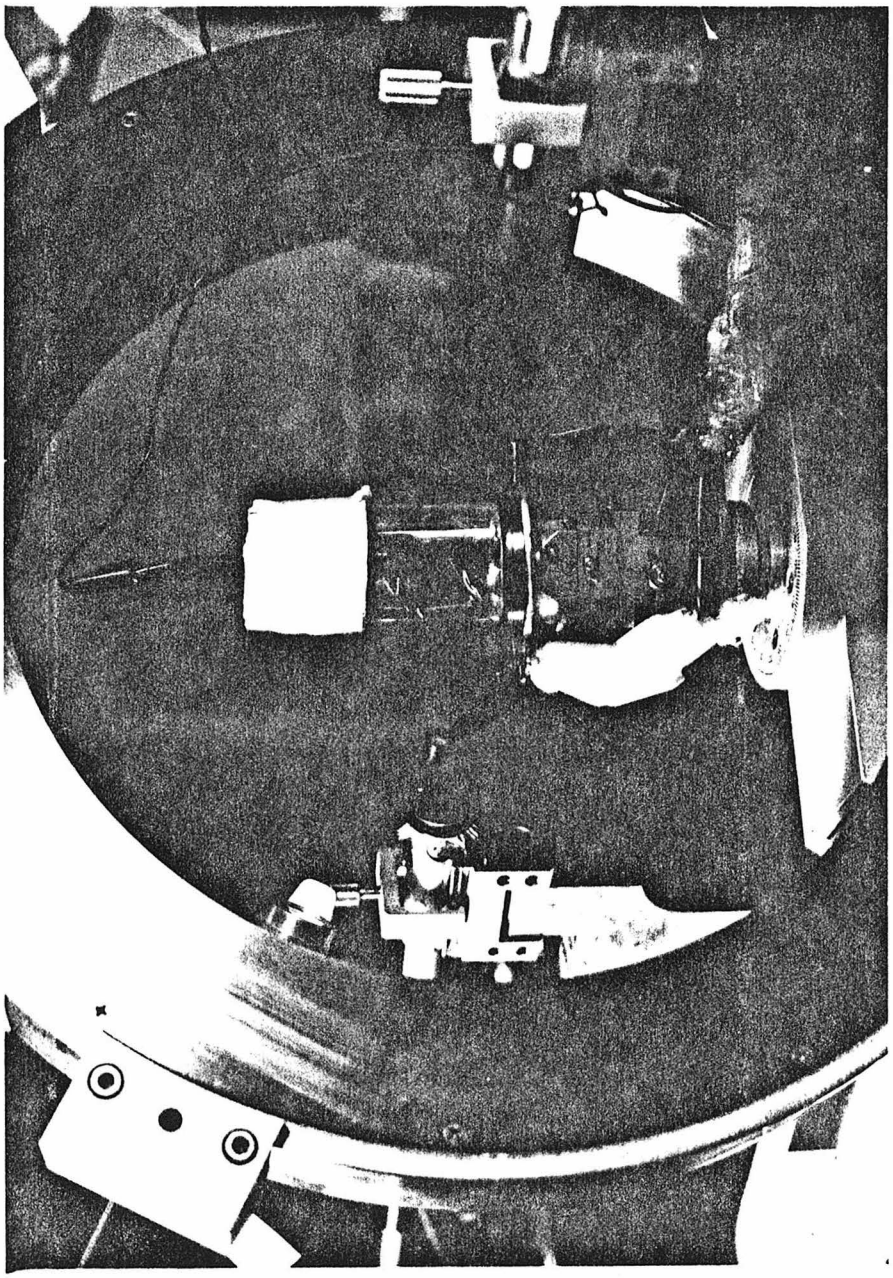


Figure 3

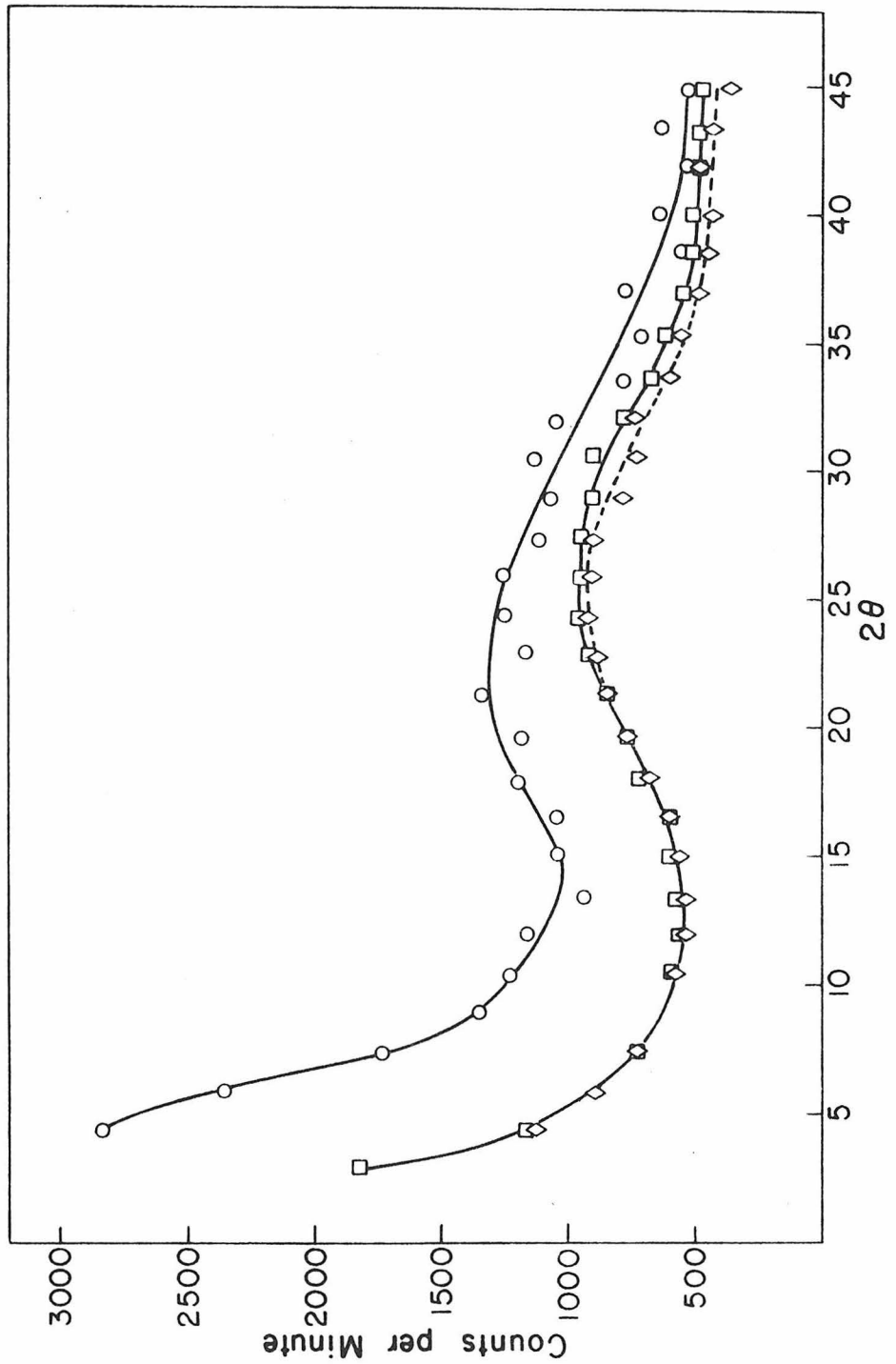

Counts per Minute
3000
2500
2000
1500
1000
500
5
10
15
20
25
30
35
40
45
2θ

Figure 4

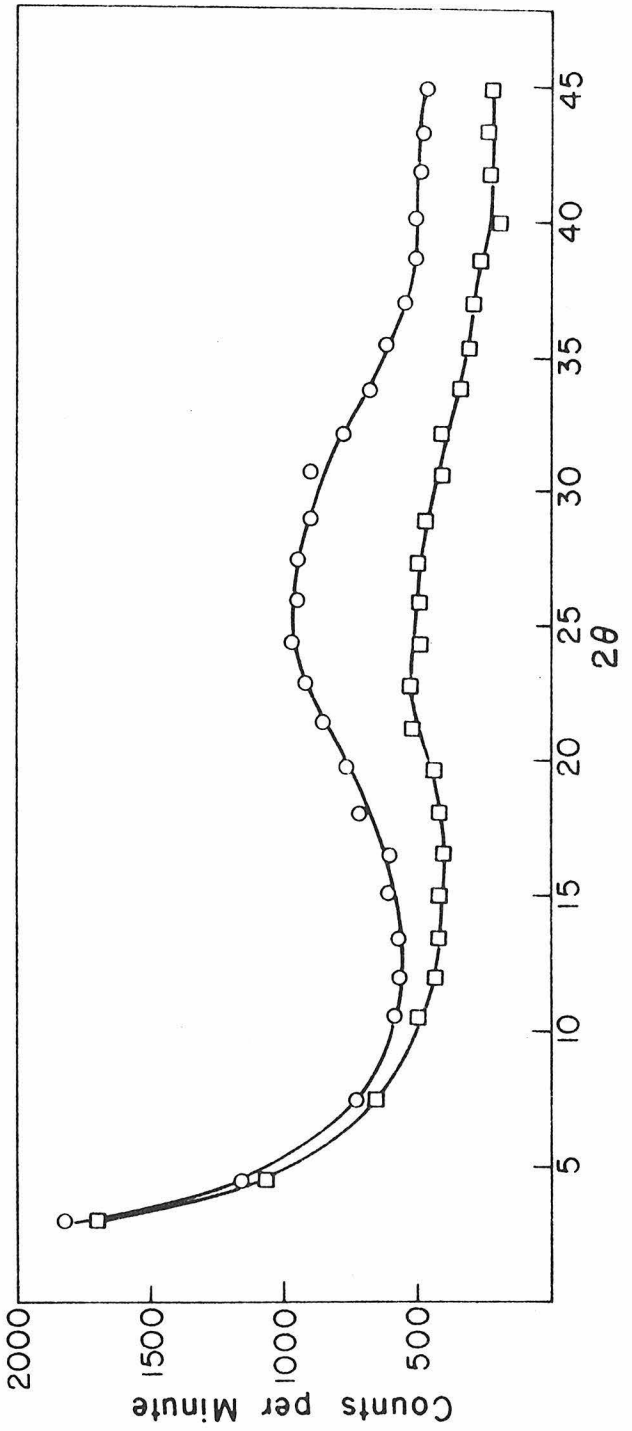


Figure 5

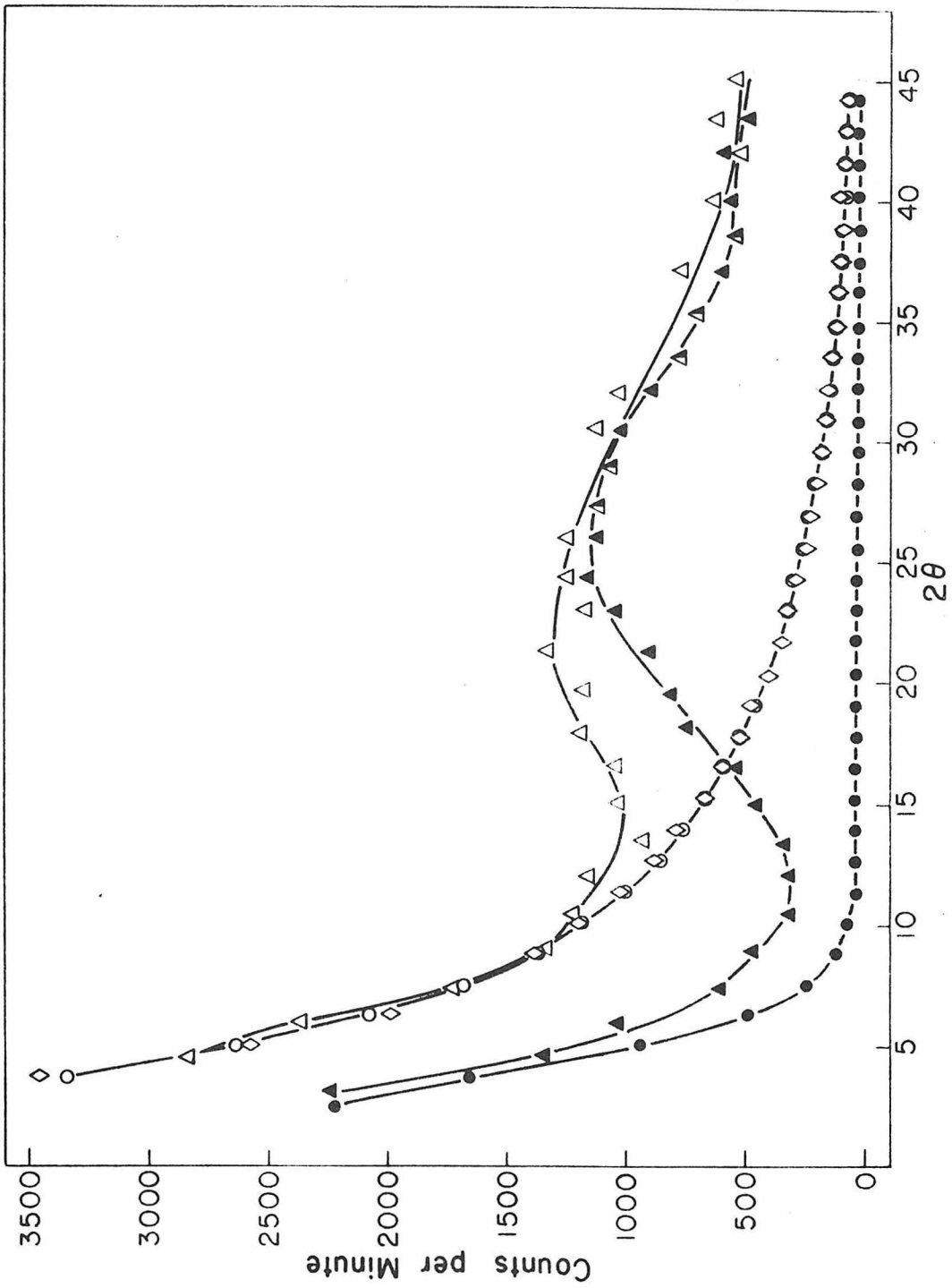
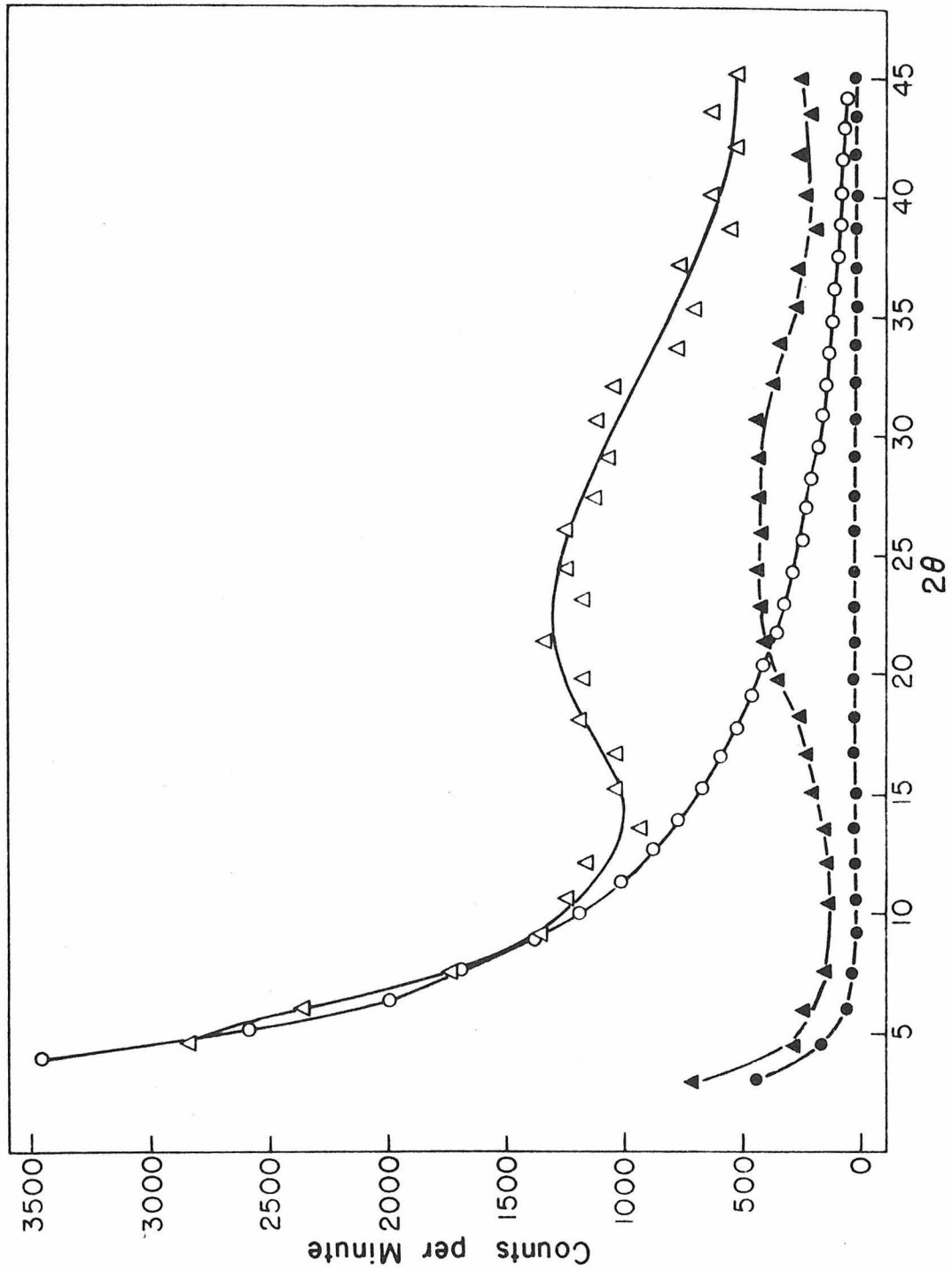

Counts per Minute
3500
3000
2500
2000
1500
1000
500
0
5
10
15
20
25
30
35
40
45
2θ

Figure 6


Counts per Minute
3500
3000
2500
2000
1500
1000
500
0
5
10
15
20
25
30
35
40
45
2θ

APPENDIX VII

Structure-Function Relationships in the Serine Proteases

R. M. Stroud, M. Krieger, R. E. Koeppe II, A. A. Kossiakoff,

and J. L. Chambers

Cold Spring Harbor Symposium on Proteases and Biological Control, 13-32 (1975).

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):

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

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,

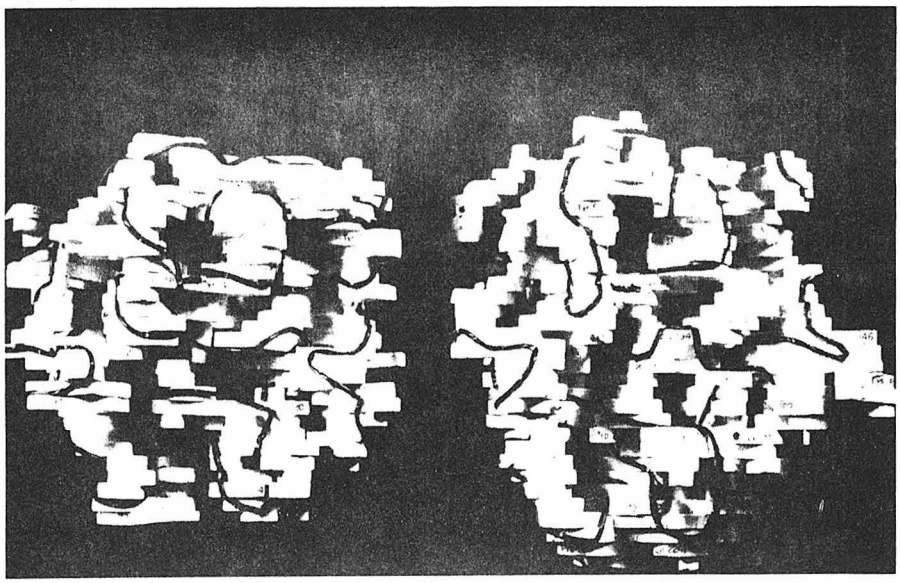


Figure 1
A comparison of the 5-Å models of trypsinogen (*left*) and DIP-trypsin (*right*) shows close structural homology in most areas of the molecule. Striking structural differences are observed in only two areas of the molecule. The first is in the binding pocket region, which is formed by residues 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

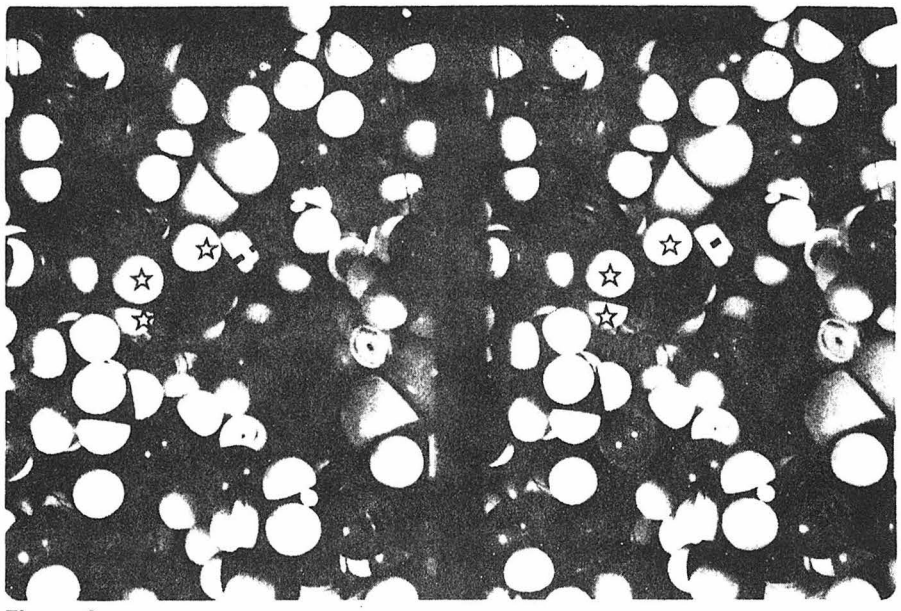


Figure 2

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

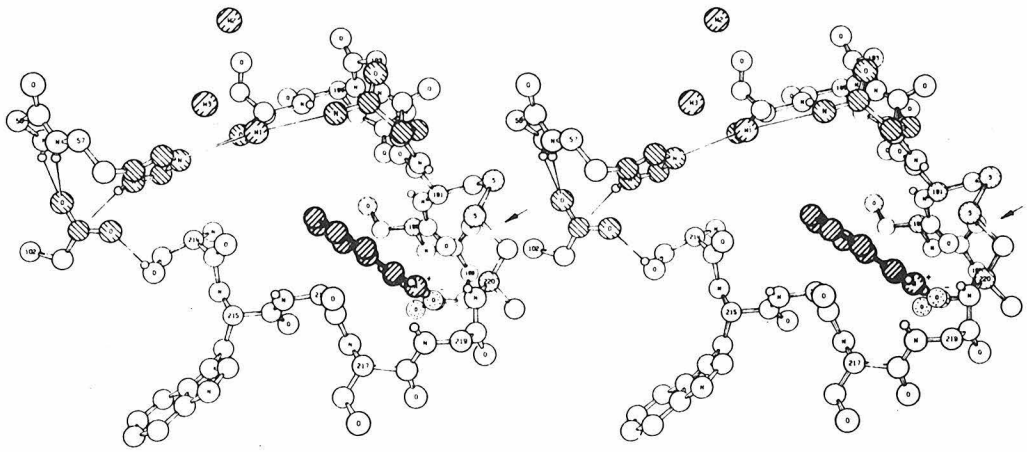


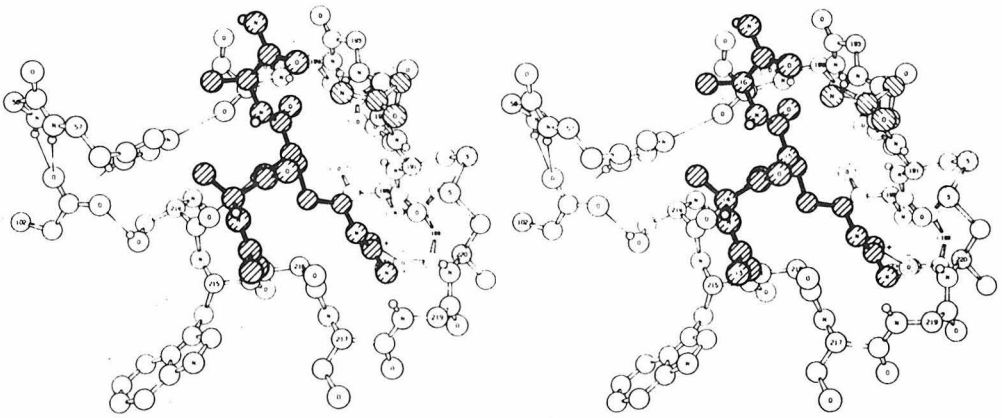
Figure 3

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.



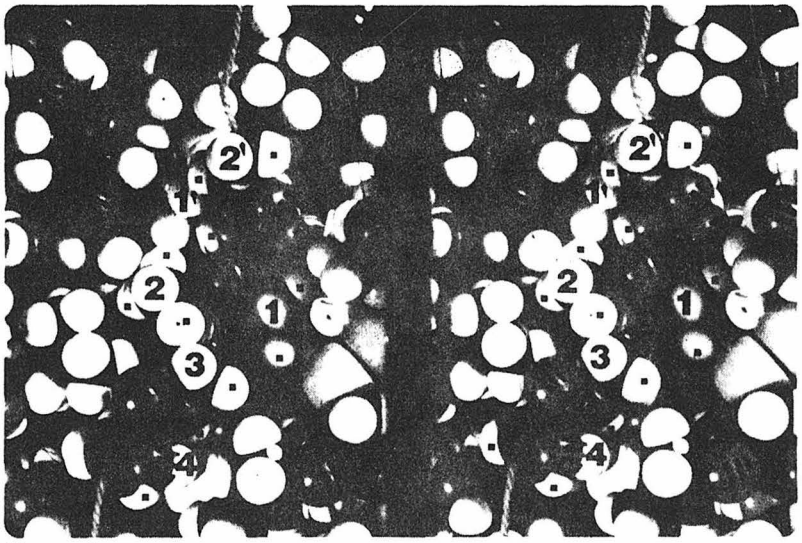


Figure 4

(*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

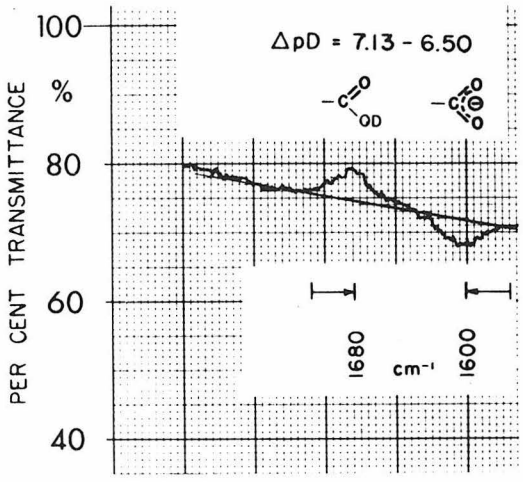
nuclear magnetic resonance studies that the imidazole of His-57 in α-lytic protease (a bacterial homolog of the pancreatic serine proteases) does not ionize until the pH is lowered below 4.0. Their research led them to propose that the group of pK_a 7.0 was Asp-102.

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.



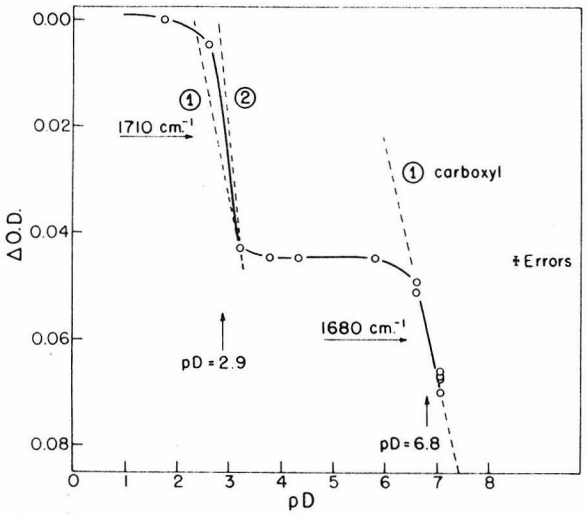


Figure 6
Titration curve for free carboxyls in semicarbazide-trypsin obtained by plotting differential absorbance at 1710 cm^{-1} and 1680 cm^{-1} from a series of infrared difference spectra. Reference solutions were at pD 3.2, 3.8, 4.4 and 6.6. The gradient of the titration at low pD was 1.5–2.0 times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

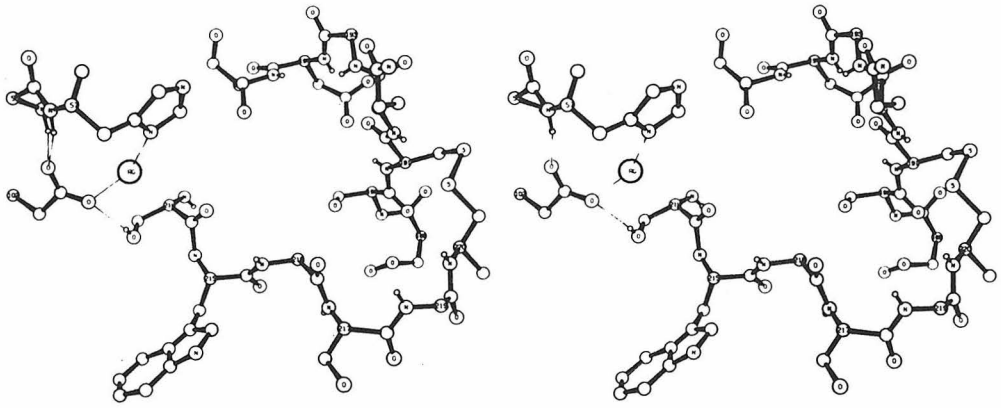


Figure 7
A stereo ORTEP drawing of the catalytic site of silver DIP-trypsin, shown in the same orientation as Figure 1. The DIP group has been omitted for clarity. The γ-oxygen of Ser-195 is in the position found for DIP-trypsin, close to that found in tetrahedral intermediates or in acyl enzymes (see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

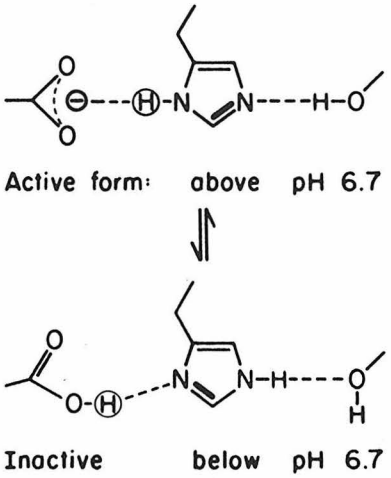


Figure 8
Ionization of the active center in the range pH 4.0–8.0 as discussed in the text.

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

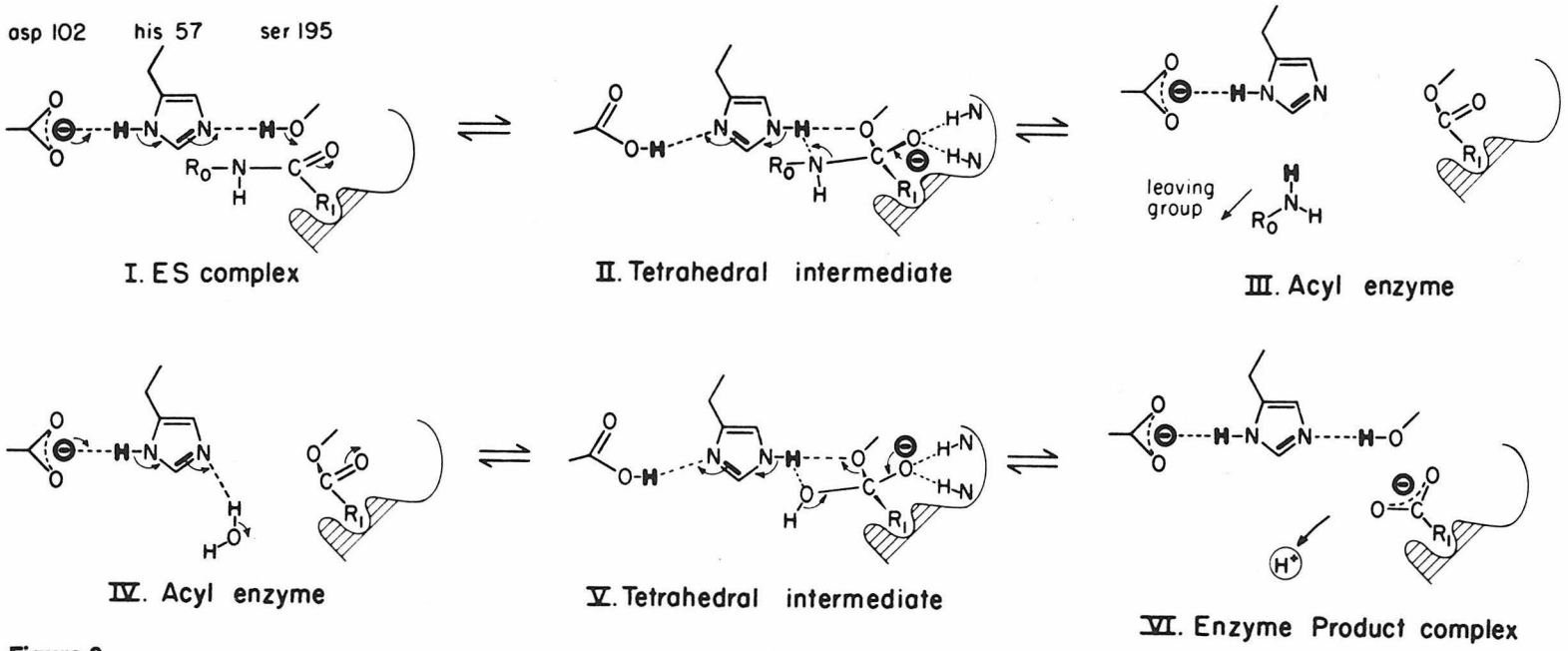


Figure 9
A mechanism for serine protease hydrolysis of peptides or amides. In this representation, the proton shuttle is concerted.

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 VIII

Plasminogen Binding to Balb/c 3T3 and SV 3T3 Cells

Plasminogen Binding to Balb/c 3T3 and SV3T3 Cells*

Monty Krieger

ABSTRACT: Plasminogen binding to Balb/c 3T3 and SV3T3 cells was studied using ^{125}I-labeled dog plasminogen. The binding appears to be independent of the plasmin-dependent morphological changes (PDMC) which are exhibited by many transformed cells. The plasminogen which binds to the 3T3 cells reaches approximately steady-state concentration while the plasminogen bound to the SV3T3 cells is completely degraded during three days of incubation to macromolecules which are the same size as the large and small chains of active plasmin and to 3-iodo-L-tyrosine. The results of a sublethal "cell surface" trypsinization assay suggest that the plasminogen and its macromolecular degradation products are bound to the surfaces of the 3T3 and SV3T3 cells, and that the plasminogen degradation on the SV3T3 cells may actually be the result of cell surface plasminogen activation.

---

*This research was performed in collaboration with Jerry Tobler and Robert Stroud.

INTRODUCTION

Dr. E. Reich and his colleagues have shown that the enhanced proteolytic activity of cells transformed by oncongenic DNA, RNA viruses, or chemical carcinogens is important in mediating certain morphological changes characteristic of transformation (1,2). Unkeless et al. (3) have shown that transformed cells synthesize an enzyme which activates serum plasminogen to plasmin. The increased proteolytic activity of transformed cells due to plasmin is responsible for these morphological changes (4,5). Cells which have undergone these morphological changes are spherical, tend to clump, and retract from the substratum. When cells are grown in plasminogen-free serum or in serum containing high concentrations of plasmin inhibitors, the morphological changes do not occur (5).

Ossowski et al. (6) have reported that simian virus 40 (SV40) transformed hamster embryo fibroblasts can bind more ^{125}I-labeled plasminogen than their normal counterparts, and that soybean-trypsin inhibitor decreased binding to the transformed cells. These binding data and reports of proteolytic activity associated with the plasma membranes of transformed cells (7) suggested the possibility of a specific cell surface-plasmin interaction. We have studied the binding of ^{125}I-plasminogen to normal and SV40-transformed

Balb/c 3T3 fibroblasts and its relationship to the plasmin-dependent morphological changes (PDMC) exhibited by the transformed cells.

EXPERIMENTAL

Binding Studies: The Balb/c 3T3 and Balb/c SV3T3 cells used were obtained from G. Todaro and were maintained in an incubator (37°C, 5% CO_2 atmosphere) in Dulbecco's modified Eagle's medium (Gibco) with 4.5 times the glucose concentration and supplemented with 10% fetal calf serum (10% FCS). The cells were periodically checked and found to be free of microplasmal infection (8, 9). Plasminogen isolated from dog serum by lysine sepharose affinity chromatography (10) at 4°C was iodinated using lactoperoxidase (Calbiochem, Lot 387009) and carrier free $Na^{125}I$ (Amersham-Searle)(11). The specific activity was generally 1×10^8 cpm/mg. Radioactivity was measured with a Nuclear Chicago Automatic Gamma Well Counter, Model 1085, and all protein concentrations were determined by the Lowry method (12) using bovine serum albumin (Sigma, Lot 43C-8120) as a standard. The plasminogen was radio-labeled one to three days before each experiment. The enzymatic activity of the plasminogen activated with Varidase (Lederle, Lot 406-334)(13) was unaffected by the iodination. All cell incubations were carried

out in a 37°C incubator (5% CO_2 atmosphere), and all other procedures were conducted at room temperature unless otherwise noted.

To determine the extent and the time course of plasminogen binding to the cells, the cells were plated at 4×10^5 cells/60 mm petri dish (Nunclon) in 5 ml of 10% FCS supplemented medium. After one day, the medium was removed and the cultures were washed twice with Tris buffer (0.025 M Tris, 0.14 M NaCl, 0.005 M KCl, 0.0007 M Na_2HPO_4, 0.0005 M $MgCl_2$, 0.0009 M $CaCl_2$, pH 7.4). Fresh medium supplemented with either 10% dog serum (Gibco) (10% DS) or 2% dog serum (2% DS), and ^{125}I-labeled dog plasminogen (final concentration = ~100 nM) were then added (t = 0). After incubation periods of 24, 48, and 72 hours, the cell morphology was scored to determine the extent of plasmin-dependent morphological change (PDMC) according to the criteria of Ossowski et al. (5), the incubation medium was removed, and the cells were detached from the petri dish using a rubber policeman. The cells were washed in 20 ml Tris buffer suspensions and pelleted (2500 rpm for five minutes) between each wash until they were free of unbound radio-labeled material. ("Off Plate" washing procedure.) Five to ten wash cycles were normally required to bring the supernatant counts down to background. After resuspension in a one-to-one mixture of 10% FCS supplemented medium and a 1% solution of trypsin (two times

recrystallized, Worthington Biochemical Corporation) in phosphate buffered saline (0.016 M Na_2HPO_4, 0.0015 M KH_2PO_4, 0.137 M NaCl, 0.0027 M KCl, 0.00049 M $MgCl_2$, 0.0009 M $CaCl_2$, pH 7.4), the number of cells was determined using hemocytometers (at least four independent determinations), and the radioactivity in the suspension was measured. The results of the binding studies will be presented as equivalent molecules of plasminogen bound per cell. These data are calculated from the specific activity of the plasminogen and the amount of cell associated radioactivity.

Similar binding studies were also conducted with SV40 transformed hamster embryo fibroblasts (SVHEF) kindly provided by Dr. D. Rifkin. The SVHEF cells were maintained under conditions identical to those for the Balb/c cells. For the binding studies, the SVHEF cells were incubated in 2% DS-supplemented medium and were washed in one of two ways. In one set of experiments, the cells were washed free of unbound radio-labeled material while attached to the substratum ("On-Plate" washing procedure). The cells were then released from the plate by trypsinization, suspended in the trypsin solution and counted. In other experiments, the Off-Plate wash was used.

Because an increase in plasminogen binding to the SV 3T3 and SVHEF cells appeared to accompany PDMC on the third day of the

binding experiments, a number of PDMC inhibitors and an accelerator of PDMC were included in the incubation media to assess the relationship between plasminogen binding and PDMC. Soybean-trypsin inhibitor (STI) (1,2) and pancreatic-trypsin inhibitor (PTI) can inhibit the PDMC expressed by SV 3T3 and SVHEF cells grown in DS-supplemented medium. PDMC is also inhibited if the cells are incubated in FCS rather than DS-supplemented medium (2). In order to inhibit the PDMC during plasminogen binding experiments, STI (Worthington Biochemical Corporation, lot 54J358) or PTI (Sigma, lot 81C-8200) was added to the DS-supplemented incubation medium at t = 0, or 10% FCS was substituted for DS in the medium. The concentrations of STI (14.6 μM for SV 3T3 cells) and PTI (6.7 μM for SV 3T3 and 15 μM for SVHEF cells) used to inhibit PDMC were determined by dose-response studies. The commercial STI was chromatographed on Sephadex G-75 at 4°C to remove contaminants. Varidase, which contains the plasminogen activator streptokinase, was used to accelerate the time course of PDMC by one day during a plasminogen binding experiment. The Varidase concentration required to do this (83.3 streptokinase units per dish) was determined by a dose-response experiment. If a higher Varidase concentration is used, the cells can be reduced to debris within 24 hours.

The binding of other ^{125}I-labeled proteins to the 3T3 and SV 3T3 cells was measured to compare with the binding of plasminogen. The proteins, their specific activities, and their concentrations in the incubation media were: PTI (1.5×10^8 cpm/mg, 0.61 μM), STI (1.2×10^7 cpm/mg, 14.6 μM), avidin-biotin complex (9.1×10^7 cpm/mg, 160 nM), ovalbumin (4.4×10^7 cpm/mg, 209 nM, two times recrystallized, Worthington Biochemical Corporation, lot 0435A953). The avidin-biotin complex was made by mixing an excess of biotin (Sigma, lot 94C-0187) with avidin (Sigma, lot 104C-0117) and purifying the iodinated complex on a Sephadex G-25 column at 4°C. The experimental procedure was identical to that used for plasminogen.

Characterization of Cell-bound ^{125}I-labeled Material: The results of the plasminogen binding experiments can only be interpreted properly when the iodinated species associated with the cells have been characterized. Therefore, the molecular weight distribution of the cell-bound iodinated molecules was analyzed by gel electrophoresis. After the radio-labeled cells were washed free of unbound radioactivity, they were dissolved in a 3% SDS-β-mercaptoethanol Tris buffer (3% SDS, 0.1 M Tris, 2.5% glycerol, 5% β-mercaptoethanol, 0.002% Brom Phenol Blue, pH = 6.8), and the macromolecules were denatured and reduced by heating for three minutes at 100°C. To reduce the viscosity of the sample, the

solution was then vigorously passed through a Pasteur pipette. The cell solutions were then electrophoresed in 5 mm diameter tubes on a 10% polyacrylamide separating gel overlayered with a 3% stacking gel according to Laemmli (14). The schlieren line due to the stacking buffer system was not allowed to run out of the gel during electrophoresis. The gels were frozen immediately after electrophoresis, sliced into 1 mm long pieces and counted. Small aliquots of the incubation media were treated similarly.

The dependence of the electrophoretic mobility on molecular weight was determined from Coomassie blue-stained gels (staining solution: 0.25% Coomassie blue, 25% propanol, 10% acetic acid; destaining solution: 10% acetic acid) using myosin (210,000 daltons), phosphorylase A (93,000 daltons), bovine serum albumin (68,000 daltons), catalase (60,000 daltons), aldolase (40,000 daltons), and myoglobin (17,000 daltons) as molecular weight standards. To further characterize the very low molecular weight iodinated molecules found associated with the SV 3T3 cells, small (25-100 μl) aliquots of the SDS-cell or SDS-supernatant solutions were chromatographed on Biogel P-2 columns using micro-step-elution chromatography. The columns were 3.5 cm long and were prepared by packing approximately 300 μl of Biogel P-2 (excludes molecules with molecular weights larger than approximately 1800) into a

conical polyethylene microcentrifuge tube (approximately 400 μl total capacity) which had a small (approximately 0.8 mm diameter) hole punched in the bottom with a small plug of glass wool blocking the hole. To obtain a chromatogram, the sample is layered on top of the packing material, the microcentrifuge tube is inserted into a larger 1 ml conical glass centrifuge tube and the apparatus is spun for one to two minutes in a clinical centrifuge (approximately 2500 rpm). Under the centrifugal force, the solution of macromolecules passes through the packing material into the bottom of the conical outer glass tube. The lip on the polyethylene tube keeps the microcentrifuge tube above the level of the liquid in the bottom of the outer glass tube. The portions of the sample remaining in the column can be eluted in steps by layering on additional column buffer (in this case, 3% SDS, 0.1 M Tris, pH = 6.8), transferring the polyethylene microcentrifuge tube to another 1 ml outer tube and repeating the centrifugation. The fractions can then be assayed, for example, by counting. This method is a modification of the desalting procedure described by Neal and Florini (15). Using a 100 μl sample and 50 and 100 μl elution steps, the lower molecular weight components of the SDS-cell solution were clearly resolved from the larger molecules in a seven-step chromatogram. The low molecular weight fraction from SV 3T3 cells incubated with ^{125}I-plasminogen for three days was

further analyzed by ascending paper chromatography (butanol:acetic acid:water, 4:1:2, v/v)(16). Standard samples of monoiodotyrosine (3-iodo-L-tyrosine, Aldrich lot 082837) and diiodotyrosine (3,5-diiodo-L-tyrosine, Aldrich lot 071637) were co-chromatographed with the low molecular weight fraction and were visualized with UV light and ninhydrin spray staining (0.2% in acetone, v/v).

Trypsinization Assay: A sub-lethal "surface" trypsinization assay was developed in an attempt to determine what fraction of the cell-associated radioactive material was cell-surface bound, and how much was internalized. 3T3 and SV 3T3 cells washed free of unbound radioactivity were resuspended in a one-to-one 10% FCS-supplemented medium:1% trypsin solution. At various time intervals, aliquots were assayed for release of cell-associated radioactivity by pelleting the cells at 4°C for 30 minutes at 7000 g and determining the relative amounts of radioactivity in the supernatant and in the pellet. Cell viability under these conditions was checked by measuring the relative plating efficiency and growth of the cells in 10% FCS-supplemented medium after the trypsinization treatment.

RESULTS AND DISCUSSION

Plasmin-dependent Morphological Changes (PDMC): Dog serum is an "activating" serum (2) for SV 3T3 and SVHEF cells; therefore, when these cells are grown in dog serum-supplemented media, they undergo PDMC. We found that the SVHEF cells exhibited dramatic PDMC between the second and third days of incubation in either 10% DS or 2% DS-supplemented medium. While minor PDMC occurred in the SV 3T3 cultures between 24 and 48 hours after transfer to DS-supplemented medium, the major PDMC also occurred between the second and third days. PDMC was not as dramatic with the SV 3T3 cells as with the SVHEF cells. 14.6 μM STI or 6.7 μM PTI completely inhibited PDMC with the SV 3T3 cells. 15 μM PTI was required to inhibit the PDMC occurring with the SVHEF cells and 0.6 μM PTI was sufficient to inhibit the PDMC manifested by a line of mycoplasma-infected Swiss SV 3T3 cells. Swiss SV 3T3 and Swiss 3T3 cells were provided by J. Vinograd and maintained in culture as described for the Balb/c cells. These results are consistent with the findings that PDMC depends upon serum plasminogen activation by a cell synthesized activator; thus, the different cell lines require different concentrations of plasmin inhibitors to prevent PDMC, presumably because they synthesize different

amounts or types of the activator. Fetal calf serum is a "nonactivating" serum (2) for SV 3T3 and SVHEF cells, and no PDMC was observed when the cells were grown in 10% FCS. In contrast to previous reports of the effects of trypsin inhibitors on cell growth (17), we found no differences in the doubling times for Balb/c SV 3T3 (~17 hours) or SVHEF (~14 hours) cells grown in media which included 2% DS, 10% DS, 10% or 2% DS supplemented with STI or PTI, or 10% FCS. Varidase, a plasminogen activator, could accelerate the onset of PDMC in Balb/c SV 3T3 cells; however, the concentrations of Varidase necessary to produce a one-day shift in the PDMC time course (83.3 streptokinase units/dish) inhibited cell growth.

The untransformed Balb/c and Swiss 3T3 cells do not exhibit PDMC; however, trypsin inhibitors did delay the detachment of the cells from the petri dishes when the cells were allowed to grow for extended periods (5 to 8 days). This suggests that a serine protease may be involved in the detachment process for normal cells.

In one experiment, the fibrinolytic activity in the 10% DS-supplemented incubation medium was assayed using the ^{125}I-fibrin plate method (1). The fibrinolytic activity in the medium was low for the Balb/c 3T3 cells throughout the three-day experiment, and for the Balb/c SV 3T3 cells for the first two days of the experiment. On

the third day of incubation, there was a substantial increase in the fibrinolytic activity in the medium in which the SV 3T3 cells were grown. This increase accompanied the PDMC. Using the harvest fluid fibrin plate assay (1), in which the cells are incubated in media without serum supplement for eighteen hours, the media for both the Balb/c 3T3 and SV 3T3 cells showed high fibrinolytic activity. It can be argued that the use of the incubation medium for the fibrinolytic assay is superior to using harvest fluid or growing the cells on fibrin-coated dishes because using the incubation medium allows a direct assessment of the fibrinolytic activity without perturbing the environment of the cells.

Plasminogen Binding: After incubation for one, two, or three days in ^{125}I-plasminogen and 10% DS-supplemented medium, the Balb/c 3T3 and SV 3T3 cells were removed from the petri dishes by scraping, thoroughly washed and counted to determine the specific activity. The results averaged over three to six separate experiments are illustrated in Figure 1. The plasminogen binding curves (plasminogen bound per cell vs. time) for the 3T3 cells were either "V" shaped or showed a continuous decrease in binding with time, in which case the binding decreased more dramatically between the first and second days than between the second and third days (distorted L-shaped curves). The SV 3T3 binding curves were

generally "V" shaped, although the relative heights of the sides of the "V" did vary. Binding curves for SVHEF cells (Figure 2) show that plasminogen binding increases throughout the three-day incubation with the largest increase occurring between the second and third days.

Inhibitors of PDMC (STI, PTI, and fetal calf serum) were added to the incubation media in several binding experiments in order to study the relationship between PDMC and plasminogen binding. The ratios of the extent of plasminogen binding in the presence of inhibitors to that observed without the inhibitors are listed in Table I. The values in the table are the averages of the ratios observed on each of the three days of the experiments. The effect of PTI on plasminogen binding was found to depend on when the inhibitor was added to the incubation media, and the cell washing procedure. A "pulse" of PTI added two hours before the cells were assayed for plasminogen binding had no effect on the binding. When the PTI was added at $t = 0$ and left in the medium throughout, it had little if any statistically significant effect on the plasminogen binding to the Balb/c 3T3, Balb/c SV 3T3 or SVHEF cells washed using the Off-Plate procedure. When the On-Plate wash was used, PTI added to the incubation media of Balb/c SV 3T3 or SVHEF cells at $t = 0$ reduced the plasminogen binding by about

50%. Ossowski et al. (6) reported that STI blocked the uptake of ^{125}I-labeled plasminogen by SVHEF cells when the binding was assayed using the On-Plate wash procedure. Their experiments differed from those reported here in that the dog serum used to supplement the incubation media was subjected to affinity chromatography and dialysis to remove endogenous plasminogen and the time course of binding was followed for 24 hours rather than 72 hours.

STI had essentially no effect on the plasminogen binding to Balb/c SV 3T3 cells and apparently induced a slight increase in binding to Balb/c 3T3 cells (Table I). In these experiments, STI was added at $t = 0$ and the cells were washed with the Off-Plate technique. The effects of fetal calf serum on plasminogen binding are difficult to assess because of the substantial variability in the results (Table I). Varidase treatment of the cells can accelerate the onset of PDMC, but it did not affect the shapes of the binding curves for the Balb/c 3T3 and SV 3T3 cells. At the concentrations of Varidase used in the binding experiments (83.3 streptokinase units/dish), PDMC occurred approximately 24 hours earlier than normally, and cell growth was inhibited.

The discrepancies between the Off-Plate and the On-Plate assay results may be due to two problems associated with the

On-Plate technique. In a control experiment, SV 3T3 cells which were incubated for three days in ^{125}I-plasminogen and 10% DS-supplemented medium were washed while attached to the petri dish, then removed from the dish with a rubber policeman. The cell free plates were extensively washed with Tris buffer and then trypsinized. The trypsinization solution contained ^{125}I-radioactivity; therefore, the On-Plate wash assay may have monitored dish-associated as well as cell-associated radioactivity. Additional support for this observation was found in a study of binding to SVHEF cells in replicate petri dishes. The "plasminogen binding" was higher when the cells were washed on the plates than when the Off-Plate wash was used. An additional problem with the On-Plate wash is that free-floating or loosely-bound cells are washed away, thus biasing the cell population with firmly-attached cells. This selection can be particularly significant after PDMC occurs and many cells have retracted from the petri dish. In contrast, the entire cell population is assayed when the Off-Plate procedure is used.

The extent of plasminogen binding to the Balb/c 3T3 and SV 3T3 cells (10^4 - 10^5 molecules/cell), determined using the Off-Plate procedure, was within the range of binding observed for other ^{125}I-labeled proteins (Table II). There is no apparent correlation between the extent of binding and the molecular weight or the

isoelectric point.

Characterization of the Cell-bound Radio-labeled Molecules: ^{125}I-radio-labeled molecules associated with the cells or in the incubation media were analyzed by polyacrylamide gel electrophoresis, Biogel P-2 chromatography and paper chromatography after solubilization in SDS and β-mercaptoethanol reduction. The variation in electrophoretograms of solubilized Balb/c 3T3 and SV 3T3 cells with time is shown in Figure 3. The electrophoresis shows that the plasminogen (molecular weight = 93,800 ± 6,500) was degraded by the Balb/c SV 3T3 cells into at least three species: 1) a 71,300 ± 3,800 molecular weight molecule which is the size of the larger of the two polypeptide chains of plasmin ("B protein"); 2) a 25,000 ± 800 molecular weight molecule which is the size of the smaller of the plasmin chains ("C protein"), and 3) low molecular weight material ("LMW"). Similar electrophoretograms were prepared from Balb/c 3T3 and SV 3T3 cells incubated with STI to prevent PDMC, and from the incubation media of the experiments. The peaks in the electrophoretograms were integrated and the relative amount of each species (as measured by radioactivity, not concentration) was calculated (see Figure 4). Radio-labeled material of very low electrophoretic mobility (HMW) was observed in the

electrophoretograms. This material has not been identified and may be an artifact due to excess β-mercaptoethanol in the SDS-cell solutions (R. Vandlen, personal communication). The relative amount of plasminogen associated with the Balb/c 3T3 cells remained fairly constant throughout the incubation while the plasminogen associated with the transformed cells was completely degraded after three days. The degradation of the plasminogen was accompanied by an increase in the relative amount of the C-protein, whereas the relative amount of the B-protein remained fairly constant. The addition of STI to the incubation medium at a concentration sufficient to inhibit PDMC had very little effect on the molecular weight distribution of the ^{125}I-labeled molecules; therefore, active plasmin in the medium cannot be responsible for the degradation of the cell-associated plasminogen. The STI results also show that the plasminogen degradation is independent of PDMC.

55% of the total radioactivity in the Balb/c SV 3T3 cells solubilized after 72 hours of incubation was small enough to be retained by the Biogel P-2 (≤ 1800 molecular weight) micro-chromatography column. This material co-migrated with monoiodotyrosine during ascending paper chromatography. Since approximately 60% of the total radioactivity in Balb/c SV 3T3 cells after three days of incubation migrates as low molecular weight material during gel

electrophoresis (Figure 3), almost all of the low molecular weight material is monoiodotyrosine.

One possible explanation for the degradation of the cell-associated plasminogen might have been that the plasminogen in the incubation medium was being degraded and that the distribution of molecules bound to the cells simply reflected the distribution of molecules in the medium because of direct cell uptake. Electrophoretograms of the incubation media, however, show that the distribution of molecules in the media is different from that on the cells (see Figures 4 and 5). Therefore, the degradation of plasminogen in the medium does not directly account for the degradation of the cell-associated plasminogen.

Electrophoretograms of the media (Figure 5) also show that plasminogen was partially activated to plasmin in the media in which the Balb/c SV 3T3 cells grew, but not in the media from the Balb/c 3T3 cells. STI did not inhibit this activation of plasminogen. Therefore, the plasminogen activator synthesized by Balb/c SV 3T3 cells appears to be similar to the activator synthesized by Schmidt-Ruppin RSV-A infected chicken embryo fibroblasts in that both activators are insensitive to macromolecular trypsin inhibitors (3).

Because the concentration of plasminogen molecules associated with the Balb/c 3T3 cells reaches an approximately steady state

during the incubation while all of the cell-bound plasminogen is degraded by the Balb/c SV 3T3 cells, a thorough understanding of the plasminogen binding will require the determination of the location of the bound molecules on or in the cells, the mechanism of degradation and the principal sources of the differences between the normal and transformed cells.

In a preliminary attempt to localize the cell-bound ^{125}I-labeled protein, the cells were trypsinized with the goal of gently stripping off cell surface proteins without releasing molecules which had been internalized. The time course of ^{125}I-release from labeled cells showed an exponential release of radioactivity which plateaued when approximately 50 - 80% of the cell-bound radioactivity was released by the trypsinization (Figure 6).

The kinetics of release were similar for the Balb/c 3T3 and SV 3T3 cells. Half of the counts which could be released by trypsin (trypsinizable counts) were released after six minutes and the plateau was reached by approximately 50 minutes of trypsinization. If the period of trypsinization was extended beyond 100 minutes, there was an additional release of radioactivity. The cell viability began to fall after 60 minutes of trypsinization; therefore, the increased release of radioactivity after prolonged trypsinization was probably caused by cell damage and release of internalized

radioactivity. A twenty-minute standard trypsinization incubation time was therefore chosen because approximately 90% of the trypsinizable counts were released after twenty minutes, and there was no loss of cell viability.

The amounts of radioactivity released by trypsin (averaged over five separate binding experiments) are listed in Table III. The percentages of the cell-associated ^{125}I-labeled material which was found in large macromolecules (with molecular weights greater than 15,000) are also included in the table for comparison. The results of the trypsinization assay and the gel electrophoresis are strikingly similar. This would suggest that almost all of the ^{125}I-labeled cell-associated macromolecules are bound to the external surfaces of the cells and that the monoiodotyrosine is inside the cells. Because two of the macromolecules (the B and C proteins) are plasminogen degradation products, plasminogen degradation to the B and C proteins must occur at the cell surface. This interpretation assumes that the trypsinization assay is working properly; that is, that the trypsinization is efficiently stripping off only cell surface-bound molecules without releasing those which have been internalized. While the bulk of the data in Table III supports this assumption, the 72-hour SV 3T3 data are not consistent with it, for the percentage of ^{125}I-labeled material trypsinized off is larger than the percentage

of ^{125}I-labeled macromolecules. Further experiments are required to validate, or invalidate, the 72-hour data point and permit greater elaboration of models for the plasminogen degradation.

One question which arises is that, if the cell-associated B and C proteins are the large and small fragments of plasmin, then the alteration of plasminogen on the cell surface is its activation to plasmin. This suggests a specific type of proteolytic activity which is localized on the cell surface. Could this be due to the plasminogen activator of transformed cells? If so, that in turn would suggest that the activator might have an important role at the cell surface, rather than just in the medium. We are currently conducting autoradiographic experiments which should further test the validity of the cell-surface trypsinization assay and help to more precisely locate the cell-associated radio-labeled molecules.

Acknowledgments: We thank Dr. Jerome Vinograd for generously allowing us to use his cell culture facilities, and Dr. Ray Owen for the use of the gamma spectrometer. Standard plasmin samples were provided by the American Red Cross. We thank Marlyn Teplitz and Maura K. Dwyer for expert technical assistance, and Dr. Ray Teplitz for helpful advice and discussions.

REFERENCES

1. Unkeless, J.C., Tobia, A., Ossowski, L., Quigley, J.P., Rifkin, D.B., and Reich, E. (1973) J. Exp. Med. 137, 85-111.
2. Ossowski, L., Unkeless, J.C., Tobia, A., Quigley, J.P., Rifkin, D.B., and Reich, E. (1973) J. Exp. Med. 137, 112-126.
3. Unkeless, J., Dano, K., Kellerman, G.M., and Reich, E. (1974) J. Biol. Chem. 249, 4294-4305.
4. Quigley, J.P., Ossowski, L., and Reich, E. (1974) J. Biol. Chem. 249, 4306-4311.
5. Ossowski, L., Quigley, J.P., and Reich, E. (1974) J. Biol. Chem. 249, 4312-4320.
6. Ossowski, L., Quigley, J.P., Kellerman, G.M., and Reich, E. (1973) J. Exp. Med. 138, 1056-1064.
7. Burger, M.M. (1971) in Growth Control in Cell Cultures (G.E.W. Wolstenholme and J. Knight), pp. 45-63. Churchill Livingstone, Edinburgh and London.
8. Brown, S., Teplitz, M., and Revel, J. (1974) Proc. Nat. Acad. Sci. USA 71, 464-468.
9. Russell, W.C., Newman, C., and Williamson, D H. (1975) Nature 253, 461-462.

10. Deutsch, D. G. and Mertz, E. T. (1970) Science 170, 1095-1096.

11. Marchalonis, J. J. (1969) Biochem. J. 113, 299-305.

12. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) J. Biol. Chem. 193, 265-275.

13. Silverstein, R. M. (1975) Analytical Biochemistry 65, 500-506.

14. Laemmli, U. K. (1970) Nature 227, 680-685.

15. Neal, M. W. and Florini, J. R. (1973) Analytical Biochemistry 55, 328-330.

16. Sato, S. and Tamiya, N. (1970) J. Biochem. 68, 867-872.

17. Schnebli, H. P. and Burger, M. M. (1972) Proc. Nat. Acad. Sci. USA 69, 3825-3827.

18. Summaria, L., Arzadon, L. A., Bernabe, P., Robbins, K. C., and Barlow, G. H. (1973) J. Biol. Chem. 248, 2984-2991.

19. Green, N M. (1957) Biochem. J. 66, 407-415.

20. Chauvet, J., Nouvel, G., and Acher, R. (1964) Biochim. Biophys. Acta 92, 200-201.

21. Ozawa, K. and Laskowski, M. Jr. (1966) J Biol. Chem. 241, 3955-3961.

22. Steiner, R. F. and Edelhoch, H. (1963) J. Biol. Chem. 238, 925-930.

23. Green, N. M. (1964) Biochem. J. 92, 16c-17c.

24. Wooley, D. W. and Longsworth, L. G. (1942) J. Biol. Chem. 142, 285-290.

25. Guntelberg, A. V. and Linderstrøm-Lang (1949) Compt. Rend. Trav. Lab. Carlsberg, Ser. Chim. 27, 1-25.

26. Cannan, R. K., Kirbrick, A. C., and Palmer, A. H. (1941) Ann. N. Y. Acad. Sci. 41, 243-266.

TABLE I

Effect of PDMC Inhibitors on Plasminogen Binding

Cell Type	Assay Technique	Inhibitor[a]	Plasminogen Binding Ratios[b] (with inhibitor/without inhibitor)
Balb/c 3T3	Off-Plate	PTI	1.65 ± 0.6
		PTI^{pulse}	1.15 ± 0.4
		STI	1.8 ± 0.2
		FCS	1.1 ± 0.6
Balb/c SV 3T3	Off-Plate	PTI	1.4 ± 0.3
		PTI^{pulse}	0.9 ± 0.2
		STI	1.15 ± 0.4
		FCS	1.9 ± 0.7
	On-Plate	PTI	0.4 ± 0.5
SVHEF	Off-Plate	PTI	1.1 ± 0.1
		PTI^{pulse}	1.2 ± 0.15
		FCS	2.6 ± 1.8
	On-Plate	PTI	0.45 ± 0.2
		PTI^{pulse}	1.0 ± 0.3
		FCS	1.5 ± 1.0

[a] The inhibitors were added at t = 0 unless indicated by the superscript "pulse", in which case the inhibitor was added two hours before the binding assay. Abbreviations: PTI, pancreatic trypsin inhibitor; STI, soybean trypsin inhibitor; FCS, fetal calf serum.

[b] Ratios from the 24, 48, and 72-hour observations were averaged together.

TABLE II

Binding of ^{125}I-labeled Proteins to Balb/c 3T3 and Balb/c SV 3T3 Cells

Molecule	Molecular Weight	Isoelectric Point	Concentration In Incubation Medium (μM)	Typical Binding (Molecules/cell)	
				Balb/c 3T3	Balb/c SV 3T3
Plasminogen	94,000	6.3-8.1[a]	0.1	1×10^5	6×10^4
PTI	6,500[b]	10.5[c]	0.6	5×10^6	2×10^6
STI	22,461[d]	4.6[e]	14.6	5×10^6	1×10^6
Avidin-biotin	68,000[f]	>10 [g]	0.16	1×10^5	9×10^4
Ovalbumin	45,000[h]	4.5[i]	0.21	2×10^4	8×10^3

[a] Values for human plasminogen (18).

[b] See reference 19.

[c] See reference 20.

[d] See reference 21.

[e] See reference 22.

[f] See reference 23.

[g] See reference 24.

[h] See reference 25.

[i] See reference 26.

TABLE III

Sub-lethal, 20-Minute Trypsinization Assay Results

Cell Type	Incubation Time (hours)	Cell-associated Radioactivity Released by Trypsinization[a] (% of total)	^{125}I-labeled Cell-associated Macromolecules[b] (% of total)
SV 3T3	24	74 ± 5	82 ± 7[c]
	48	54 ± 8	58 ± 0.4[c]
	72	64 ± 15	27 ± 13[c]
3T3	24	78 ± 8	83
	48	58 ± 7	81
	72	64 ± 7	61

[a] Average values from five independent experiments.

[b] Macromolecules are defined to have molecular weights greater than 15,000. The data were obtained by integrating the peaks on electrophoretograms.

[c] Average values from two independent experiments ± the deviation of either datum from the mean.

FIGURE CAPTIONS

Fig. 1: Time course of ^{125}I-plasminogen binding to Balb/c 3T3 and Balb/c SV 3T3 cells. The cells were incubated in medium supplemented with 10% dog serum and 100 nM ^{125}I-labeled dog plasminogen and assayed for cell-associated radioactivity using the Off-Plate washing procedure described in the text. The specific radioactivity of the cells was converted to plasminogen molecule equivalents per cell based on the specific activity of the plasminogen. There were two distinct types of binding curves for the Balb/c 3T3 cells (left): High binding on the first day followed by decreased binding on the second and third days (distorted L shape, upper curve) or V shaped binding (lower curve). Each of these curves is the average from three independent experiments which involved a total of ten separate petri dishes averaged together for each datum in the upper curve and five dishes per point for the lower curve. The V shaped binding curve for the Balb/c SV 3T3 cells is the average from six independent experiments in which fifteen separate petri dishes were averaged together for each point. For both the Balb/c 3T3 and SV 3T3 cells, the standard deviation for replicate plates within an experiment was about 20% of the total binding.

Fig. 2: Time course of ^{125}I- plasminogen binding to SVHEF cells. The cells were incubated in medium supplemented with 2% dog serum and 100 nM ^{125}I-labeled dog plasminogen and assayed for cell-associated radioactivity using the Off-Plate washing procedure as described in the text. The specific radioactivity of the cells was converted to plasminogen molecule equivalents per cell based on the specific activity of the plasminogen. The width of the error bars represents two standard deviations from the mean calculated by averaging together this binding curve and two binding curves in which PDMC inhibiting concentrations of PTI were added at t = 0 or two hours before the binding assay. All three curves were observed during the same experiment and show that PTI had no statistically significant effect on plasminogen binding.

Fig. 3: Electrophoretograms of SDS solubilized, β-mercapto-ethanol reduced Balb/c 3T3 and Balb/c SV 3T3 cells. After 24, 48, and 72 hours of incubation in ^{125}I-plasminogen and 10% dog serum-supplemented medium, the cells were washed using the Off-Plate technique, dissolved in a 3% SDS-β-mercaptoethanol-Tris buffer, boiled for three minutes and 100-200 μl aliquots containing 1500 to 6300 counts per minute were electrophoresed on 10% polyacrylamide tube gels. The relative radioactivity in each 1 mm gel slice is

presented as a function of the position of the slice relative to the Brom Phenol Blue tracking dye (R_f = 1.0). Because the amounts of radioactivity in the samples were different for each gel, relative radioactivity scales, which are normalized to the total radioactivity on the gels, were used.

Fig. 4: Time dependence of the relative concentrations of the Balb/c 3T3 (— — —) and Balb/c SV 3T3 (————) cell-associated ^{125}I-labeled molecules. The peaks in the electrophoretogram in Fig. 3 were integrated and the relative amounts of each species (% of total) are presented as a function of time. Nearly identical electrophoretograms of Balb/c 3T3 (— — —) and Balb/c SV 3T3 (————) cells which had been incubated with PDMC-inhibiting concentrations of STI (14.6 μM) were similarly analyzed.

Fig. 5: Time dependence of the relative concentrations of ^{125}I-labeled molecules in the incubation media. Samples of the incubation media from an ^{125}I-plasminogen binding experiment on Balb/c 3T3 (— — —) and Balb/c SV 3T3 (————) cells were denatured in 3% SDS-β-mercaptoethanol-Tris buffer and electrophoresed on 10% polyacrylamide gels. The peaks in the electrophoretograms were integrated and the relative amounts of each species

(% of total) are presented as a function of time. Electrophoretograms of Balb/c 3T3 (— — —) and Balb/c SV 3T3 (———) incubation media which included PDMC-inhibiting concentrations of STI (14.6 μM) were similarly analyzed. The data represented in this figure and in Figs. 3 and 4 were observed during the same experiment.

Fig. 6: Cell-surface trypsinization assay. Balb/c SV 3T3 cells were incubated in ^{125}I-plasminogen and 10% dog serum-supplemented medium for three days and then washed using the Off-Plate technique. The release of radioactivity by trypsinization was followed as a function of time. An exponential function was fit to the data by the method of least squares (solid line) and showed a half-time for the trypsinization of six minutes.

Figure 1

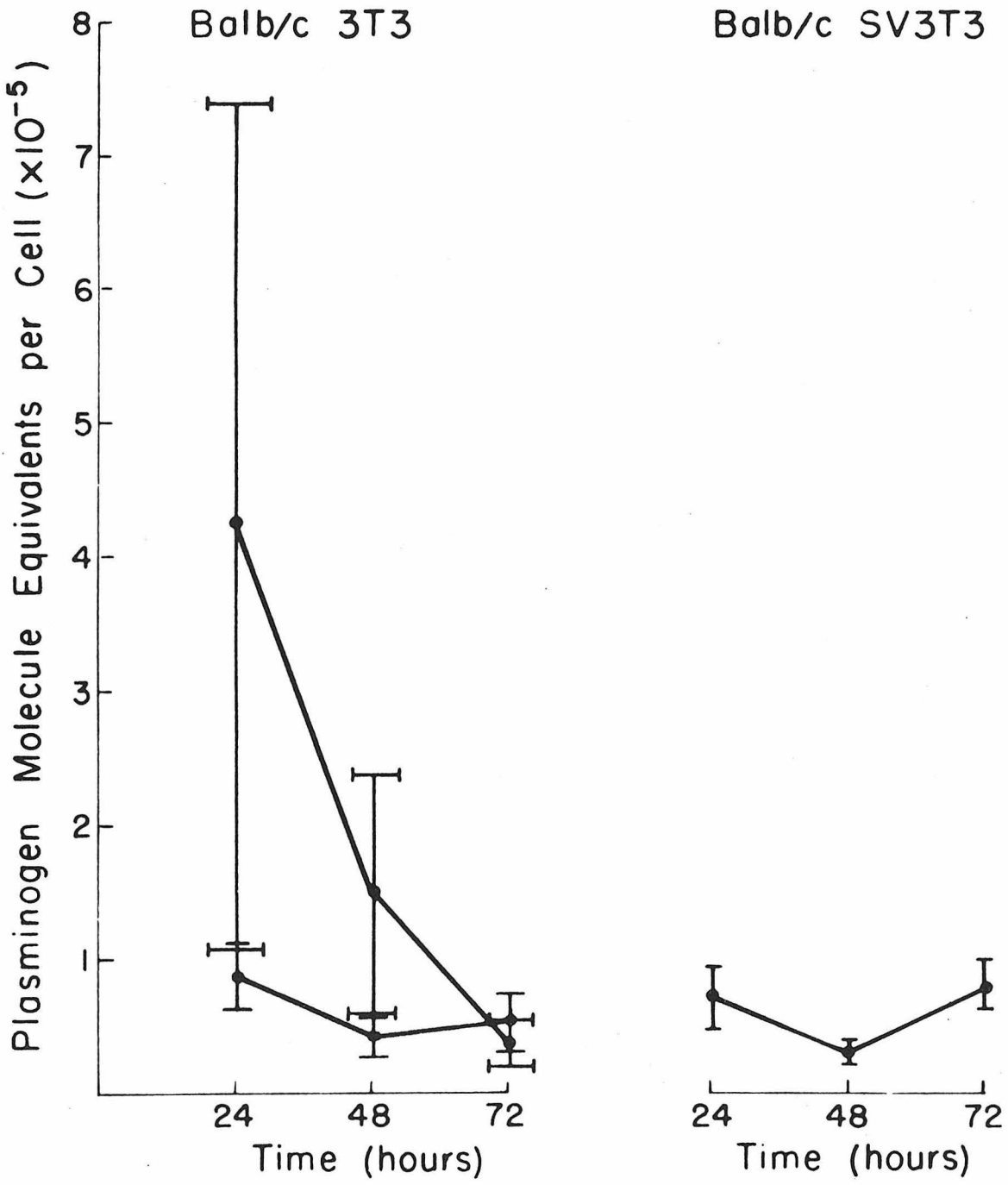

Balb/c 3T3
Balb/c SV3T3
Plasminogen Molecule Equivalents per Cell (x10⁻⁵)
8
7
6
5
4
3
2
1
24
48
72
Time (hours)
24
48
72
Time (hours)

Figure 2

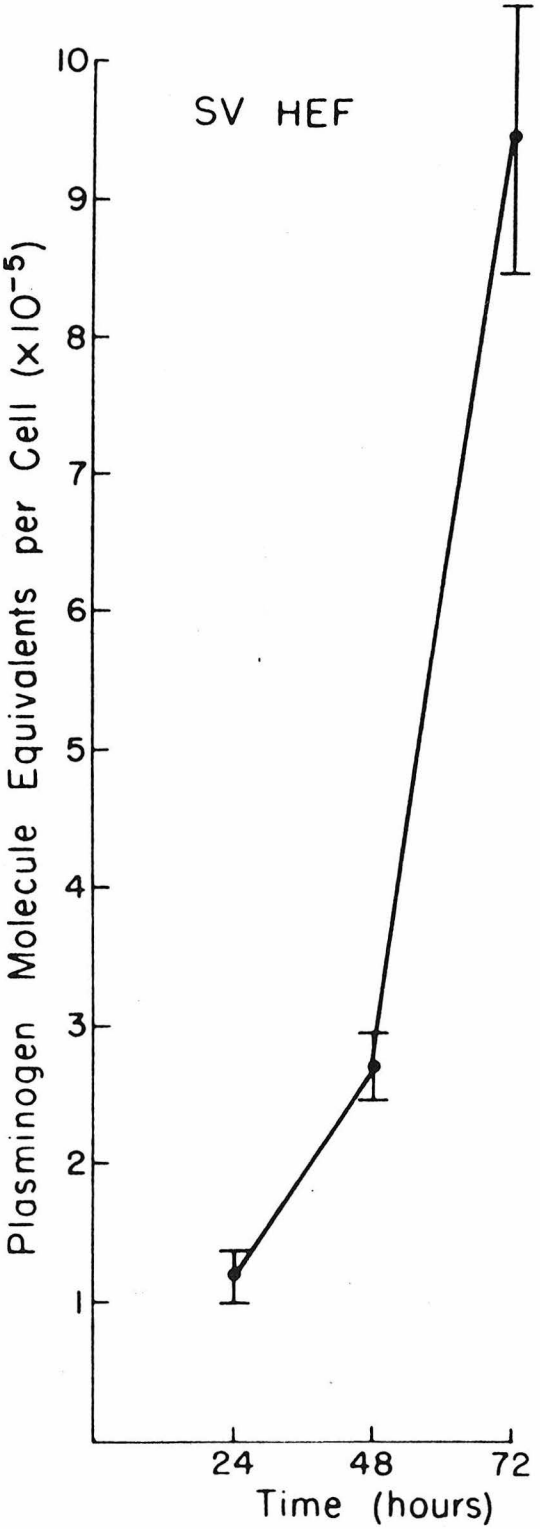

SV HEF
Plasminogen Molecule Equivalents per Cell (x10⁻⁵)
10
9
8
7
6
5
4
3
2
1
24
48
72
Time (hours)

Figure 3

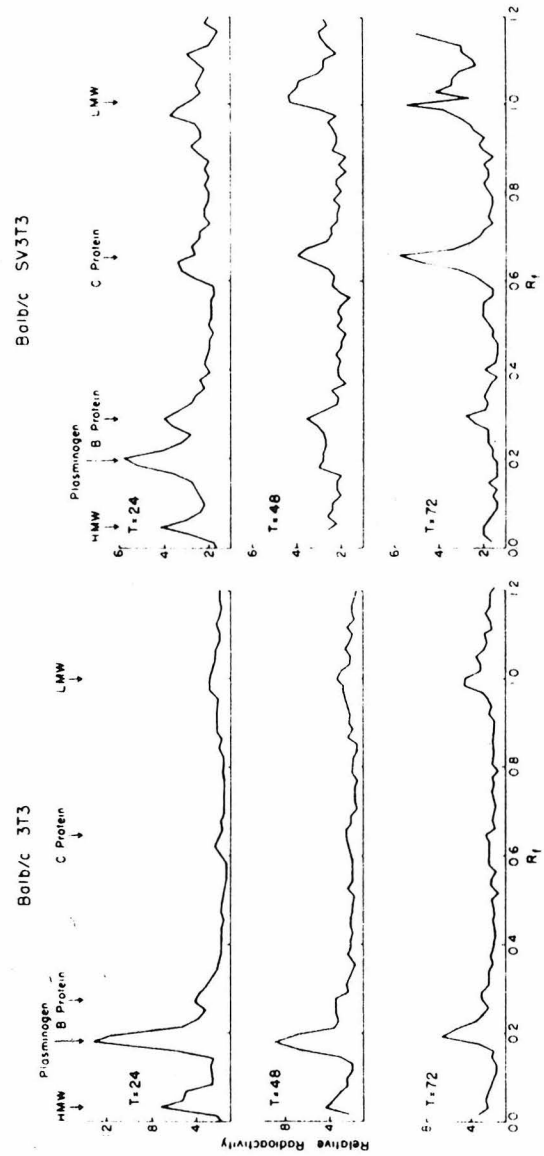

Balb/c 3T3
Balb/c SV3T3
Plasminogen
HMW
B Protein
C Protein
LMW
T=24
T=48
T=72
Relative Radioactivity
R_f

Figure 4

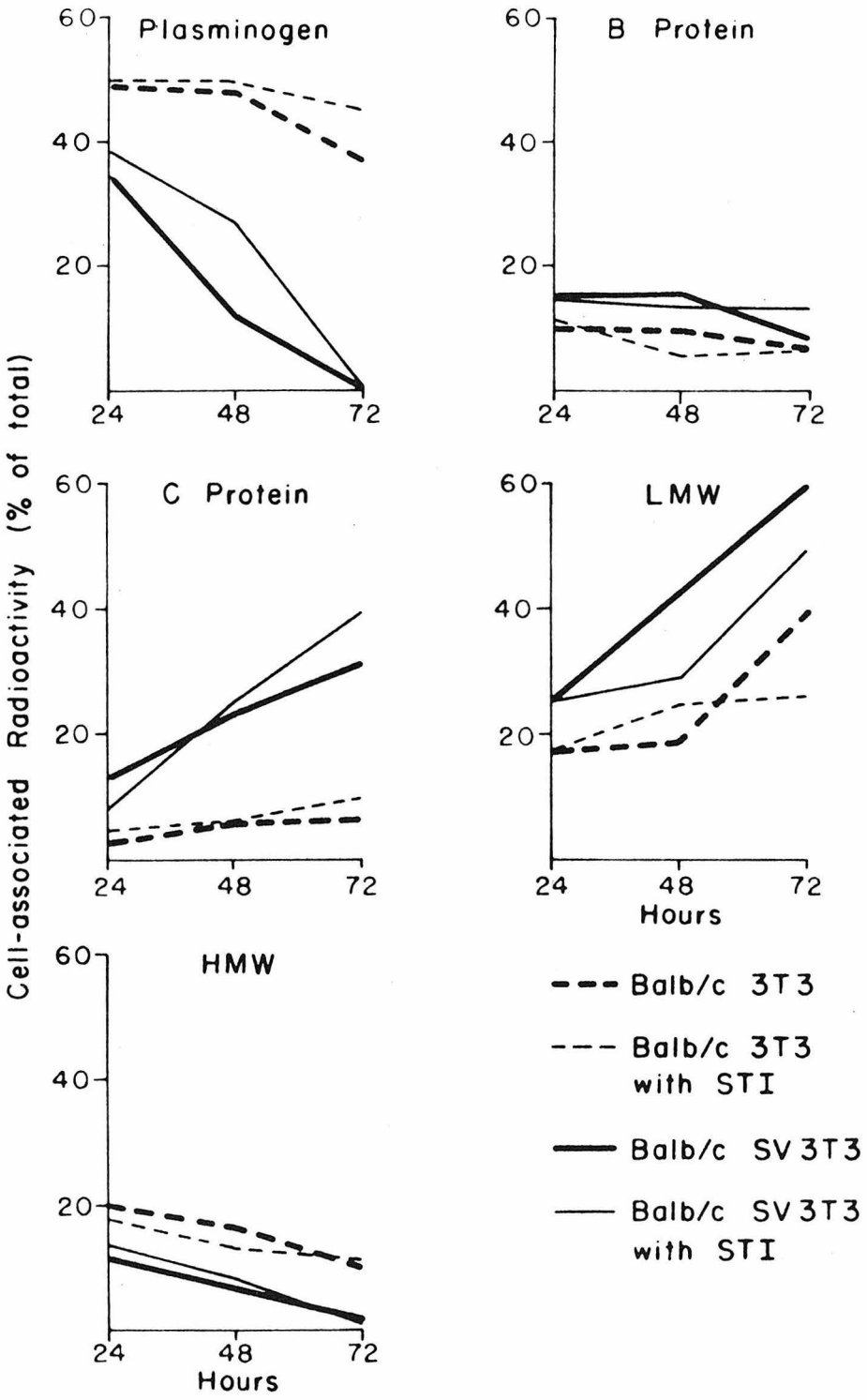

Plasminogen
B Protein
C Protein
LMW
HMW
Cell-associated Radioactivity (% of total)
60
40
20
24
48
72
Hours
Balb/c 3T3
Balb/c 3T3 with STI
Balb/c SV3T3
Balb/c SV3T3 with STI

Figure 5

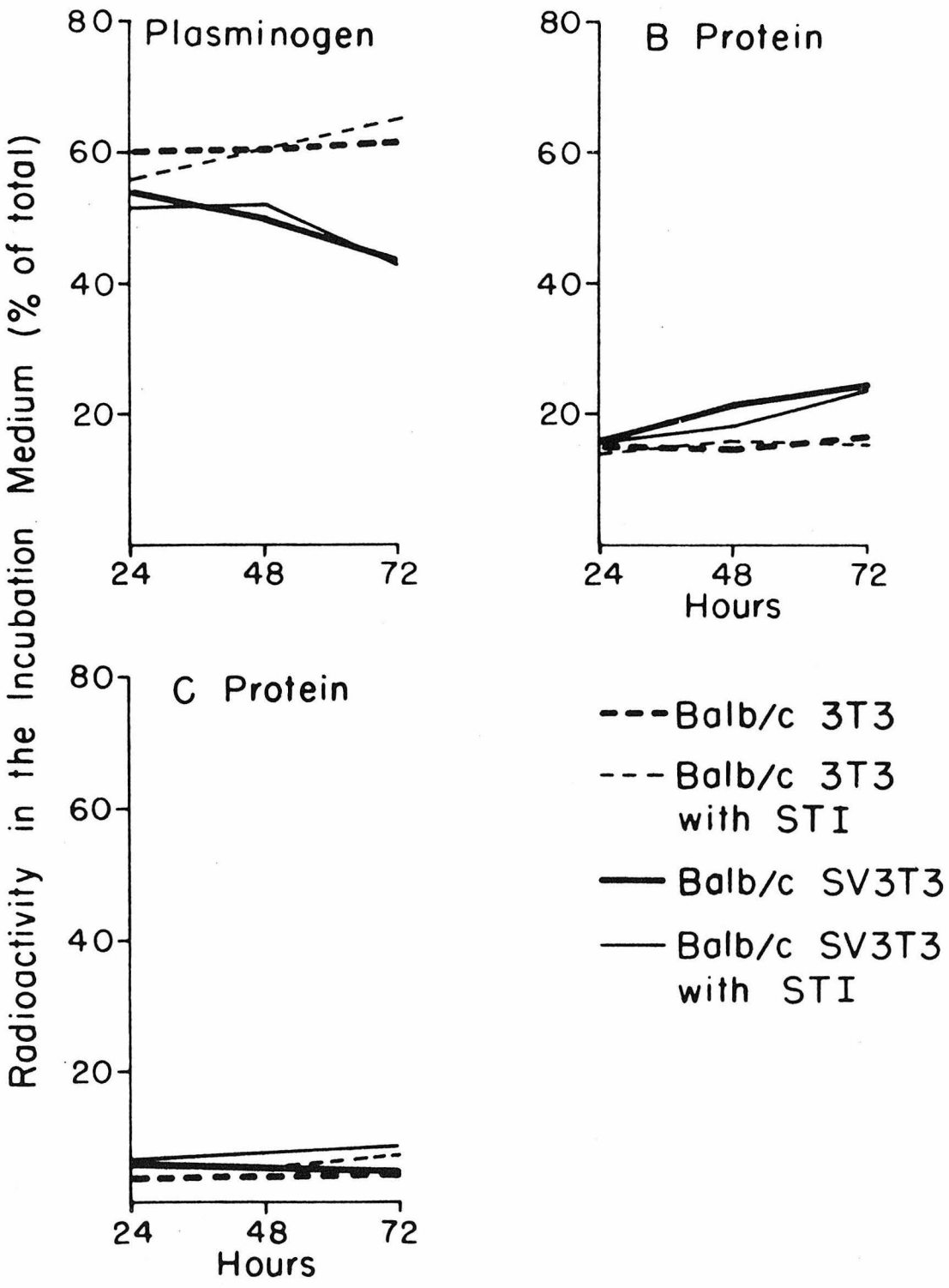
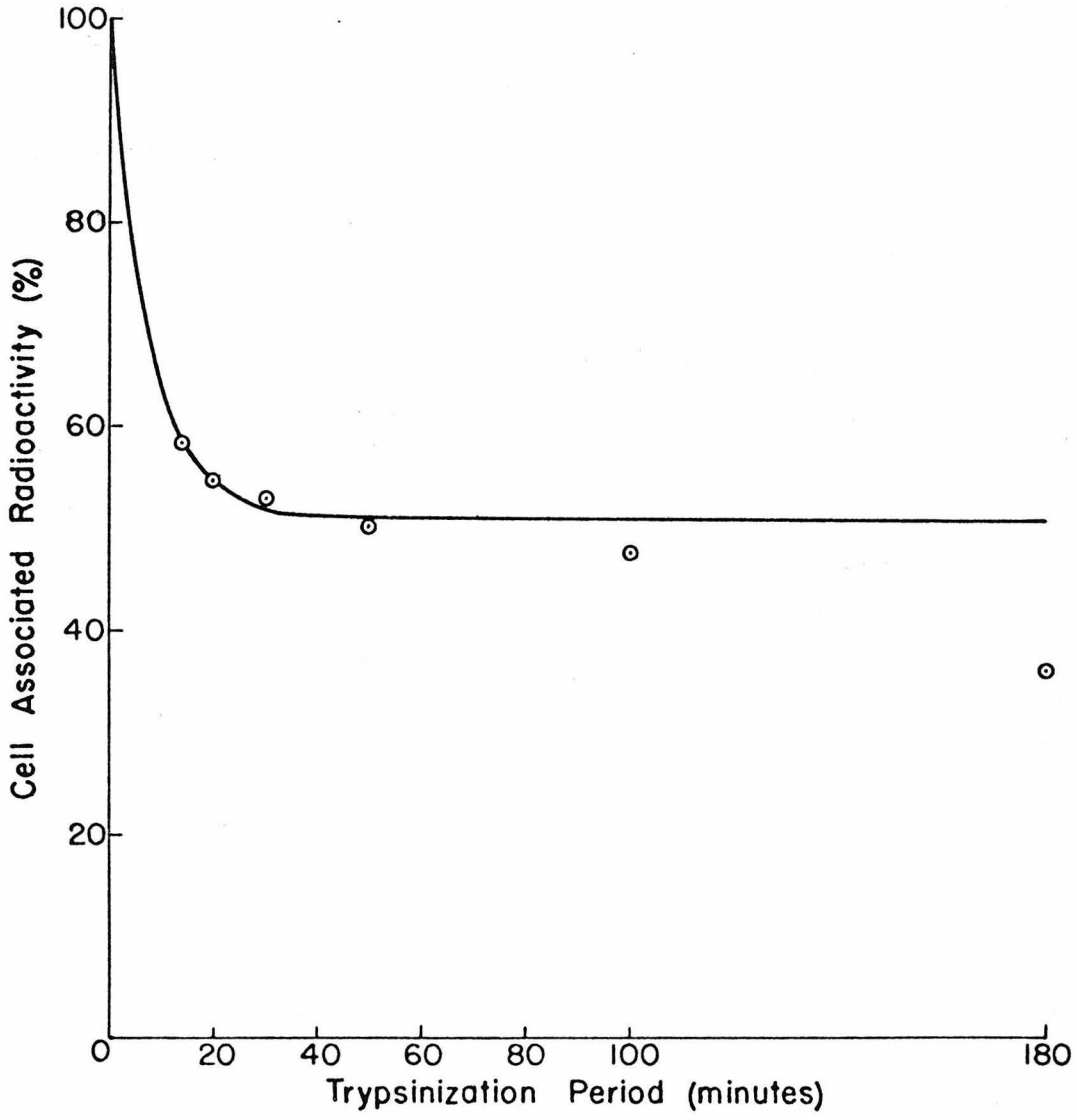

Plasminogen
B Protein
C Protein
80
60
40
20
24
48
72
Hours
Radioactivity in the Incubation Medium (% of total)
Balb/c 3T3
Balb/c 3T3 with STI
Balb/c SV3T3
Balb/c SV3T3 with STI

Figure 6


100
80
60
40
20
0
Cell Associated Radioactivity (%)
20
40
60
80
100
180
Trypsinization Period (minutes)

Abstracts of Propositions

Monty Krieger

I. The Function of Trypsin's Specificity Binding Pocket: The catalytic role of specific substrate side chains. In order to determine more precisely how substrate side chains function in tryptic catalysis, the hydrolysis of "fragmented" substrates should be studied. The substrates to be investigated should include substituted glycine amides and analides and the pH dependence of the reactions should be considered.

II. The Synthesis and Structure of a Complex Molecule: Trumbelene. A pathway for the synthesis of a highly strained hydrocarbon, trumbelene, is proposed along with a discussion of the electronic structure of this unusual, currently hypothetical, molecule.

III. Solution Complexes of DNA and Ionene: A simple assay system and its application. DNA forms highly-ordered, birefringent fiberous precipitates with a series of quaternary ammonium polymers called ionenes. A simple centrifugation assay for solution complexes of DNA and ionene is proposed and a direct application of the kinds of information provided by in vitro physical-chemical studies of DNA-polycation complexes is discussed.

IV. The Antileukemic Action of Glyoxal-bis(guanylhydrazones): Do these drugs act by inhibiting S-adenosyl methionine decarboxylases, DNA polymerases or trypsin-like enzymes? The antileukemic activities of glyoxal-bis(guanylhydrazones) and 4,4'-diacetyl-urea-bis(guanylhydrazone) may be due to the ability of these drugs to inhibit certain enzymes. A systematic survey of the inhibitory effects of these drugs is proposed.

V. Is Retinal Rivalry Dependent upon the Lateral Specialization of the Human Brain? Retinal rivalry appears to require high-level visual information processing. Experiments which could test for the involvement of lateral specialization in retinal rivalry using asymmetrically brain-damaged patients, commissuratomy patients, and normal subjects are proposed.