Title
Electrostatics in the catalytic mechanisms of trypsin

Author
Vasquez, John Ronald

Publication Date
1990

Peer reviewed|Thesis/dissertation

Electrostatics in the Catalytic Mechanism of Trypsin

by

John Ronald Vásquez

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHARMACEUTICAL CHEMISTRY

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

Date . . . . . University Librarian

Degree Conferred: 12/31/90

for Dad and Mom,

for Nadine,
Debbie,
Art,
Mark,
Valerie, &
Vanessa

for Catherine,

for Caroline,

for Linda,

for Mary,

for Charly,

but especially

for Catherine Louise Vásquez,
my most important love
She's put up with much more
than a little girl should have to bear.

Preface

I started to write a much longer preface, but it was turning into another thesis. I couldn't thank my friends and colleagues enough in another hundred pages. Suffice it to say, they've each helped me live so that nothing has been left unsaid. Science is a very human pursuit.

This work began as a collaborative effort between Charly Craik and Bob Fletterick. Their collaboration became Dr. Mary McGrath and myself. You can see how much of her work is presented here and judge for yourself how closely we've worked together. She's proven she could have done my part; I could not have done her's. Dr. Jeff Higaki and Dr. Luke Evnin gave me free access to their work and their friendship to help me pull off "my" half of this collaboration. Charly would tell me what to do. Jeff would tell me how to do it so that it worked. And Luke would tell me how to do it faster..... Or he'd tell me to do it faster. "Mañana is here everyday!" Daniel Wandres was there at the beginning, and I'd never have gotten started in molecular biology without him. Christopher Cilley has been here the rest of the time, and I'm not sure what part of this work is his. Dr. Sergio Pichuantes showed me how it's supposed to go when you start a post-doc. Dr. Guo Hong brought a little bit of the world to me and showed me some of the most exciting results I've ever seen on one gel. Dr. Lilia Babé brought a lot of new techniques into the lab and fought both of our tempers to teach me something new about friendship. Dr. Dianne DeCamp brought Ralph Brooker into the lab and thereby got the fluorimeter talking to the Macintosh. She's been great to talk to about science or anything else for that matter. Ann Eakin, Scott Willett, Chris Tsu, Jason Rosé, Dr. David Corey and Dr. John Perona: Charly has time and again proven his skill at finding talent. It's been a privilege working with them though we never intended to overlap so much. David has been particularly patient about moving into his new desk and bench with my stuff all over the place. I wish I had more time.

These people have been my family. My apology to them has been abbreviated: H. L. Mencken wrote, "If after I leave this vale, you chance to remember me, and choose to honor my ghost, forgive some sinner, wink at a homely girl." The men and women of the lab must have a powerful affection for some star as Mencken, for each of them have forgiven me and winked at me at each and every turn, no doubt wishing beyond hope that the storm would end soon. Each has had a peculiar patience with me, which I almost never appreciated at the time. Each has proven his or her friendship a thousand times over. I do not understand them. But my horrible memory reminds me often that I owe each of them nothing short of love and devotion. It is one good thing I have in abundance. I hope they know.

George Kenyon brought me into the department through Michael Marletta and sent me to Charly's lab for a project it's taken a good post doc, Dr. Lorenzo Chen, to accomplish. George has been phenomenal; I know how lucky I am have him as a mentor. Dan Santi has been a helpful teacher. I'm thankful for his reading of the manuscript. I'm also thankful to Peter Kollman and Jeanmarie Guenot for reading the electrostatics section and for helpful discussions.

The beginning of graduate school was so long ago that it's hard to remember. But I have a lot of friends from that time who helped get me here now. John Altman, Beth Gillece-Castro, Betsy Komives, Mark Watanabe, Lisa Peterson, Becky Redick: All doctors now, and a couple of professors. Caroline Decker and Linda Clayton who taught me a lot about life and helped me become me. They're still close friends.

In 1979, I gave the salutatory address at my high school graduation. I forgot to mention even *one* member of my family. I've worked eleven years to get a second chance to right that wrong. I felt pure joy to see so many members of my family with so many members of the lab at the commencement exercises this past June. The contributions of the people in the lab are better understood, and I've often taken the contributions of my family

for granted. So each of their names appears in the dedication. They were great to grow up with, great to live with (that isn't always so!), and harder to live without than I realized. They've each contributed money to this. They've been waiting a long time. We did it.

Many a morning, I rounded the corner to the lab, and I was thankful to Charly and George for the work they did to provide me a place to play and the toys to play with. Many nights I left in a hurry, so I could get to sleep, so I could get up, so I could get back to work. Mutant enzymes of my own design, substrate, a calibrated fluorimeter, and the stereo blasting something obscene or beautiful or both. That would have been enough. But I've had so much more here. It's a funny business that makes you pick up every five or six years and leave your friends. My time with them has been a rich and rewarding life. They taught me science and much more. Thankfully, this book doesn't mean that this life is over. It just means I'm just beginning to understand them better.

Electrostatics in the Catalytic Mechanism of Trypsin
by
John Ronald Vásquez

The internal Ser-214 of trypsin was mutated to Ala (A), Asp (D), Glu (E), Gln (Q), Lys (K), and Arg (R) to test the electrostatic basis for the 10^9x rate acceleration of trypsin-catalyzed reactions. Each protein was purified to homogeneity from *E. coli*. Isoelectric focusing indicates that the S214K and S214R trypsins have an additional positive charge relative to S214A, S214E, and S214Q trypsins which comigrate at a slightly more basic pI than trypsin. Steady-state kinetics of appropriate substrates were used to test the chemical reactivity of the enzymes while controlling for the effects of charge on the enzyme/substrate interaction and for structural perturbations caused by mutations. Only the positive charge mutations significantly reduced the chemical reactivity of trypsin. High pH reversibly inactivated S214E but not S214(Q,K, or S) trypsins suggesting the Glu-214 carboxylate deprotonates with a pK_a ~9. Moreover, an increase in K_m with decreasing pH suggests that Lys-214 disrupts the Michaelis complex at neutral pH while Gln-214 does so at acidic pH. The crystal structures of S214Q, S214E, and S214K trypsins determined by Dr. Mary McGrath show the mutant side chains folded within the protein without global distortion of the structures. Poisson-Boltzmann calculations based on these structures indicate that electrostatic interactions with protonated Lys-214 account for the loss of chemical activity and indicate that the protonation of Glu-214 and Gln-214 suggested by the pH-rate profiles is energetically accessible. Apparently, Asp-102 establishes a powerful electrostatic field within the protein which stabilizes positive charges and destabilizes negative charges at position 214. Thus, Asp-102 can stabilize positively charged His-57 in the transition states of trypsin-catalyzed reactions. The absence of bulk solvent enhances electrostatic interactions among side chains at positions 57, 102, and 214. Consistent with these results, double-mutation reconstruction experiments indicate that the chemical activity of the His-57:Ser-195 couple can be restored by a carboxylate in a

different orientation from Asp-102. Exclusion of bulk solvent from the Michaelis complex may be the fundamental basis for part of the rate acceleration seen in many enzyme-catalyzed reactions.

Table of Contents

List of Tables

List of Figures and Illustrations

CHAPTER 1

INTRODUCTION TO SERINE PROTEASE ENZYMOLOGY

THE STRUCTURAL BIOLOGY OF TRYPSIN

Trypsin catalyzes the hydrolysis of the carbonyl peptide bonds of lysyl and arginyl amino acid residues in proteins, a critical process in digestion. Trypsin is synthesized in the pancreas with a pre-zymogen signal peptide of approximately 15 amino acids,[47] but it is stored in secretory granules of pancreatic acinar cells and is secreted into the pleural space of the small intestine as the inactive precursor trypsinogen (MW ~25 kD).[88] Trypsinogen is converted in the duodenum to the catalytically active, monomeric trypsin (MW ~24 kD) upon the removal of the zymogen hexapeptide from the amino terminus by enteropeptidase (formerly "enterokinase")[84,88,130,142] or autocatalytically by trypsin.[132] The nascent amino terminal Ile-16 and Val-17 side chains (amino acid residues are numbered according to the analogous residues of chymotrypsinogen)[93] fit snugly into a hydrophobic pocket in the otherwise folded protein causing structural changes which activate the enzyme and direct the amino terminus to form a solvent-inaccessible salt bridge with Asp-194.[29,213,214]

Nature has developed four strategies for proteolysis, and the families of proteases are named for the unique chemical moiety which defines each strategy: aspartyl proteases,

metalloproteases, cysteine proteases, and serine proteases.[a,165] The serine proteases are further divided into the chymotrypsin and subtilisin subfamiles. These two subfamiles differ completely in primary, secondary, and tertiary structure except for the positions of the catalytic moieties in relation to the substrate amide bond undergoing hydrolysis.[175,245] In contrast, bacterial serine proteases are known with different primary sequences (Table 1)[107,160] which nevertheless fold in the same way as the mammalian proteases: two "Greek key [anti-parallel] β-barrels," as described by Jane Richardson,[171] which slightly twist around each other. Thus, significant portions of the polypeptide backbone in the core of trypsin, α-lytic protease,[72] and *Streptomyces griseus* trypsin[168] and proteases A[35] and B[51] are superimposable.

Table 1. *Amino acid sequence identity in the trypsin family of bacterial and eukaryotic serine proteases.* Streptomyces griseus *proteases A (SGPA), B (SGPB) and trypsin (SGT), α-lytic protease (α-LP), bovine cationic trypsin (BCT), chymotrypsin (CHYMO), and elastase (ELAST) Data from James, et al.*[107]

	SGPA	SGPB	α-LP	SGT	BCT	CHYMO	ELAST
SGPA	100%	61%	35%	18%	21%	21%	18%
SGPB	61%	100%	36%	18%	17%	18%	19%
α-LP	35%	36%	100%	14%	20%	18%	18%
SGT	18%	18%	14%	100%	33%	31%	34%
BCT	21%	17%	20%	33%	100%	45%	39%
CHYMO	21%	18%	18%	31%	45%	100%	41%
ELAST	18%	19%	18%	34%	39%	41%	100%

The primary sequence of *S. griseus* trypsin is more similar to the mammalian proteases (~30% of the side chains are identical) than to the bacterial proteases, *S. griseus* proteases A and B and α-lytic protease (~18% of the side chains are identical).[107] The mammalian subfamily is a set of genetically related proteins (~40% of the side chains are identical) including anionic (pI = 4.5) and cationic (pI = 8.5) trypsins (75% homologous to each other),[70] chymotrypsin, and elastase.[12,47,93,216] The proteases in the mammalian subfamily are differentiated by substrate specificity. These proteases characteristicly cleave

[a] Recent reports blur these distinctions. But the four basic families provide a context for understanding hybrids. Thus the uniqueness of a serine protease with a metalloprotease fold, serine carboxypeptidase,[137] is quickly appreciated.

the carbonyl peptide bonds of particular amino acids: chymotrypsin cleaves after large hydrophobic residues, elastase cleaves after small hydrophobic residues, and trypsin cleaves after basic amino acids (*see* Asp-189, p.•••). So characteristic is the substrate specificity of the proteases that *S. griseus* trypsin is named without question although it is chemically more similar to elastase than to trypsin (Table 1).

The primary determinants of the substrate specificity are located in the P_1 binding pocket which interacts only with the side chain of the substrate while structural determinants of the catalytic function of the serine proteases which interact with the amide bond to be hydrolyzed are identical (*see* DETERMINANTS OF THE CHEMISTRY OF TRYPSIN). Therefore, despite the differences in substrate specificity, the chemistry which gives rise to the catalytic function of the mammalian proteases is similar.[13]

CHEMISTRY IN THE ACTIVE SITE OF TRYPSIN

Trypsin has served as an archetypal enzyme since the late 19th century (references in ([130])) and still serves in the development of state-of-the-art biochemistry and biophysics,[48,49,126,203] as an example in textbooks,[68,231] and even in the popular press.[212,225] A large part of the success in the study of the serine proteases can be attributed to the similarities and differences among members of both chymotrypsin and subtilisin subfamilies and of the ready ability of the serine proteases to tolerate perturbations of the chemistry of substrates. In particular, the ability of the serine proteases to accommodate a variety of leaving groups has facilitated the elucidation of the chemical, kinetic, and catalytic mechanisms of these enzymes.

CHEMICAL REACTIONS CATALYZED BY TRYPSIN

Amidolysis

In vivo, trypsin hydrolyzes amide bonds in proteins (Figure 1).[88] The amide bond provides surprising chemical stability in light of the favorable free energy for hydrolysis of a peptide bond, $\Delta G \approx -2 - -3$ kcal/mol.[71] The spontaneous hydrolysis rate of the Phe-Gly

peptide bond has recently been determined at room temperature and neutral pH to be 3×10^{-9} s^{-1}, which corresponds to a half life of approximately 7 years.[115] Nevertheless, chymotrypsin hydrolyzes the Phe-Ala amide bond at ~3 s^{-1}, a rate acceleration of 10^9.[67] Similarly, the solvolysis rate for Suc-AAPF-*p*-nitroanilide was reported to be 10^{-8} of the rate of hydrolysis by subtilisin under identical conditions.[41] Similar rate accelerations can be seen in the active site of a mutant trypsin which still apparently works by the acyl-enzyme mechanism. The double mutant D102N:S195C trypsin hydrolyzed Z-GPR-AMC at a rate 10^{-8} of trypsin.[101] No direct measurement of the solvolysis rate was made for this substrate, but solvolysis was undetectable in the time course of this reaction.

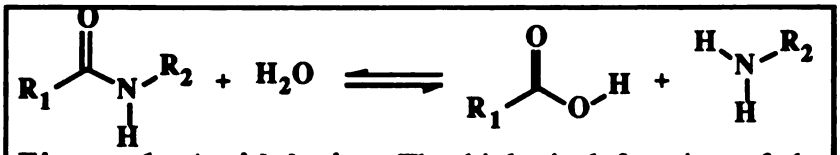


Figure 1. Amidolysis. The biological function of the serine proteases is to hydrolyze amide bonds in proteins. Here, R_1 and R_2 represent the α-carbons of two amino acid residues. For trypsin, R_1 is usually the α-carbon of arginine or lysine. R_2 may be the α-carbon of any amino acid (not proline), or in synthetic substrates, a masked chromophor or fluorophor or some other chemically useful moiety. $K_{eq} \approx 60$ in the above reaction.[71] The serine proteases *do* catalyze the back reaction which can be enhanced with dipolar aprotic solvents in aqueous solution[9] or upon methylation of the His-57 Nε.[240]

Esterolysis

Catalyzed hydrolysis of the isosteric but chemically disparate ester bond has provided an essential probe into the enzymatic mechanism of the serine proteases. Three aspects of esterolysis have aided the determination of the chemical function of the serine proteases. First, the ester bond is ~10^3 more reactive than the amide bond in solution[a,15,17] or ~10^2

Figure 2. Esterolysis. Ester bonds are isosteric but chemically more labile than amide bonds. As for amidolysis by trypsin, R_1 is usually the α-carbon of arginine or lysine while R_2 can be a fluorophor or chromophor. In addition, the carboxylic acid can be assayed nonspecifically as an acid.

[a] The hydrolysis rates for benzamide[15] and ethyl benzoate[17] have been determined in a well characterized system. For the hydroxide catalyzed reaction, benzamide hydrolyzes at 2.63×10^{-3} $M^{-1}s^{-1}$ at 109°C while ethyl benzoate hydrolyzes at 3×10^{-2} $M^{-1}s^{-1}$ at 25°C. Conservatively corrected for temperature

Footnote continued on the next page

more reactive in the active site of trypsin (tosyl- or benzoyl-arginine methyl ester vs amide)[192,202,221] or chymotrypsin (benzoyl-tyrosine ethyl ester vs amide).[118] Second, the leaving group can be a fluorophor such as 4-methylumbelliferone (Ex 360, Em 450) or a chromophor such as *p*-nitrophenol ($\Delta\varepsilon_{410}$ = 7,200) to facilitate monitoring the reaction. Third, the carboxylic acid ionizes to generate protons which can be assayed.

The flexibility of ester chemistry has been exploited in two ways. The fluorogenic ester substrate 4-methylumbelliferryl *p*-guanidinobenzoate (MUGB) reacts stoichiometrically with trypsin and so is a useful active site titrant (*see* The Kinetic Mechanism of the Serine Proteases [Ch. 1] *and* Active Site Titrations [Ch. 3]).[109] The chemical lability of the ester bond compensates for the steric incompatibility of the inhibitor side chain[140] which is necessary to make alkylation of the enzyme irreversible. In addition, coupled

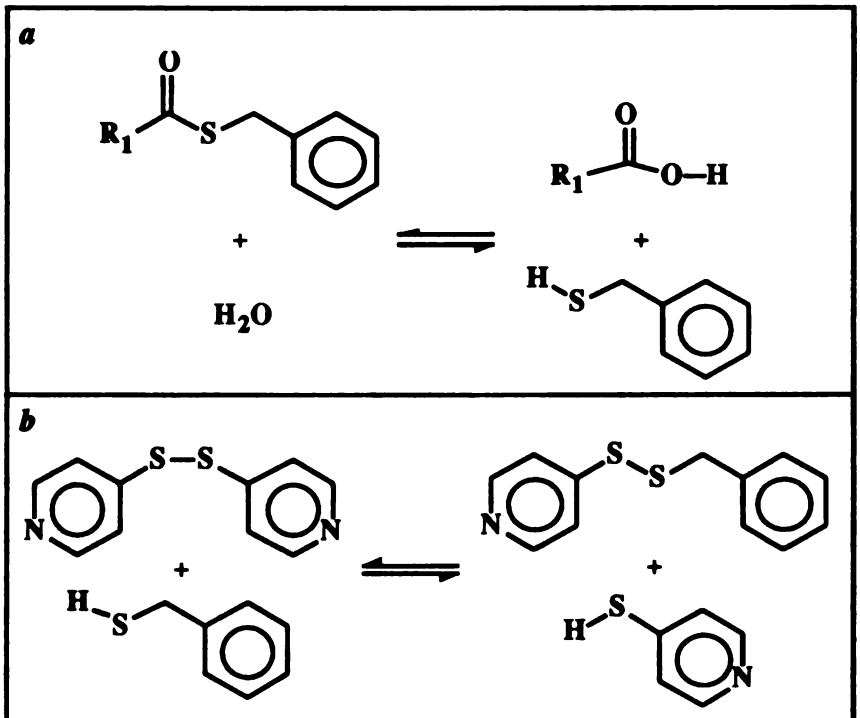


Figure 3. Thioesterolysis. ***a.*** The thiobenzyl ester hydrolyzes to yield a carboxylic acid and benzyl mercaptan. The thiobenzyl ester reacts with serine proteases at rates comparable to oxygen esters, but relative to esters, is stable to hydroxide. ***b.*** The thiobenzyl leaving group is coupled to 4,4'-dithiodipyridine to release chromophoric 4-thiopyridone. In combination with the chemical lability of the thioester bond, this system has a sensitivity comparable to fluorimetric assays. (see text for details and references)

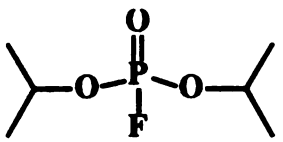
(double the rate for every ten degrees), the benzamide hydrolysis rate would be ~3 x 10^{-5} $M^{-1}s^{-1}$ at 25°C.

assays using thiobenzyl ester substrates[65] provide a sensitive assay for esterolysis.[43,92,151] The mercaptan leaving group is coupled to 4,4'-dithiodipyridine ($k_2 \approx 30{,}000\ M^{-1}\ s^{-1}$)[151] to release 4-thiopyridone, a chromophor with $\varepsilon_{340} = 19{,}800\ M^{-1}\ cm^{-1}$ (Figure 3).[43] As for esters, the leaving group had little effect on the observed rate of reaction.[151] Moreover, esters are normlly unstable to hydroxide but thioester substrates are reasonably stable at elevated pH (0.8% of thiobenzl ester is hydrolyzed per hour at pH 9.1).[65]

DETERMINANTS OF THE CHEMISTRY OF TRYPSIN

Ser-195

Phosphorylation



DIFP

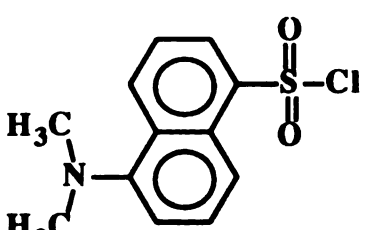
In addition to catalyzing hydrolysis reactions, the serine proteases proved to have two particularly reactive residues. Chymotrypsin was labeled stoichiometrically with ^{32}P-diisopropylfluorophosphate (DIFP).[110] The radioactive label was found incorporated in phosphoserine, and it was noted that although there were 22 serines in a molecule of chymotrypsin, apparently only one was labeled.[182] Similar results were noted for trypsin.[53] The zymogens trypsinogen and chymotrypsinogen were shown to react with DIFP at a rate four orders of magnitude lower than the rate for the mature enzymes indicating that the activation of the zymogen to the mature protease involved the chemical activation of the serine (Table 2).[155]

Table 2. *Second order rate constants for the reaction of serine proteases and zymogens with DIFP.*[155]

Protein	k_2 (M^{-1} min^{-1})
Chymotrypsin	2,700
Trypsin	300
Chymotrypsinogen	0.105
Trypsinogen	0.041

Sulfonylation

1-Dimethylaminonaphthalene 5-sulfonyl chloride sulfonylates the active site of chymotrypsin.[95,144] Chymotrypsin and trypsin were also found to be stoichiometrically reactive with a variety of sulfonyl fluorides to form sulfonyl-



1-Dimethylaminonaphthalene 5-sulfonyl chloride

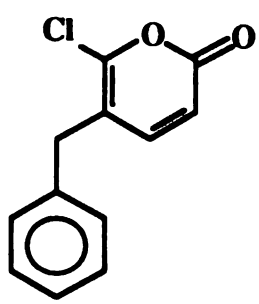
chymotrypsins and sulfonyl-trypsins[64] or with tosyl or pipsyl chloride to form tosyl- or pipsyl-chymotrypsin.[64,116] Despite the nonspecific chemical reactivity of tosyl and pipsyl chloride, again, one serine appeared to be particularly reactive.[64,117] It was noted that chymotrypsinogen was only weakly reactive toward tosyl chloride under the same conditions that labeled chymotrypsin.[117] Moreover, toluene sulfonic acid was eliminated from tosyl-chymotrypsin to form inactive anhydro-chymotrypsin which contained dehydroalanine-195.[117,215] Anhydro-chymotrypsin was inert towards chymotrypsin substrates but reacted with nucleophiles (dehydroalanine is a Michael acceptor) to yield, after total hydrolysis of the peptide bonds, the expected modified serines.[117,215] Both anhydrotrypsin[2] and anhydrochymotrypsin[2] were shown to bind specific proteinaceous inhibitors indicating that the proteins were not denatured.[1] Ultimately, the crystal structure of the bovine pancreatic trypsin inhibitor (BPTI)-anhydrotrypsin complex was determined.[104]

It was with some confidence therefore that Ser-195 was identified as the active site serine in electron density maps of tosyl-chymotrypsin derived from X-ray diffraction data.[145] Crystallographic results for the substrate analog inhibited indoleacryloyl-α-chymotrypsin demonstrated how substrate could interact with Ser-195.[99] The crystal structures of tosyl-elastase,[235] diisopropyl-trypsin,[213] and the BPTI-trypsin complex[176] have given unanimous indication of the similarities of the active sites of the serine proteases.

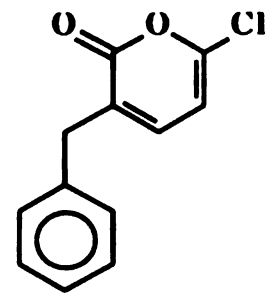
Suicide Inactivation

Serine proteases with different specificities catalyze a number of biologically important peptidase reactions. The chemistry in the active site of the serine proteases thus provides a target for rational drug-design based on the development of mechanism-based inhibitors. The design of mechanism-based inhibitors entails developing reagents which react by virtue of the chemical mechanism of the enzyme to unmask a reactive functional

group. The masked functional group is affixed to a substituent which confers specificity for the protease of interest. A number of serine protease mechanism-based inhibitors have been developed by the Abeles laboratory,[75,98,236,241] but two in particular have been exceptionally well characterized.



5-Benzyl-6-chloro-2-pyrone



3-Benzyl-6-chloro-2-pyrone

The chloropyrones were developed in an attempt to present a masked acyl chloride in the active site of chymotrypsin.[241] The chloropyrone ring is delivered to the active site of chymotrypsin by a benzyl group which binds into the substrate binding pocket. Crystallographic analysis of chymotrypsin inhibited with either 5-benzyl- or 3-benzyl-6-chloro-2-pyrone in conjunction with chemical analysis of the products of the inhibition reactions has revealed that the two chloropyrones inhibit chymotrypsin by different mechanisms. The 5-benzyl-6-chloropyrone binds into the active site of chymotrypsin to present C-6 to the nucleophilic Ser-195. Ser-195/chloropyrone ring system then undergoes an aromatic substitution reaction to alkylate Ser-195 with the loss of chloride ion (Figure 4).[173]

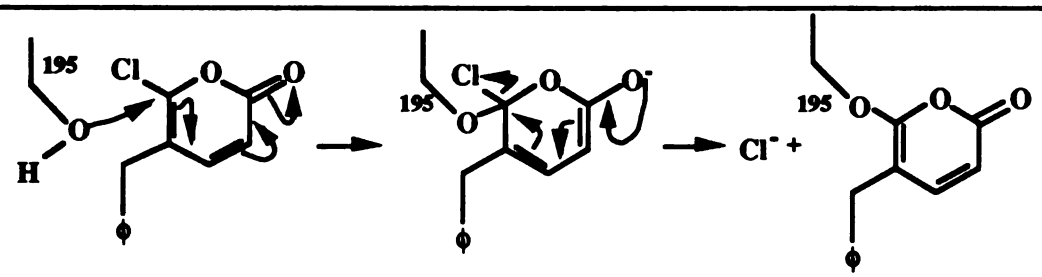


Figure 4. Chemical Mechanism of Inactivation of Chymotrypsin by 5-Benzyl-6-chloro-2-pyrone.

In contrast, 3-benzyl-6-chloro-2-pyrone binds in the active site of trypsin to present C-2 to the nucleophilic Ser-195. Attack by Ser-195 leads to ring opening and the generation of an acyl chloride which quickly hydrolyzes. The resulting carboxylate

evidently forms a salt bridge with the imidazolium of His-57 which protects the acyl-enzyme ester linkage from attack by water. The mechanism (Figure 5) is somewhat speculative, but the ring-opened product is observed in the crystal structure of inactivated chymotrypsin.[172]

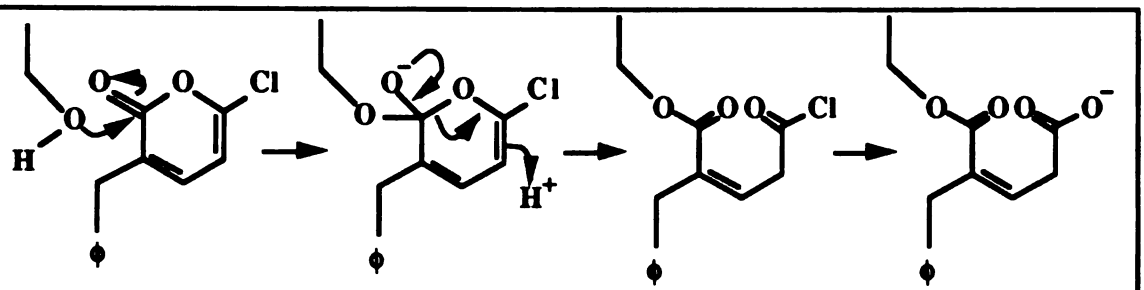


Figure 5. Chemical Mechanism of Inactivation of Chymotrypsin by 3-Benzyl-6-chloro-2-pyrone.

His-57

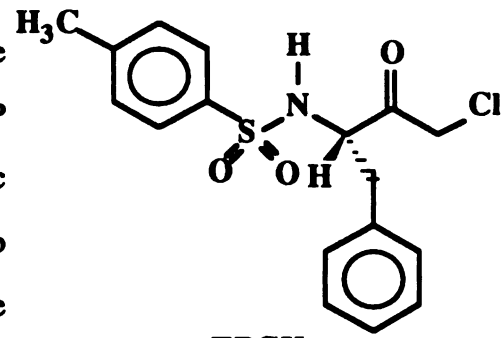
Photo-oxidation

Preliminary attempts to identify catalytically important residues included photooxidation of chymotrypsin in the presence of methylene blue.[237] Correlation of the extent of photooxidation (as released CO_2) with the percentage of destroyed histidine and tryptophan side chains and with the percentage loss of enzymatic activity implicated histidine as a catalytically important residue. Through active site titrations using *p*-nitrophenyl acetate, it was possible to show that destruction of histidine resulted in inactive enzyme while destruction of other residues resulted in partially inactive enzyme, further implicating histidine as the catalytically important residue.[124]

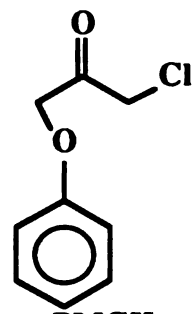
Substrate Specificity Directed Halomethyl Ketones

Shaw and coworkers exploited the specificity of chymotrypsin to present reactive functional groups in the active site of the enzyme. In the first round of experiments, a modified phenylalanine, tosylphenylalanyl chloromethylketone (TPCK), was used to deliver a chloromethylketone to the active site resulting in stoichiometric inactivation of

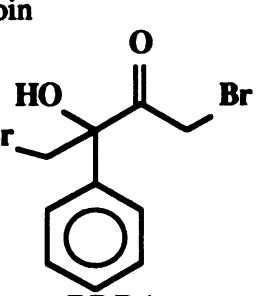
chymotrypsin with the loss of exactly one histidine from total protein hydrolysates; DIFP was shown to protect this histidine.[185] A peptic fragment of TPCK-chymotrypsin was found to have the same amino acid composition (minus one histidine) as a previously identified peptide[161] and was quickly mapped to His-57 and analogous peptides in elastase and trypsin.[198] Phenoxymethylchloromethyl ketone, PMCK, a TPCK analogue with no tosylamido moiety, was shown to alkylate chymotrypsin, but not chymotrypsinogen or trypsin, thus confirming the substrate specificity basis of the inactivation.[209] Two bromomethyl groups were presented to chymotrypsin on 2-phenyl-1,4-dibromoacetoin (PDBA), but only the bromomethylketone moiety reacts.[190] Analogously to TPCK, the chloromethylketone of lysine, TLCK, was shown to stoichiometrically inactivate trypsin with the loss of a single histidine which could be protected by the competitive inhibitor benzamidine.[194] Chymotrypsin is unaffected by TLCK and trypsin is unaffected by TPCK. Recent crystallographic results have revealed an epoxide in the active site of chymotrypsin inhibited with a chloro*ethyl*ketone. Thus, it appears that the inactivation of haloketones with serine proteases occurs by attack of Ser-195 on the carbonyl bond of the ketone, followed by elimination of chloride by the resulting oxygen anion to form the epoxide. The epoxide then reacts with the His-57 imidazole (Robert. H. Abeles, *personal communication*).



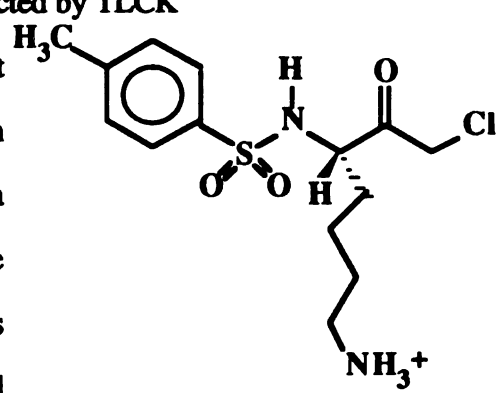
TPCK



PMCK



PDBA



TLCK

Asp-102

Asp-102 was first identified as a catalytic deteminant when a correction to the peptide sequence revealed that chymotrypsin contained a buried aspartate.[27] No chemical labeling of this residue has ever been reported. Mutation of Asp-102 to Asn-102 was called for in print as the reagents for the experiment became available.[68] The D102N mutant has 10^{-4} of the k_{cat} of trypsin with no observed effect on K_m;[49] however, the D102N mutation changes the tautomeric form of His-57.[203] The D32A mutant of subtilisin has a similar effect on k_{cat}, 3×10^{-5} of the k_{cat} of the appropriate mutant for comparison, but causes a 2x increase in K_m.[41] Thus, the Asp-102 carboxylate contributes at least 10^4 of the rate acceleration observed in the function of the serine proteases and is an important determinant of the chemistry in the active site of the serine proteases.

The Oxyanion Hole

As for Asp-102, the importance of the oxyanion hole was inferred from crystal structures. In addition to the catalytic triad of Asp, His, and Ser, subtilisin and chymotrypsin also convergently evolved a binding site, the "oxyanion hole," for the anion formed in the tetrahedral intermediates of the reactions catalyzed by serine proteases (Figure 7).[175] The oxyanion hole of the mammalian proteases is formed by the backbone amide protons of Gly-193 and Ser-195. The oxyanion hole of subtilisin is formed by the backbone amide proton of Ser-221 and a side chain amide proton of Asn-155. Phosphate or sulfonylate tetrahedral intermediate analogs were bound such that an oxygen atom is found in the oxyanion hole. Furthermore, the BPTI-anhydrotrypsin crystal structure features a non-planar peptide bond with the carbonyl oxygen in the oxyanion hole. This finding suggested that the oxyanion hole is sufficient to stabilize the carbonyl oxygen when the carbonyl bond adopts a non-planar geometry.[104] Testing the function of the oxyanion hole experimentally was difficult since the chemical inactivity of the amide groups negated chemical modification studies. However, site-directed mutagenesis of the Asn-155 side

chain in subtilisin has verified the importance of the oxyanion hole to the chemical activity of the serine proteses.[37,238]

Asp-189

A characteristic feature of trypsin is its specificity for the carbonyl bond of arginyl and lysyl amino acids. The first synthetic substrate for the pancreatic enzymes was α-hippuryl-lysine amide which was tested against a crude preparation of trypsin.[23] This observation justified the use of synthetic substrates in a preliminary survey of the substrate specificities of purified chymotrypsin and trypsin which had just become available.[131,132] Chymotrypsin was shown to hydrolyze amide bonds after tyrosine or phenylalanine residues but not after glycine, lysine, or glutamate, while trypsin was shown to react with lysine but not tyrosine amide bonds.[21] This result led to the finding that trypsin would also hydrolyze arginine amide bonds while chymotrypsin would not.[22,103] The esterase specificities of these proteins conformed to the amidase specificities, except that chymotrypsin was found to have some activity toward arginine methyl esters.[191,192] It was soon shown that benzoyl-arginine methyl, ethyl, isopropyl, cyclohexyl and benzyl esters were all hydrolyzed by trypsin at the same rate.[191] Thus, the substrate specificity of trypsin was determined to be for arginine or lysine substrates independent of the carbonyl chemistry or the leaving group.

The stereochemical basis for the observed substrate specificities of the three enzymes[196] was inferred from the similarity of amino acid sequences among the three proteases and the structural similarity between chymotrypsin[145] and elastase.[197] The three-dimensional structure of formyl-tryptophanyl-chymotrypsin indicated the position of the substrate binding pocket in space *and* in the protein sequence.[207] Hartley noted that elastase and chymotrypsin have Ser-189 at the bottom of the P_1 binding pocket while trypsin has Asp-189 and proposed the charge complementarity between the negative aspartic acid and the positive lysine and arginine side chains as the primary determinant of

the substrate specificity of trypsin.[196] This proposition was strengthened by the structure of trypsin.[213] The contributions to catalysis of the interactions in the substrate specificity pocket could be tested with modified substrates,[54,121] but modern enzymological techniques have only recently permitted forays into testing the subtleties of these specificities.[32,48,63,239] In the case of trypsin, in particular, it has been experimentally verified that trypsin specificity requires a negative charge at the bottom of the binding pocket.[63] However, although a general, theoretical description has been offered,[112] it remains to be elucidated exactly how the binding determinants of a given protease activate or fail to activate a given substrate for hydrolysis. Nevertheless, Asp-189 is a primary determinant of the chemical specificity of trypsin.

THE KINETIC MECHANISM OF TRYPSIN[68,231]

THE MICHAELIS-MENTEN MECHANISM

The kinetic mechanism of a single substrate enzyme which generates two products is adequately described by equation (1). This scheme indicates the presence of a

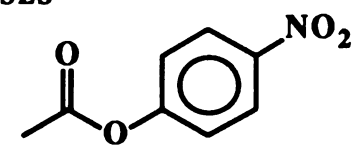
$$E + S \overset{K_s}{\leftrightarrow} ES \overset{k_{cat}}{\rightarrow} E + P1 + P2 \qquad (1)$$

noncovalently bound intermediate, the Michaelis complex ES, which forms when the substrate S adsorbs onto the enzyme E with some equilibrium constant of interaction K_s. The Michaelis complex then reacts at some limiting rate, k_{cat}, to regenerate free enzyme and form products, P1 and P2. The observed rate of reaction, v_{obs}, can be related to the concentrations of enzyme, [E], and substrate, [S], and intrinsic rate of formation of products, k_{cat}, by the Michaelis-Menten equation[154] (2) which is derived from the kinetic

$$v_{obs} = \frac{k_{cat}[E][S]}{K_m + [S]} \qquad (2)$$

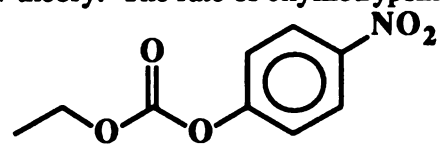
scheme in equation (1).[36] The constant, k_{cat}, is a characteristic parameter of the enzyme-catalyzed rate of reaction for a given substrate. However, it can be determined only when [E] is known. Because of variations in enzyme preparations and the conditions used for assay, the experimentally observed parameter $V_{max} = k_{cat}[E]$ is reported along with the variables (*e.g.*, temperature, ionic strength, and composition of the buffer, as well as the purity and preparation of the enzyme) which might affect it. V_{max} is the maximum observed initial rate of reaction. K_m is equal to [S] at which half of the maximum observed initial rate occurs. K_m is an intrinsic parameter of the enzyme-substrate interaction and may be determined without knowing the exact [E].

The Kinetic Mechanism of the Serine Proteases



p-Nitrophenyl Acetate

The Michaelis-Menten equation (2) describes the observed rate of product formation for a given set of conditions. It is desirable, however, to describe as many observations of the function of an enzyme as possible. And behavior of any system which cannot be explained by the current theory demands a new theory. The rate of chymotrypsin hydrolysis of *p*-nitrophenyl esters does not obey the Michaelis-Menten equation (2). The specific failure of the Michaelis-Menten model (1) and rate equation (2) is the prediction that products, P1 and P2, and free enzyme E are generated at the same time with the same rate k_{cat}. This rate is depicted by the dashed lines in Figure 6. However, the data (•) in Figure 6 show that *p*-nitrophenol is formed at different rates[a] early and late in the incubation. And, as shown in Figure 6, early in the course of the reaction, the rate of release of *p*-nitrophenol is faster than the rate of release late in the reaction. A new kinetic scheme, (3), was required to explain these observations.[94] It was postulated that the noncovalent Michaelis complex ES



p-Nitrophenyl Ethyl Carbonate

[a] Note that the *rate* is the slope that can be approximated by the line defined by two adjacent data points.

$$E + S \overset{K_s}{\leftrightarrow} ES \overset{k_2}{\rightarrow} ES' + P1 \overset{k_3}{\rightarrow} E + P2 \quad (3)$$

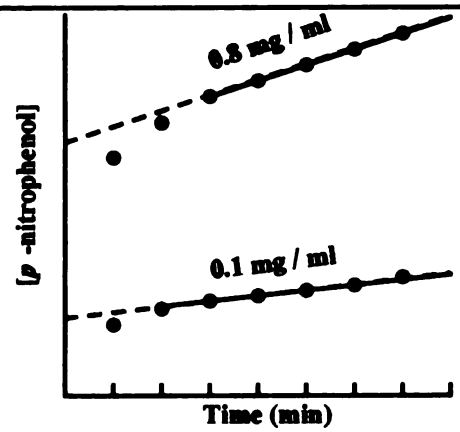


Figure 6. Amount of *p*-nitrophenol formed with time from the reaction of chymotrypsin with *p*-nitrophenyl ethyl carbonate.[94]

reacted quickly and stoichiometrically with a rate k_2 to release *p*-nitrophenol, **P1**, and form **ES'**, an enzyme intermediate. **ES'** then decomposed at a slower rate k_3 to release free enzyme **E** and **P2** (either acetate or ethyl carbonate depending on the substrate used). Consistent with this interpretation, the amount of *p*-nitrophenol formed during the early "burst" phase of the reaction was proportional to the amount of chymotrypsin used in the incubation; the y-intercepts of the two lines in Figure 6 are equal to the concentration of enzyme active sites in the incubation.[a]

Although the observed rate of reaction is always the rate of formation of *p*-nitrophenol, the rate of formation of *p*-nitrophenol is determined by factors which are either mitigated or which limit the observed rate of reaction. The initial phase of the reaction is limited by the concentration of substrate and enzyme (which determine the concentration and rate of formation of ES) and the intrinsic rate of the reaction of ES, k_2. The concentration of substrate was kept saturating and constant, and by varying the concentration of enzyme, the observed early rate could be accelerated (Figure 6). Using the stopped flow technique, the rate of decomposition of **ES'** was found to be equal to the observed rate of reaction. Thus, in the late phase of the reaction, the factor limiting the formation of *p*-nitrophenol was shown to be the concentration of free enzyme [E] which was determined by the rate of decomposition of the intermediate **ES'**.[87]

---

[a] A modern active site titration uses substrates which react with trypsin in the dead time of mixing to produce products which can be monitored fluorimetrically. (see Chapter 3, Active Site Titrations)

Despite the new model, the enzyme [E] and substrate [S] concentrations are still correlated with the observed rate of product formation, k_{cat}, by the Michaelis-Menten equation (2). However, the different limiting rates affecting v_{obs} are incorporated into the terms k_{cat} and K_m by equations (4) and (5) where k_2 is the rate of acylation, k_3 the rate

$$k_{cat} = \frac{k_2 k_3}{k_2 + k_3} \tag{4}$$

$$K_m = \frac{K_s k_3}{k_2 + k_3} \tag{5}$$

deacylation, and K_s is the dissociation constant for the Michaelis complex under rapid-equilibrium conditions, $k_{-1} >> k_2$.[246] For amide hydrolysis, $k_3 >> k_2$, and k_{cat} in equation (4) is a good approximation for the acylation rate, k_2, while K_m in equation (5) is a good approximation for K_s. For ester hydrolysis, $k_2 >> k_3$ and k_{cat} in equation (4) is a good approximation for the rate of deacylation, k_3, while K_m in equation (5) becomes $K_s k_3/k_2$. This kinetic scheme has proved ultimately useful and the verification of this scheme, the chemical description of the intermediates in this scheme, and the analysis of the chemical and kinetic perturbations of this scheme have proven to be didactic if not definitive enzymology.

THE CHEMICAL MECHANISM OF TRYPSIN[68,231]

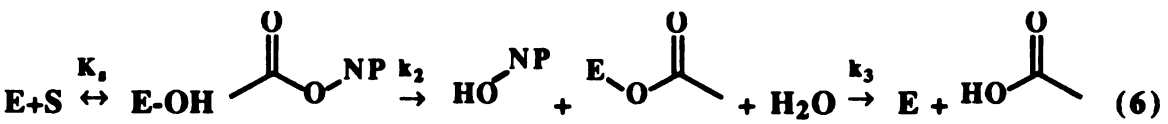
The chemical mechanism of an enzyme is a description of the chemical reactions and intermediates which lie on the reaction coordinate for the conversion of substrate to products. Fersht[68] adapted for the serine proteases Jencks'[111] three criteria for "proving" that a covalent intermediate is formed in an enzyme catalyzed reaction.

1. *The intermediate is isolated and characterized.*
2. *The intermediate is formed sufficiently rapidly to be on the reaction pathway.*
3. *The intermediate reacts sufficiently rapidly to be on the reaction pathway.*

The acyl-enzyme intermediates of the serine proteases have been isolated and identified, and the rates of formation and decomposition have been determined.

As discussed above, chemical labeling had indicated that Ser-195 was uniquely reactive.[182] Given the chemistry of carbonyl compounds and having detected a kinetic intermediate, Hartley and Kirby proposed the acyl-enzyme mechanism (6).[94]

$$E+S \overset{K_s}{\leftrightarrow} ES \overset{k_2}{\rightarrow} P1 + ES' \overset{k_3}{\rightarrow} E + P2$$



(6)

The kinetic intermediates can be identified according to this mechanism for the hydrolysis of *p*-nitrophenyl acetate or *p*-nitrophenyl ethylcarbonate: **ES'** is acyl-enzyme, **P1** is *p*-nitrophenol, and **P2** is acetic acid or ethylcarbonic acid. Although derived from the analysis of the hydrolysis of nonspecific substrates, the mechanism in equation **(6)** provided a theoretical foundation for work using specific substrates.

KINETIC DETECTION OF INTERMEDIATES

Determination of the Microscopic Rate of Acylation

Further work on the nonspecific substrate, *p*-nitrophenyl acetate, resulted in the development of two methods for the study of serine protease mechanism. The stopped flow technique uses a device to mix reagents rapidly in a flow-through spectrophotometer cell. The flow is stopped suddenly at the same time an oscilloscope is triggered to record data.[76,77] In the quenched flow technique, the reagents are mixed rapidly and passed through a capillary of known length into a quenching solution (thus the reaction has been allowed to proceed for a short, known amount of time).[77]

Using these rapid sampling devices, the rate of formation of acyl-enzyme can be monitored directly in three ways: the release of colored products upon acylation,[87] the formation of a colored acyl-enzyme,[25] and the displacement of a colored inhibitor.[24] The rate of specific ester binding and acylation (trypsin with N-benzoyl-arginine ethyl ester $k_2 \approx 175\ s^{-1}$; chymotrypsin with N-benzoyl-tryosine ethyl ester $k_2 \approx 600\ s^{-1}$) was faster than the overall reaction, but the microscopic rate of acylation for specific amides on both trypsin (N-benzoyl-arginine amide $k_2 \approx 0.7\ s^{-1}$) and chymotrypsin (N-benzoyl-tyrosine amide $k_2 \approx 0.24\ s^{-1}$) was the same as the overall rate.[24,85,86]

Determination of the Microscopic Rate of Deacylation

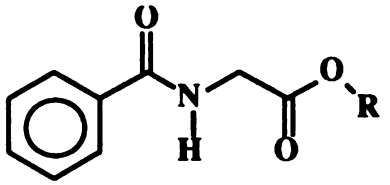
Indirect evidence for the existence of the acyl-enyme intermediate was obtained from the steady-state hydrolysis of various esters which would form the same acyl enzyme. Five different labile cinnamates were shown to acylate chymotrypsin at rates faster than the overall rate depending on the chemistry of the leaving group ($k_2 \approx 300\ s^{-1}$),[18] yet the chymotrypsin-catalyzed hydrolysis of these cinnamates occurs at the same, slower rate ($k_3 \approx 13\ s^{-1}$) indicating the decomposition of a common intermediate.[14,19] Similar results were obtained for the hydrolysis of a number of hippurate esters by chymotrypsin.[59] In addition, acyl-chymotrypsins with various electron donating and withdrawing groups were found to deacylate with rates which correlated with the chemistry of the acyl group.[39] The rate of decomposition of the acyl-enzyme could be followed directly by monitoring the loss of the spectroscopic signal due to the acyl-enzyme chromophor.[119] Ultimately, steady state k_{cat}'s for specific ester substrates were correlated with the microscopic deacylation rates (chymotrypsin with N-acetyl-tryptophan esters $k_{cat} = k_3 = 28\ s^{-1}$; with N-acetyl-phenylalanine esters $k_{cat} = k_3 = 72\ s^{-1}$).[247]

Chemical Detection of Intermediates

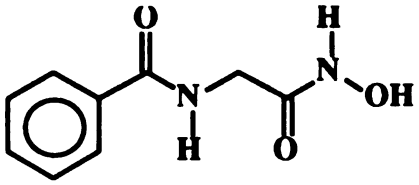
Partitioning of Intermediates between Competing Acceptors[67]

The analytical identification of chemicals has often been aided by the reaction of the unknown compound with a known reagent. The products of the reaction are characterized

and the nature of the products interpreted to give an indication of the chemical structures of the unknown. Thus, the acyl-enzyme has been scavenged using various reagents to generate free enzyme and the acylated scavenging agent. Hippuryl-chymotrypsins prepared from various hippurate esters would react with hydroxylamine, NH_2OH, to form hippurylhydroxamic acid and chymotrypsin; moreover, despite the various leaving groups (methyl, ethyl, isopropyl, *etc.*, hippurate esters were tested), the same fraction of the ester was converted to the hydroxamate. The conversion of the same fraction of the ester to the hydroxamate indicated that the scavenging reagent was reacting with a common intermediate formed from the reaction of the various esters with chymotrypsin *and* that the scavenging reagent reacted with this intermediate at some characteristic rate in competition with water. The chemical nature of the products indicated that the intermediate was the acyl-enzyme.[59,60] Specific substrates for chymotrypsin were used to form acyl enzymes which were scavenged by methanol or hydroxylamine[16] or by alanine, glycine, or hydrazine[67] to give identical results. The amount of scavenged acyl-enzyme was independent of the acylating agent, and a particular scavenging agent reacted with a given acyl-enzyme at a characteristic rate.[67]



Hippurate Ester



Hippuryl Hydroxamic Acid

Physical Analysis of Acyl-enzyme Intermediates

Characterization of specific substrate acyl-enzymes has been surprisingly successful in light of the labile nature of these compounds. Acetyltryptophanyl chymotrypsin was trapped at low pH (pH 2 – 4) and characterized spectroscopically and kinetically.[119] Low temperature was used to trap the acyl-enzyme of elastase with a specific substrate which even allowed crystallographic characterization.[3] Both temperature and pH had to be used to trap an acyl-trypsin intermediate which was characterized by ^{13}C-NMR.[139] However,

by methylating the His-57 imidazole,[158] it has been possible to observe this NMR signal at room temperature.[240]

The crystal structures of substrate-enzyme analogs have provided an indication of the possible interactions of substrates with the enzyme at almost every definable position on the reaction coordinate. These structures invariably have Ser-195 coordinated or acylated, with the substrate side chain located in the binding pocket as expected. These structures are sterically or chemically stabilized so that they survive long enough for x-ray diffraction data to be collected. No deviation from this theme has ever been reported.[a,8] Indoleacryloyl-chymotrypsin provided the first model for a serine protease acyl-enzyme structure[99,207] which was recently complemented by the structure of guanidinobenzoyl-trypsin.[140] The complex of bovine pancreatic trypsin inhibitor-trypsin (BPTI-trypsin)[105,176] provides a model for the Michaelis complex of substrate-lysine:trypsin interactions while K15R-BPTI-trypsinogen[30] provides a model for the Michaelis complex of arginine-substrate trypsin. The complexes of acyl-prolylalanylprolylphenylalanyl keto-*Streptomyces griseus* protease A,[108] amidinophenylpyruvyl-trypsin,[143] N-butyloxycarbonylalanylprolylvalyl borono-α-lytic protease,[31] and acylphenylalanyl and acylleucylphenylalanyl trifluoromethylketochymotrypsin[33] offer structural models for the enzyme tetrahedral intermediate on the pathway to formation of the acyl-enzyme. Recently, cinnamoyl-chymotrypsins, which provide a model for chymotrypsin acyl-enzyme, have been used in a study offering exciting prospects.[211] These compounds can be photoactivated in the crystalline form as well as in solution to trigger deacylation. These results directly indicate the similarity of the crystalline and solution chemistry of these compounds and offer the possibility for use in Laue diffraction experiments.[210]

---

[a] Recent NMR data indicates that nonspecific boroester substrates may form complexes with a bond between B and N_ε of His-57.[8]

The crystal structures of substrate-enzyme analogues provide such a wealth of information that they are frequently misconstrued as true intermediate and even transition state structures. Such misguided interpretation has been so frustrating that this extraordinary body of work has been dismissed, somewhat accurately, if too succinctly, as "interesting, but mechanistically irrelevant."[188] The crystallographic and spectroscopic characterization of substrate-enzyme analogs *can* provide unambiguous identification of allowed structures, steric chemistry, and bonding[7] which can be observed nowhere else, but they do not definitively identify the forces or bonding in the *actual* transition state. To be sure, a single crystal structure is almost unmatched in information content by any other result, but until chemistry and physics are closed bodies of knowledge, the interpretation of crystal structures will beg for functional information.

For example, the crystal structures of tosyl-chymotrypsin,[145] tosyl-elastase,[197] and DIP-trypsin[213] provided the all important stereochemical structures of these molecules. The similarities of the three pancreatic enzyme structures and the consistent active site labeling correlated a great deal of chemical modification and kinetic data. It should be remembered, however, that these structures were solved based on a great deal of what would now be called "old fashioned" biochemistry. And it was the seemingly trivial chemical result that only *one* of 22 serines could be labeled[182] and protected by competitive inhibitors[116] that strengthened the conclusion that the first of these structures had actually pinpointed the active site serine. In turn, crystal structures have suggested a mechanism for how this particular serine was chemically activated to attack carbonyl bonds by the imidazole of His-57[27] and the amide protons of the oxyanion hole.[175] The *sum* of the results from disparate techniques has provided a powerful model for describing the kinetic, chemical, and physical properties of every intermediate on the reaction coordinate of serine protease-catalyzed reactions. The structures of analogs have revealed possible bonding interactions for the intermediates in the chemical mechanism, but the acyl-enzyme

mechanism in Figure 7 has been proven for peptides, amides, and esters according to the rules listed by Fersht.

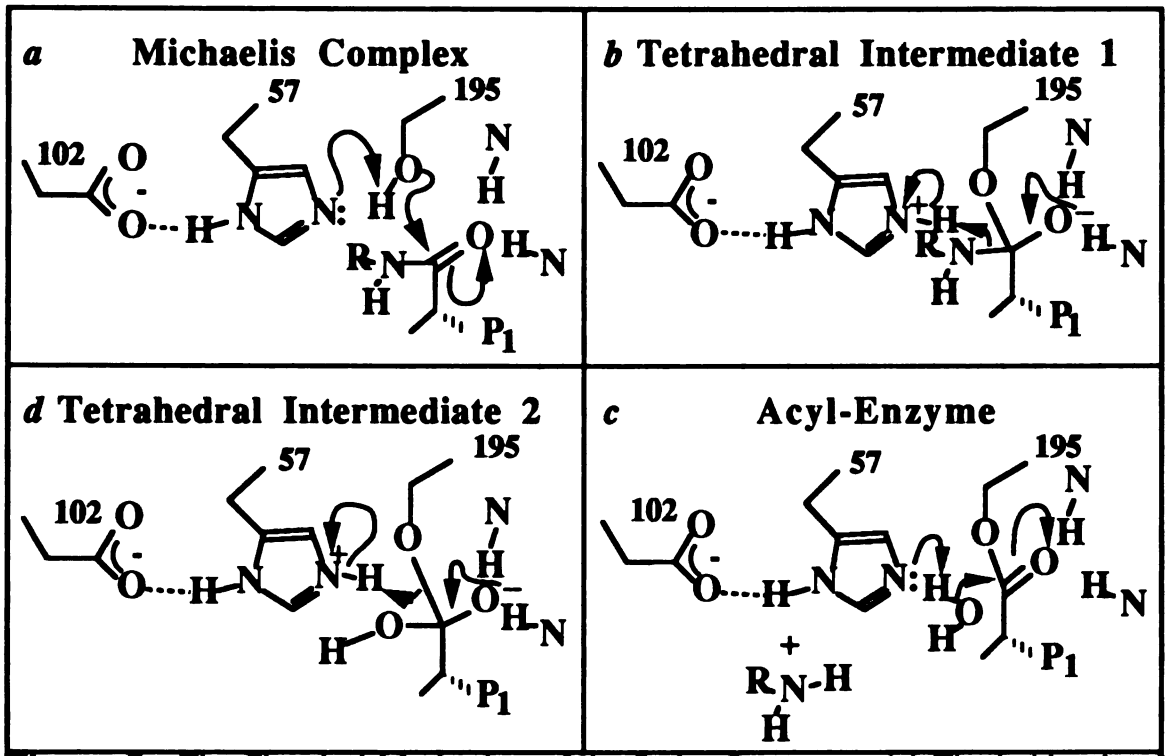


Figure 7. The Chemical Mechanism of the Serine Proteases. (*a*) The imidazole of His-57 acts as a general base to activate the hydroxyl group on Ser-195 for nucleophilic attack on the amide bond of the substrate to form the tetrahedral intermediate. (*b*) The imidazolium then acts as a general acid to catalyze the collapse of the tetrahedral intermediate to form the acyl-enzyme intermediate. (*c*) The imidazole of His-57 then activates a water molecule to attack the acyl-enzyme carbonyl bond to form the tetrahedral intermediate for deacylation. (*d*) The imidazolium again acts as a general acid to catalyze the collapse of this tetrahedral intermediate to form the carboxylic acid which diffuses off of the enzyme (not shown). **R** may be a fluorophor as is used in these studies. P_1 denotes the substrate residue in the S_1 subsite on the protease. P_1 would be lysine or arginine for trypsin, tyrosine or phenylalanine for chymotrypsin.

Free Energy Diagrams for Serine Protease-Catalyzed Reactions

The microscopic rates of serine protease-catalyzed reactions can be determined by the method of added nucleophiles as developed by Bender.[16] The added nucleophile is the P1 product of the reaction of the substrate with the protease, *e. g.*, methanol is added to methyl ester hydrolysis reactions and alanine amide is added to AcPhe-AlaNH_2 hydrolysis reactions. By varying the concentration of the added nucelophile, the rates of the back

reactions can be determined. The reaction coordinate diagrams in Figure 8 were derived from such data.[67] Similar results have been reported for specific esterolysis by chymotrypsin.[54] Because Gibbs' free energy (G) is plotted, the rate determining step of each reaction is the highest barrier from trough to peak, *i. e.*, the greatest $\Delta G^{\ddagger}$.[a] Therefore, it can be seen that the rate determining step for amides is acylation while the rate determining step for esters is deacylation. The barrier for formation of the Michaelis complex is approximately the same for esters and amides; the difference in free energy between the free states (or for the Michaelis complexes) for amides and esters can be attributed to the difference in chemical lability of the ester and amide bonds. Similarly, the difference in product energies for amides and esters are due to the protonation of amide products.[67]

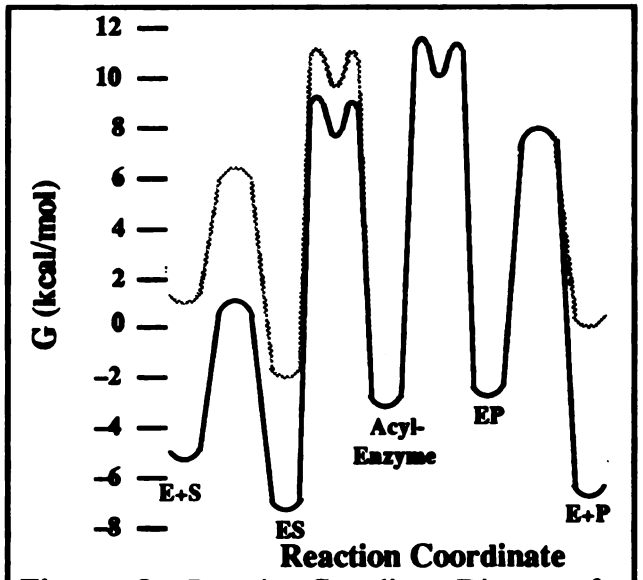


Figure 8. Reaction Coordinate Diagrams for Chymotrypsin Hydrolysis of AcPhe-AlaNH$_2$ (–) and AcPhe-OMe (--). Data from Fastrez and Fersht.[67]

THE CATALYTIC MECHANISM OF TRYPSIN

The free energy diagram derived from the chemical mechanism reveals that the catalyzed reaction has lower activation energies ($\Delta G^{\ddagger} \approx 16$ kcal/mol)[67] than the uncatalyzed reaction ($\Delta G^{\ddagger} \approx 29$ kcal/mol).[115] The catalytic mechanism is a description of exactly *how* the activation energies are lowered in the catalyzed reaction. Given the tenets of transition state theory,[127,136,138] this description may involve the definition of the forces which stabilize the transition state relative to the ground state of the reaction.[113,162,243]

---

[a] Had potential energy been plotted, the absolute highest barrier would be the rate determining step.

THE GHOST IN THE MACHINE

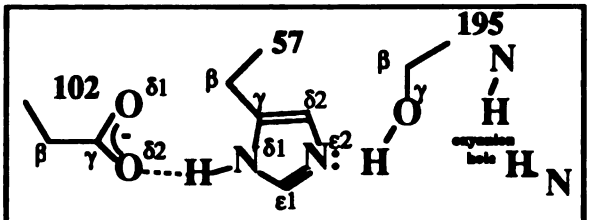


Figure 9. Definition of atom positions in the catalytic triad.

Before describing the physically definable mechanisms for transition state stabilization, there are three tantalizing results which hint at a "programming" of the enzyme for catalysis. First, an analysis of the structures of BPTI inhibited proteases and other structures[28] has hinted at a side chain motion which may be catalytically important.[7,28,125] In most structures, the His-57 Nε forms a hydrogen bond with a water molecule sitting at the position of the leaving group amide. In the BPTI-inhibited proteases, the His-57 Nε forms a hydrogen bond with the Ser-195 Oγ suggesting that substrate binding shifts the imidazole to a position which will activate the Ser-195 Oγ for nucleophilic attack. Second, consistent with the moving histidine mechanism, NMR data indicate that the His-57 Nδ forms a hydrogen bond with Asp-102 Oδ2 in the ground state, but that this hydrogen bond is lost in tetrahedral inhibitor complexes.[7,8] Third, the structure of the BPTI-anhydrotrypsin complex reveals that the amide bond which would be attacked has adopted a pyramidal structure even in the absence of the nucleophilic hydroxyl. This result was essentially interpreted as indicating that the oxyanion hole is able to stabilize the deformation of the carbonyl double bond by the structural interactions between trypsin and the inhibitor.[104] Thus, subtle motions of catalytic residues and subtle changes in the protein structure may contribute to the catalytic function of the enzyme in ways that will be difficult to study directly.

CHARGE-RELAY

The chemical mechanism of trypsin presented in Figure 7 requires the formation of charges. In some schools of thought, charges do not normally approach so closely without neutralizing. It was noted early in the mechanistic proposals for the structure of the serine proteases that these charges could easily be neutralized by shifting the proton from the

His-57 imidazolium to the Asp-102 carboxylate to yield electrostatically neutral carboxylic acid and imidazole.[27] In addition, the enormous acceleration of amide hydrolysis rates by serine proteases seemed to suggest that something other than simple general base catalysis was at work. The notion that the Asp-102:His-57 couple appears to operate as a general super-base system which is capable of relieving the generation of charge in the transition state is intellectually satisfying if theoretically deficient.

Experimental Evidence

One major aspect of the charge relay hypothesis is that there is direct experimental evidence for its existence. The proton inventory technique is an experimental method with well-developed theoretical underpinnings[230] for counting protons in the transition states of chemical reactions. In practice, the kinetics of a reaction are determined in water which contains various mole fractions of D_2O. A curve is fit to all of the data based on a model which accounts for the number of protons moving in the transition state and the fraction of those which are deuterons.[230]

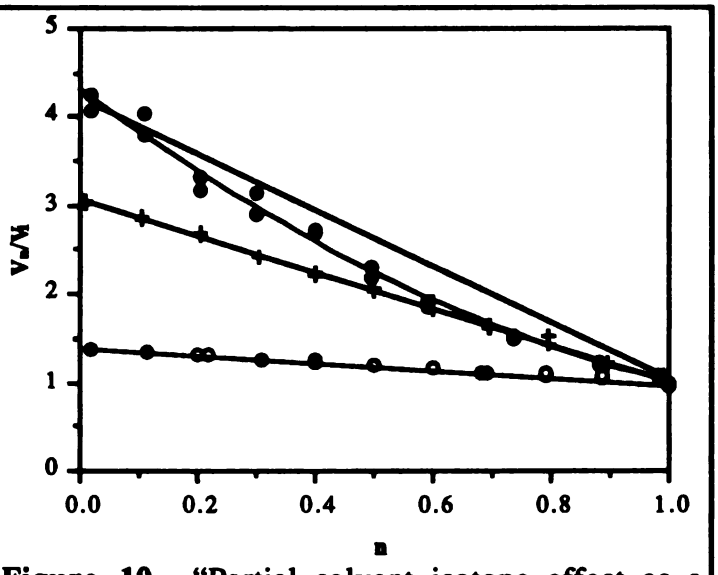


Figure 10. "Partial solvent isotope effect as a function of atom fraction of deuterium *n* for the trypsin catalyzed hydrolysis of" *p*-nitrophenyl acetate (o), N-benzoyl arginine ethyl ester (+), and N-benzoyl-phe-val-arg p-nitroanilide (•). Straight lines are consistent with models with one proton in the transition state; the polynomial fit is consistent with a model with two protons in the transition state.[58]

Two major problems arise in the interpretation of such results. First, although it is possible to fit a model requiring two protons, for example, to such a curve, it is not possible to *assign* those protons to any particular two that might be involved in the chemical changes in the active site during catalysis. Moreover, a given curve may be fit with

different models, so the "inventory" can be somewhat inaccurate.[4,128] In particular, it has been pointed out that deuterium exchange at a subset as small as 20 of the 400 exchangeable protons of a serine protease may reduce the number of protons observed via an inverse isotope effect or increase the number of protons observed by secondary isotope effects to yield plots indistinguishable from those in Figure 10.[128] Therefore, interpretations indicating "full functioning of the catalytic triad"[206] or linking secondary site interactions[204] or P_1 interactions[205] to charge relay are unjustified. Indeed, the rate of hydrolysis of N-acetyl-tyrosine ethyl ester chymotrypsin with a methylated His-57 imidazole has been shown to respond to various mole fractions of D_2O according to a model invoking multiple protons in the transition state.[186] This result does not show that the proton inventory results are not detecting charge relay; however, it does exemplify the problems with proton inventory studies noted above.

Electrostatic Transition State Stabilization

In contrast to the situation with charge relay, a well-developed theory for electrostatic transition-state stabilization by serine proteases has evolved in the absence of direct experimental data.[201,233] This thesis purports to present data supporting the electrostatic transition state stabilization theory, and so a discussion of electrostatic theory is presented in Chapter 2.

Direct experimental evidence for catalysis by electrostatic forces in the serine proteases is rare to nonexistent. The primary cause of this shortage of data is the screening effect of bulk water (Chapter 2). To date, most electrostatic mutations have been made on the surface of proteins and have been observed to perturb the pK_a of the imidazolium of His-57 and to affect the interactions with charged substrates.[177,178,220,239]

Electrostatic stabilization of the transition states of serine protease-catalyzed reactions has been the subject of extensive theoretical study.[122,201,224,233] However, while theoretical calculations may demonstrate which effects are sufficient for catalysis, they do

not determine which effects are necessarily at work in catalysis. In a complementary fashion, mutagenesis experiments have usually involved removing catalytic groups and so have revealed the magnitude of the effect a given substituent has on catalysis, but not the nature of the effect.[37,41,49,203,238] Transition state analogues reveal the possible behavior of a given constellation of amino acid residues in a protein but do not represent the actual transition state of interest. Nevertheless, the analysis of such complexes has provided an indication of the structure[31] of such complexes and the bonding[7,8,66,126,200] interactions which are possible. In particular, the NMR spectroscopic analysis of the boronic acid complex of α-lytic protease has given direct experimental evidence for the strength of the ionic interactions in the active site of the serine proteases. Boron becomes negatively charged upon accepting a fourth ligand and positioned next to the imidazole of His-57, shielded from bulk solvent, anionic boron apparently stabilizes the imidazolium ion up to and beyond pH 10.5 (a pK_a shift >3.5 units).[8] Thus the magnitudes of such interactions in the active site of a serine protease are directly indicated but not for the actual transition state of a reaction. The discovery that charges could be buried in trypsin has provided a probe into the electrostatic nature of the transition state through the crystallographically defined, low-dielectric medium of the protein.

MUTAGENESIS OF THE SERINE PROTEASES

Being such well-characterized, archetypal enzymes, almost all new techniques and theories, including site-directed mutagenesis, have been tested on the serine proteases. The first site-directed mutagenesis experiments on the serine proteases were conservative changes exploring substrate specificity[48] and catalytic function.[220]

SUBSTRATE SPECIFICITY OF THE SERINE PROTEASES

The crystal structures of BPTI-trypsin[176] and K15R-BPTI-trypsinogen[30] indicated that the ζ-amino group of a lysine substrate interacted with the protein through bound water molecules while the guanidinium moiety of an arginine substrate interacted directly with

Asp-189 at the bottom of the binding pocket. It was reasoned that the specificity for lysine and arginine substrates might be altered by disrupting the positions of the bound water molecules or by blocking the direct interaction with Asp-189 by introducing methyl groups at appropriate positions. These surface methyl groups were introduced in the binding pocket by mutating glycine residues to alanine which would presumably introduce a minimal perturbation of the structure. The anticipated shift in specificity was observed but was mediated in k_{cat} (again hinting at the ghost in the machine). In a more drastic experiment, the attempted electrostatic charge reversal of substrate specificity by replacing Asp-189 at the bottom of the substrate-binding pocket with Lys-189 resulted in specificity not for negatively charged substrates, but rather for hydrophobic substrates.[84] In a refreshingly understandable result, genetic selection for trypsin specificity in a library of randomized binding pocket mutants generated various combinations of amino acids in the bottom of the binding pocket which invariably included a negatively charged residue in the most active isolates.[63]

Alteration of the *specificity* of a serine protease based upon engineering the microscopic interactions between substrates and proteins has not been accomplished; rather, such experiments have generally resulted in the loss of specificity. Subtilisin BPN' normally has little affinity for anionic substrates, and so, the change of net charge at two residues in the P_1 binding site from –1 to +1 resulted in a 10^3 fold increase in k_{cat}/K_m for glutamate substrates while the corresponding change of k_{cat}/K_m for glutamine, methionine, and lysine tetrapeptide substrates was approximately 10.[239] Similarly, the relief of steric hindrance in the P_1 site of α-lytic protease, which normally accepts only small hydrophobics, has resulted in a newfound tolerance for large hydrophobic residues in the P_1 position of tetrapeptide substrates with an "extraordinarily...surprisingly...dramatic" retention of activity toward tetrapeptide alanine substrates.[32] In neither of these systems is it possible to test single residue substrates because the enzyme does not hydrolyze single

residue substrates. If the substrate specificity of the enzyme is dictated by the higher order subsites (S_2, S_3, ...), the effect of mutated amino acids on the interactions between substrate and enzyme at these subsites cannot be assessed. To date, similar published studies on trypsin have not exploited the ability of the enzyme to hydrolyze single residue substrates efficiently.[48]

Substrate specificity *has* been introduced into subtilisin in a clever manipulation of the chemical composition of the Michaelis complex. Deletion of the catalytic histidine imidazolium introduces substrate specificity for histidine in the P_2 position of substrate.[40] The turnover for even the best substrate in this system is quite low, however (k_{cat} 3 x 10^{-3} of subtilisin; K_m 16x subtilisin; k_{cat}/K_m 2 x 10^{-4} of subtilisin).

Catalytic Power of the Serine Proteases

Initial experiments on the catalytic function of the serine proteases involved electrostatic surface mutations which presumably would be well tolerated by the enzyme.[220] This series of experiments on subtilisin resulted in a thorough characterization of the effects of electrostatic surface mutations on the pK_a of the imidazolium of His-57.[69,80,81,177,178,220]

The catalytic[a] moieties of the serine proteases have all been mutated directly. The oxyanion hole in the chymotrypsin family of proteases is defined by backbone amide protons, but, in the subtilisin family, it is partially defined by an asparagine side chain. Mutation of this asparagine has indicated the importance of the oxyanion hole to catalytic function of the serine proteases.[37,238] The amino acid side chains of the catalytic triad have also been mutated. The catalytic aspartic acid was mutated to asparagine (D102N) with no perturbation of the structure or structural function (as determined by substrate binding), but with a loss of 10^4 of wild type activity (Chapter 4).[49,203] The catalytic Ser-195 has been mutated to a cysteine, singly and in concert with the D102N mutation. The thorough

[a] That is, those residues that increase the chemical activity of Ser-195 or the substrate carbonyl bond.

characterization of these mutants suggests that the more nucleophilic cysteine residue is sterically incompatible with the active site of trypsin and blocks the oxyanion hole.[101,149] All of the amino acid side chains of the catalytic triad and the asparagine of the oxyanion hole have been mutated singly and in combination, providing an indication of the minimum determinants of catalytic function in serine proteases.[a,41,42] Ile-31, proximal to the catalytic triad in subtilisin has been mutated to Leu-31 causing an unexplained increased in the catalytic activity of the enzyme.[218] Arg-96, a proximal amino acid side chain of the trypsin catalytic triad was mutated to His-96 which apparently chelates metal ions along with the active site His-57 resulting in ion-mediated control of the enzymatic activity.[102]

---

[a] These mutants were analyzed using a single substrate without active site titrations and the results have come under attack invoking tRNA mischarging or codon misreading (J. Tang, *personal communication*).[184]

CHAPTER 2

ELECTROSTATIC THEORY APPLIED TO THE SERINE PROTEASES

Based on recent reviews,[96,193] the physical basis of the electrostatic potential is described followed by a derivation of the Poisson-Boltzmann equation. Finally, two numeric models of the electrostatic potential are presented follwed by a discussion of their application to the catalytic mechanism of the serine proteases.

THE PHYSICAL BASIS OF THE ELECTROSTATIC POTENTIAL

DEFINITION OF ELECTROSTATIC POTENTIAL

The intuitive feel for the electrostatic potential can get lost in the various descriptions and mathematical treatments purporting to describe it. Often, the electrostatic potential is confused with the *electric field.* It is worthwhile, therefore, to attempt to provide a rigorous, quickly reproducible definition which can be fruitfully considered during any discussion of electrostatic results. The author has found one particular definition which fulfills this need.[170]

The electrostatic potential is a scalar function whose gradient is the electric field.

A scalar field is defined solely by numerical values; *i.e.*, at any point in space, a single number defines the value of the function being considered. The whole set of numbers and

their location defines the function. This definition also requires a feel for the mathematical operation, *gradient*, which is written as in (7) for the gradient of a scalar function, ϕ,

$$\nabla\phi = \vec{\mathbf{i}}\frac{\partial\phi}{\partial x} + \vec{\mathbf{j}}\frac{\partial\phi}{\partial y} + \vec{\mathbf{k}}\frac{\partial\phi}{\partial z} \qquad (7)$$

where $\vec{\mathbf{i}}$, $\vec{\mathbf{j}}$, and $\vec{\mathbf{k}}$ are the unit vectors of the Cartesian coordinate system and ∇ is del, the symbol for the gradient operator. The application of the gradient operator is therefore decomposed into computing, at a given point, the slope of the scalar field in the x, y, and z directions, and summing the vectors arising from those three calculations. Note, therefore, that the electric field is a *vector* field; *i.e.*, the electric field function is defined by a vector at each point in space. A definition for gradient is provided by the same authors:[170]

> *The gradient of a scalar function is a vector whose magnitude is the maximum directional derivative at the point being considered and whose direction is the direction of the maximum directional derivative at the point.*

The verbose definition is required because the potential field is a continuous three-dimensional function. Alternatively, the gradient can be thought of as the *slope* of the function at that point, the slope always being taken in the direction in which the value is largest. Once the intuitive notion of the gradient is firmly embedded, the scalar electrostatic potential field and the vector electric field become firm, rigorous, reproducible constructs which must be generated by any attempt to calculate, describe, or depict them.

Physical Determinants of the Electrostatic Potential

The electrostatic potential field (*i.e.*, the values of the electrostatic potential function) of an aqueous system is defined by the presence and interactions of three entities: the induced dipoles, the permanent dipoles, and the charges. Induced dipoles arise from the polarizability of the electron clouds of atoms. Permanent dipoles are defined by the

arrangement of different nuclei in molecules. Charges are thought of as the principal determinants of the electrostatic potential, yet the interactions with the less energetic elements defining the electrostatic potential are what determine the observed potential. In particular, orientation of water dipoles effectively screens charges.[74,167] In contrast, the inability of protein bound dipoles to reorient relative to any position defined by the structure of the protein prevents such screening, effectively amplifying electrostatic interactions between charges within the protein interior. In addition, mobile charges in bulk solvent and the dipolar orientation of bulk water molecules play some role in determining the electrostatic potential within the solvent-inaccessible regions of protein.

THEORETICAL MODELS OF THE PHYSICAL BASIS OF THE ELECTROSTATIC POTENTIAL

THE MATHEMATICAL MODEL - THE POISSON-BOLTZMANN EQUATION

The Poisson-Boltzmann equation provides a thorough mathematical treatment of the three determinants of the electrostatic potential. This imposing equation is more approachable if built up gradually while taking into account more electrostatic effects.[96,193] The fundamental equation of electrostatics is LaPlace's equation (8) for

$$\nabla^2\phi(\mathbf{r}) = 0 \qquad (8)$$

regions of uniform dielectric constant with no free charges, where position is indicated as the point, $\mathbf{r}$, and ϕ is the electrostatic potential. Immediately, this equation changes the electrostatic problem from calculating the electric field to finding a potential function for which the gradient of the differential is zero (9).

$$\nabla^2\phi \equiv \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial^2 y} + \frac{\partial^2\phi}{\partial z^2} = 0 \qquad (9)$$

The solution to the equation is derived from the boundary conditions. For *fixed* charges in a uniform dielectric medium, the electrostatic potential must satisfy Poisson's equation

$$\nabla^2\phi(r) + \frac{4\pi\rho(r)}{\varepsilon} = 0 \qquad (10)$$

where $\rho(r)$ is the charge density as a function of position and ε is the dielectric constant of the medium. For a point charge, q, the solution of this equation is Coulomb's Law where

$$\phi = \frac{Cq}{\varepsilon r} \qquad (11)$$

C is a constant depending on the units and r is the distance from the charge. Note that the solution of the Poisson equation is a scalar field of values definable by a function relating the charge (q), distance (r), and dielectric constant (ε). The gradient of this function is the electric field, a vector field. In this context, Coulomb's law is immediately identified as a special case of a more general theory, and therefore, its failure in a system which has media of different dielectric constants is more easily comprehended. In fact, the dielectric constant can even be a continuously changing function of position as in (12) which is not

$$\nabla \cdot \varepsilon(r)\nabla\phi(r) + 4\pi\rho(r) = 0 \qquad (12)$$

so easily made by changes of Coulomb's law. The divergence operator, $\nabla \cdot$, is defined such that the divergence of the scalar dielectric constant function is the gradient

$$\nabla \cdot \varepsilon(r) = \frac{\partial\varepsilon(r)}{\partial x} + \frac{\partial\varepsilon(r)}{\partial y} + \frac{\partial\varepsilon(r)}{\partial z} \qquad (13)$$

while distribution of the divergence operator through the argument regenerates the Laplacian component of the equation:

$$\nabla \cdot \nabla\phi(r) = \nabla^2\phi(r). \qquad (14)$$

The theory for mobile charges in solution introduces the hyperbolic sine dependence on the electrostatic potential multiplied by κ^2 such that

$$\nabla \cdot \varepsilon(r)\nabla\phi(r) - \kappa^2 \sinh[\phi(r)] + 4\pi\rho(r) = 0 \qquad (15)$$

where κ has the dimension cm^{-1} and $1/\kappa$ is the Debye-Hückel screening distance for mobile charges in bulk solvent.

$$\kappa^2 = \frac{8\pi N_a e^2 I}{1000\varepsilon k T} \qquad (16)$$

The parameter κ is a function of Avogadro's number, N_a, the charge on the proton, e, the ionic strength, I, the dielectric constant, ε, Boltzmann's constant, k, and the absolute temperature, T. The number 1000 is added to account for units. In general then, the Poisson-Boltzmann equation is at least traceable to physical parameters which account for the three effects which determine the electrostatic potential. In addition, the problem of defining the electrostatic potential function is presented as the problem of finding a function *or the numerical values of a function* which fit the above equation.

Analytic solutions of the Poisson-Boltzmann equation to find the function, ϕ, are not currently possible for true protein structures (ϕ = ?). Analytic solutions can be found if simplifying assumptions are made about the shape of the dielectric media. In particular, the problem of a low dielectric sphere immersed in an infinite expanse of a high dielectric medium with charges positioned at various depths in the sphere can be solved exactly. This is Tanford-Kirkwood theory[219] which, although not generally useful for problems involving proteins, serves as a rigorous test for models based on numerical solutions to the Poisson-Boltzmann equation.

NUMERIC MODELS

Protein Dipole Langevin Dipole Model (PDLD)

In the Protein Dipole Langevin Dipole Model (PDLD) developed by Warshel and coworkers,[234] the coulombic charge and a point dipole based on the polarizability of the atom type is assigned to each atom in a protein.[232] The polarizability field is then

calculated by iteration where the dipole moments and charges of each atom at one step are used to calculate an electric field which is used to calculate the orientation of the dipoles for the next step. The calculations converge to a self-consistent field, usually within 7 iterations.[232] Bulk solvent is ES' as a cubic lattice of point dipoles 3Å apart mimicking the density of water. Water grid points less than 3Å from the protein are removed, and dipoles in the first shell are allowed to relax and even to shift position to better pack the protein based on interactions with the charges on the protein. Then all of the dipole grid points are allowed to interact with the protein charges in another iterative calculation. The water dipoles are calculated using the Langevin potential formula relating the orientation of the water permanent dipole to the electric field. Thus the name of the method. The coulombic treatment is justified because no explicit dielectric is assumed; rather, the charges *and* polarizabilites of all atoms and water molecules are modeled explicitly and the effects which give rise to the dielectric constant are calculated on a microscopic level. In addition, the technique uses quantum mechanics to calculate the charges and bonding of the atoms involved in all bond making and bond breaking processes on the enzyme including putative transition states. This quantum mechanical system is presented to the classical electrostatic dipole fields of the protein and solvent as a charge distribution which can easily be taken into account in the PDLD method. Thus, the dipole model need only be used to describe the stabilization of putative transition states.[232] In sum, the PDLD method accounts for inducible and permanent dipoles explicitly as dipoles and accounts for mobile charges as a factor affecting the susceptibility of water grid dipoles to reorient. Permanent dipoles of the protein are modeled explicitly by the charges of the atoms.

Finite Difference Poisson-Boltzmann Model (FDPB)

As implied by the name, the Finite Difference Poisson-Boltzmann Model (FDPB) calculates a numeric solution for the linearized Poisson-Boltzmann equation applied to a protein in a solution containing mobile charges. As in the PDLD method, charges are

assigned to each atom of the protein, including hydrogens. The primary assumption of the FDPB model is the division of protein and solvent into discrete dielectric regions divided by the solvent accessible surface of the protein. The macroscopic dielectric constant of water, 80, is assigned to the solvent, and a low dielectric constant of 2 – 4 is assigned to the protein interior. These assignments can be justified based upon experimental results and the physical boundary conditions of the system.

The Dielectric Constant of Water

The primary cause of the high dielectric constant of the solvent is the ability of the permanent dipoles of the water molecules to reorient and screen charges in solution.[57,97] The contribution to screening by orientation versus polarizability has been experimentally determined by measuring the dielectric constant of water excited by microwave radiation. In the typical dielectric constant determination, a potential is applied across two plates separated by a dielectric medium. The amount of charge which can be held by the plates is proportional to the dielectric constant of the medium. Thus, vacuum has a dielectric constant of 1, because there are no molecules to mitigate the potential across the plates. Apolar hydrocarbons have dielectric constants of 2 – 4 because they contain no permanent dipoles which can reorient, but do have polarizable electron clouds which counter the potential across the plates. Water has a dielectric of 80 because the reorientable dipoles can realign to counter the potential. Actually, to account for a dielectric constant of 80, only 1 in 5 x 10^6 molecules rotate,[50] or each molecule rotates 0.0001°.[96] Microwave radiation excites the rotational energy of water molecules, and water which has absorbed microwave radiation has a dielectric constant of ~4.

The Dielectric Constant of Protein

Thus, inducible dipoles are accounted for in the FDPB method as the dielectric constant. The permanent, orientable dipoles of bulk water are excluded from the interior of the protein at the solvent accessible surface. The low dielectric interior of the protein is

therefore defined by the polarizability of the electron clouds of the protein atoms and the absence of orientable dipoles. As in the PDLD method, permanent dipoles in the protein are represented by explicit charges which define the dipoles. However, the permanent dipoles in bulk solvent are accounted for by the high dielectric constant assigned to solvent. Mobile charges in solution are represented in the calculations by the Debye-Hückel term. In most implementations of the FDPB method, the hyperbolic sine dependence is omitted such that sinh[ϕ(r)] is calculated as [ϕ(r)], a reasonable approximation at low ionic strength.

Numerical Method

The solution to the Poisson-Boltzmann equation is determined by numerical evaluation of the integration of the linearized Poisson-Boltzmann equation. The description presented here is taken from the text by Honig and coworkers who developed the method.[120]

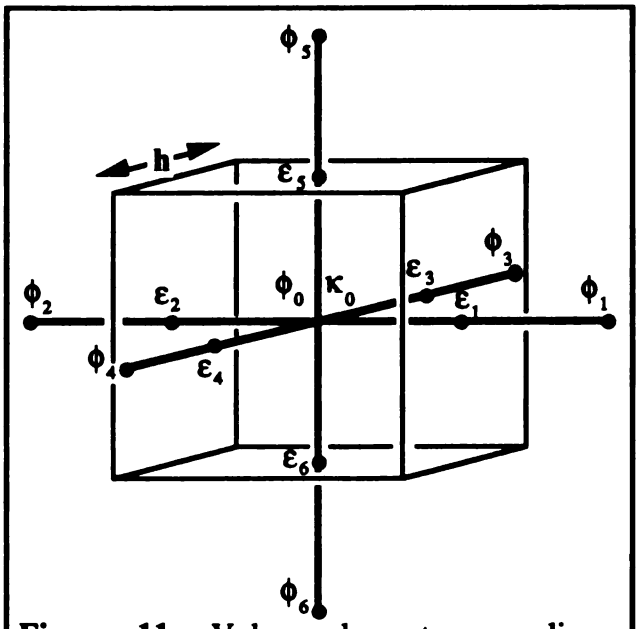


Figure 11. Volume element surrounding a grid point in a finite difference Poisson-Boltzmann calculation.[120]

In practice, the protein structure is immersed in a cubic grid no more than 65 points on a side. A tiny, cubic volume element is defined around each point with the center of each face of the cube lying on the line between grid points (Figure 11). The integrated form of the linearized Poisson-Boltzmann equation (approximating **sinh**[ϕ(r)] as ϕ(r)) is written as follows:

$$\iiint \nabla \cdot \varepsilon(r)\nabla\phi(r)\ d^3r - \iiint \kappa(r)^2\ \phi(r)\ d^3r + 4\pi\iiint \rho(r)\ d^3r = 0 \qquad (17)$$

The first term is transformed **(18)** using Gauss's law to relate the total charge enclosed by the surface to the integral of the surface normal component of the electric field.[170]

$$\iiint \nabla \cdot \varepsilon(r)\nabla\phi(r)\ d^3r = \iint \varepsilon(r)\nabla\phi(r)\ dA \qquad (18)$$

dA is the surface element for which the normal vector is evaluated. This is an integrable function, and the surface considered is each face of the cube so that the first term of **(17)** becomes

$$\iint \varepsilon(r)\nabla\phi(r)\ dA = \sum \varepsilon_i(\phi_i - \phi_0)h \qquad (19)$$

where $i = 1 - 6$, and **h** is the distance between grid points. The value of the dielectric constant is set to 4 if the midpoint between grid points lies inside the solvent accessible surface of the protein and to 80 if the midpoint lies outside the solvent accessible surface. The second term of **(17)** is approximated by evaluating the Debye-Hückel parameter **(16)** at the grid point and evaluating the volume integral of the potential $\phi(r)$ as in **(20)**.

$$\iiint \kappa(r)^2\ \phi(r)\ d^3r \approx \kappa_0\phi_0h^3 \qquad (20)$$

The third term of **(17)** integrates the charge density within the cube **(21)** which is simply the charge contained within the cube q_0.

$$4\pi\iiint \rho(r)\ d^3r = 4\pi q_0 \qquad (21)$$

Thus the integral of the linearized Poisson-Boltzmann calculation **(17)** is reduced to

$$h\sum \varepsilon_i(\phi_i - \phi_0) + h^3\kappa_0\phi_0 + 4\pi q_0 = 0 \qquad (22)$$

which is rearranged and evaluated as

$$\phi_0 = \frac{\sum \varepsilon_i\phi_i + 4\pi q_0}{\sum \varepsilon_i + \kappa_0{}^2h^2} \qquad (23)$$

for each grid point. An initial guess is made for the potential of each grid point, and then new values are calculated for each point using **(23)** until a self consistent field is obtained. Thus a scalar electrostatic potential field is calculated. Values for the potential at other

positions are obtained by interpolating from the values of the appropriate grid points. By making the grid spacing smaller and calculating the potentials of the grid points on the edge of the cubic lattice from a previous lower resolution calculation, atomic resolution of the electrostatic potential field is possible.[82]

Validity of FDPB Results

Three results justify the use of the FDPB method. First, analytic solutions of the Poisson-Boltzmann equation are possible for simplified geometries. In particular, Tanford-Kirkwood theory considers the special case which can be solved exactly: a low dielectric sphere immersed in an infinite expanse of high dielectric medium with charges embedded at various depths in the sphere.[219] Such analytic cases can and should be used to test the faithfulness of the numeric approximation to the rigorous execution of the mathemetical theory. The FDPB method has been tested against the Tanford-Kirkwood case. In general, errors are less than 5%; however, at distances less than 3Å, the error climbs to 20%.[82] Second, the FDPB method has been used successfully to calculate the electrostatic effects of surface mutations[177,178] on the pK_a of the catalytic imidazolium in subtilisin.[80,81,208]

Third, both the PDLD method[233] and the FDPB method[201] have been used to calculate the contribution of a negatively charged Asp-102 to the transition state energy of a rat anionic trypsin-catalyzed reaction. The results from these calculations agree well with each other and predict Asp-102 contributes approximately 4 kcal/mol of electrostatic stabilization to the His-57 imidazolium in the transition state. The PDLD results on trypsin were compared with results on subtilisin which revealed a similar electrostatic contribution of the catalytic aspartic acid to transition state stabilization.[233] The FDPB results on rat anionic trypsin were compared with results on cow cationic trypsin which is 12.5 units more positive. These calculations indicated that the electrostatic potential field in the active site was determined by Asp-102 independently of the charge in the substrate binding pocket

or the surface charges.[201] Surface charges of the isozymes are effectively screened by bulk solvent[193] resulting in a nearly neutral electrostatic potential in the active site of each isozyme which is then determined by the charge on Asp-102. If it is assumed that the protein and solvent have the same dielectric constant, then the surface charges of each isozyme generate oppositely charged potentials in the active site which suggests that the observed rates of reaction for each isozyme would be different. These rates are very similar, however, thus validating the assumptions concerning the different dielectric constants for proteins and bulk solvent and verifying that the coulombic model fails, given these assumptions. Interestingly, the PDLD method uses a coulombic model, but explicitly models the polarization and screening of bulk solvent using reorientable dipoles.[232]

CHAPTER 3

PERTURBATION OF THE ELECTROSTATIC FIELD IN THE ACTIVE SITE OF TRYPSIN

INTRODUCTION

Electrostatic stabilization of the transition states of serine protease-catalyzed reactions has been the subject of extensive theoretical study.[122,201,224,233] However, while theoretical calculations may demonstrate which effects are sufficient for catalysis, they do not determine which effects are necessarily at work in catalysis. In a complementary fashion, mutagenesis experiments which involve removing catalytic groups reveal the magnitude of the effect a given substituent has on catalysis, but not the nature of the effect.[37,41,49,203,238] Although electrostatic theories can account for the effects of mutations of the catalytic triad,[41,49,233] it remains to be shown whether or not the electrostatic potential affects the transition states of the fully constituted catalytic triad. Ideally, one would like to perturb, in an informative way, the electrostatic potential of the active site residues.

Ser-214 is a solvent-inacessible amino acid in trypsin which forms hydrogen bonds with Asp-102 and a buried water molecule.[29] This structural motif has been demonstrated crystallographically in both the trypsin and subtilisin families of serine proteases,[153] and this hydroxyl group is evolutionarily conserved in all eukaryotic serine proteases and even

in the 3c gene of some viral cysteine proteases (Table 3) suggesting that it is an essential moiety.[10] Nevertheless, the trypsin structure is able to accomodate mutation of Ser-214 to amino acids of varying bulk and charge.

Table 3. *Conservation of Serine 214. The protease genes from several sources were analyzed to determine overall homologies.*[10,153] *The residues found at position 214 in proteases from the various listed sources are presented below.*

Source Residue at 195		Residue at 214
Serine	**Cysteine**	
trypsin, chymotrypsin, elastase, *Streptomyces griseus* trypsin	tobacco etch virus-N1a, tobacco-vein-mottling virus-N1a, foot-and-mouth disease virus-3c, Theiler's murine encephalomyelitis virus-3c, encephalomyocarditis virus-3c	Serine
Subtilisin, *Streptomyces griseus* protease A, *Streptomyces griseus* protease B, α-lytic protease	bovine enterovirus-2a, coxsackie-virus-2a, human polio virus-2a, human rhinovirus strains 2 & 14-2a	Threonine
	cow pea mosaic virus-24K, hepatitis A virus-3c, bovine enterovirus-3c, coxsackievirus-3c, echovirus strain 9-3c, human polio virus-3c, human rhinovirus strains 2 & 14-3c	Valine
Staphylococcus aureus V8 protease		Tryptophan

The successful integration of a buried lysine into the structure of S214K trypsin near the active site[148] provides the opportunity to study the effects of a fixed, buried charge on the function of the intact catalytic triad, without the overwhelming complications due to interactions of the charge with bulk solvent. The S214E and S214Q trypsins provide controls for the structural perturbations caused by the S214K mutation.[148] The Glu-214 and Gln-214 side chains cause structural perturbations which are similar to, or more severe than, those caused by the Lys-214 side chain. An analysis of the kinetic data for these mutants permits us to distinguish the structural effects of the Ser-214 mutations from the electrostatic effects.

MATERIALS AND METHODS

THE MUTAGENESIS EXPERIMENT - AN OVERVIEW

The design of the actual mutation is at the discretion of the investigator. The interpretation of experimental data will always benefit over time as theories are tested and refined. The answer to the question, "What will happen if we change X to Z?" is invariably, "No one knows. You'll have to do the experiment to find out." Any inspiration will suffice.

In the case of the trypsin system used in this thesis, design of the mutant to characterization of the purified protein can be a relatively brief time. Computer modeling of the mutant protein, based at some point on electron density maps or nuclear magnetic resonance cross peaks, must serve as the starting point for any structure-based directed mutagenesis experiment; however, our understanding of protein structure, function, and folding at this date does *not* justify the elimination of first line experiments based solely on a theoretical structure. Too many unexpected results occur, and some mutations are simply interesting by design and *must* be tried in the face of *any* counter arguments.

However, from its conception, through its execution, and during its evaluation, the oligonucleotide-directed mutagenesis experiment itself is fraught with pitfalls. As with any powerful technique, simple mistakes can lead to large errors in understanding. It is important, therefore, to treat every step with the care of the original experiments. As the neophyte musician must learn, and the virtuoso knows, it is important to play what is on the page and to make changes consciously with deliberation and precise intent.

The mutagenic oligonucleotide must be designed carefully and accurately and the resulting piece of DNA must prime faithfully. The template must be genetically pure and its sequence known. The mutant DNA product must be genetically purified and accurately characterized. The plasmid carrying the mutant gene must be harbored by the host strain without mutation and the gene faithfully transcribed. The evaluation of the mutant protein

product must be done carefully. Preliminary screening in crude extracts is always subject to complications from either potential contamination or endogenous activity. The effect of a mutation cannot be considered to be evaluated until the protein is purified and characterized with full knowledge of the limitations of the characterization.

GENERATION OF MUTANT PROTEINS

Directed Mutagenesis

Theory

Kunkel has developed an efficient method for introducing mutations into DNA.[133] The technique is based on the preparation of uracil-containing DNA, an unnatural form which is stable *in vitro* and in *dut⁻ ung⁻ E. coli* but is unstable in *dut⁺ ung⁺ E. coli*. In normal *E. coli*, deoxyuridine triphosphate (dUTP) is synthesized as an intermediate in the biosynthesis of deoxythymidine triphosphate (dTTP).[159] In the absence of dUTPase which is encoded by the *dut* gene, the concentration of dUTP in the cell increases, and the amount of uracil incorporated into cellular DNA also increases.[195,223] Uracil is normally removed from DNA by uracil N-glycosidase, which is encoded by the *ung* gene, and the resulting gap triggers excision repair.[222,223] In excision repair, rather than simply replace the deleted base, the repair proteins resynthesize the entire context of the gap on the uracil-laden strand. It follows, therefore, that a uracil-laden strand of DNA will be biologically unstable while a non-uracil-containing strand will not only be stable, but will be used as the template during excision repair of the uracil-laden strand. The trick is to prepare a double-stranded piece of DNA in which only one strand contains uracil and in which the mutation is located in the non-uracil-containing strand. The Kunkel mutagenesis technique therefore relies on the preparation of single-stranded, uracil-laden DNA, and polymerization of the complementary strand primed by a mutagenic oligonucleotide. Bacteriophage M13 makes single-stranded DNA from a double-stranded template.[123] A small segment of the M13

genome, the *ori*, encodes a recognition sequence which is sufficient to direct single-stranded synthesis of double-stranded plasmid DNA by the machinery of a helper phage.[55] To maximize the production of single-stranded plasmid DNA, the helper phage genome contains a defective *ori*, an *ori* with a non-optimal sequence.[52] Recognition of the *ori* is like a synthetic chemical reaction under thermodynamic control – the protein:DNA binding reaction favors the plasmid *ori*. The trypsin gene with a bacterial promoter and transcription terminator has been incorporated into the plasmid $p^{m}T_3$ (Figure 12).[62,63]

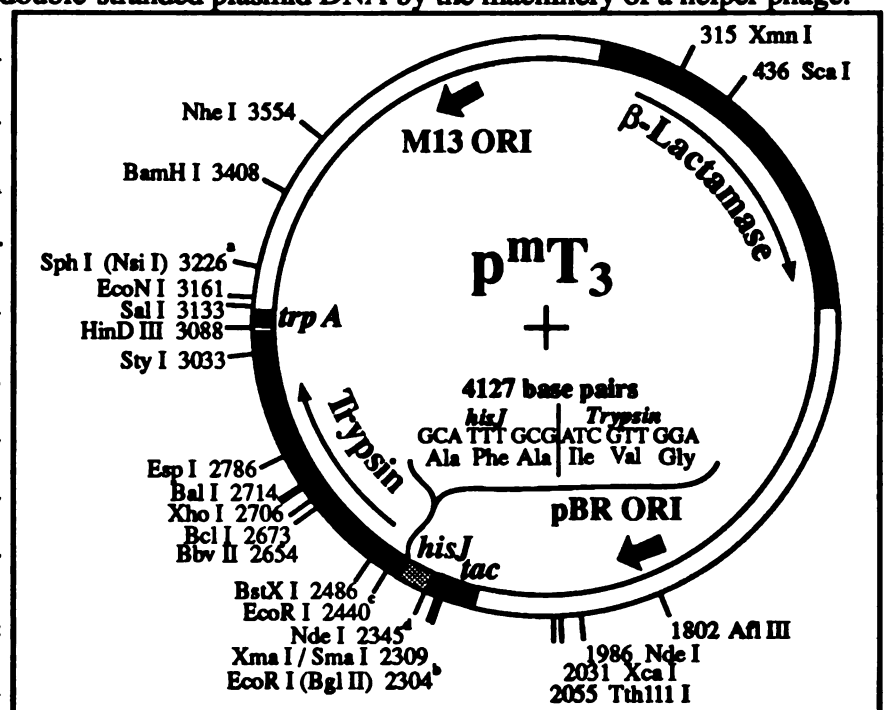


Figure 12. Restriction Map of the $p^{m}T_3$ Family of Plasmids. The positions of the β-lactamase, *hisJ* signal peptide, and trypsin genes, the *tac* promoter and *trp A* transcription terminator are indicated by shading. The cleavage site of the *hisj* signal peptide-trypsin polyprotein is depicted. The M13 *ori* rescues the non-coding strand of the trypsin gene; therefore, mutagenic oligonucleotides should be made such that Met = ATG. Sites are unique except as noted.

[a] Deleted and replaced with Nsi I in $p^{m}T_3$' which is commonly used in the Craik lab as pZ_3

[b] Replaced with Bgl II at 2303 in $p^{m}T_c$ only

[c] Not unique in $p^{m}T_3$

[d] In $p^{m}T_3$-Nde I only

Practice

Preparation of Uracil-laden Single-stranded DNA

Genetics of *E. coli* Strains and Plasmids. The expression plasmid pT_3 encodes a *his J* signal peptide-trypsin fusion protein under the control of the *tac* promoter and *trp A* transcription terminator.[101] The $p^{m}T_3$ plasmid incorporates an M13 *ori* into the

pT_3 plasmid (Figure 12).[63] This plamid is best propagated in *E. coli* strains which tolerate trypsin and which carry a *lac* i^q repressor gene[38,156] for the *tac* promoter; *e.g.*, X90 (F' *lac* i^q, *lacZY*, *proAB* / Δ*lac-pro*, *ara*, *nalA*, *argE*(am),*thi*, rif^r). In order to prepare single-stranded uracil-laden DNA from this plasmid, it was necessary to prepare a *dut ung* strain with a *lac* i^q (Luke Evnin, *personal communication*). Strain LE112 (F' *lac* i^q, *lacZ*::Tn5, proAB / *dut*1, *ung*1, *thi*-1*relA*1)[a] was prepared for this purpose by curing strain CJ236[114,134] of plasmid pCJ105 which encodes a chloramphenicol selectable F pilus.[61] The F episome was obtained by mating cured CJ236 with SY903.[181] Note that the genome of LE112 does not contain a deletion of the gene for proline biosynthesis as does X90. Therefore, it is not possible to select the F episome solely by plating on minimal media. However, the Tn5 transposon also carries onto the F episome a kanamycin resistance[20] gene and a *lacZ* gene. LE112 can therefore be selected by kanamycin resistance and assayed using the X-gal/IPTG induction of blue colonies. Calcium competent LE112 cells should therefore be prepared from blue colonies of fresh streaks on minimal media containing 10 μg/ml kanamycin, 1 mM IPTG, and 25 μg/ml X-gal (40 μl of a 20 mg/ml solution in DMF per plate) (blue color selection protocol under X-gal α-complementation)[179] as they were originally identified.[61] In addition, for the preparation of uracil-laden, single-stranded plasmid of interest, transformants of LE112 should be selected using the selectable marker for the plasmid (50 μg/ml of ampicillin for $p^{m}T_3$) in addition to the minimal medium described above. It is possible that recombination between the *lacZ* on the episome and the genomic copy will result in blue colonies which are not F'; therefore, blue selection is neccessary, but not sufficient for the presence of the F'. *Several transformants of LE112 should be tested, and the ultimate indication of the presence of the F pilus will be infectivity using male specific phage such as M13.*

---

[a] Strain LE113 was prepared in the same mating experiment, should have the same genotype, and operationally has behaved identically to LE-112.

Superinfection. This F episome in LE112 is relatively stable and need not be selected during the rescue of single-stranded DNA. The following rescue protocol has worked consistently for p^mT_3 and pBSCK-3, a high copy number plasmid encoding creatine kinase. Grow an overnight culture of p^mT_3 in LE112 in 3 mls of 2xYT (1% yeast extract, 2% tryptone, 0.5% NaCl in water) containing 50 µg/ml ampicillin. In the morning, dilute 2.5 mls into 25 mls of 2xYT (27.5 ml total volume) containing 50 µg/ml ampicillin. Add 50 µl of 1.0×10^{11} pfu/ml VCSM13 helper phage[a] immediately and grow with vigorous shaking at 37°C for 6 – 8 hours.[b]

Isolation of Single-stranded DNA. Single-stranded DNA is prepared essentially as described.[52] *E. coli* cells are removed from the culture by centrifuging 2 min at 5,000 x *g* and then centrifuging the supernatant 15 min at 5,000 x *g* . The protein-coated phage are precipitated from the second supernatant by adding 7.5 ml of 20% PEG-8000, 3.5M NH_4OAc and incubating for 15 min at room temperature. The protein is then phenol/chloroform extracted from the DNA. The composition of the DNA in the preparation should be determined by testing a small aliquot of the precipitation reaction.

Small Scale Test for the Plasmid DNA : Helper Phage DNA Ratio. Microfuge 1 ml of the precipitation reaction at maximum speed for 5 min. Aspirate the supernatant, quick spin the pellet, and aspirate the remainder of the supernatant (a P-1000 works well). Resuspend the pellet in 300 µl of TE by repreatedly passing the liquid in and out of a pipettor tip. Add 300 µl of buffered phenol[c] vortex to homogeneity, add 300 µl of 1:24 *iso*-amyl alcohol (IAA):chloroform, and vortex to homogeneity. Microfuge 5 min at

---

[a] Obtained from Stratagene, Inc., San Diego, CA.

[b] Some workers report that it is better to wait 30 minutes before adding phage to allow the cells to recover (L. Evnin, *personal communication*), a parameter to be tested during any troubleshooting.

[c] 1% β-mercaptoethanol, 0.1% 8-hydroxyquinoline, ~100 mM Tris, pH 8.0: In the commercial bottle, extract 100 ml of melted, ultrapure phenol 2x with 100 mls of 1M Tris, pH 8.0. After each extraction, aspirate with a 25 ml pipet as much of the the upper, aqueous phase as possible without removing the phenol. Extract the buffered, saturated phenol 2x with 100 mls of deionized, distilled water. Add 1 ml of neat β-mercaptoethanol and 0.1 g of 8-hydroxyquinoline.)

maximum speed, transfer the upper, aqueous phase to a fresh tube, add 300 µl of 1:24 IAA:chloroform, vortex to homogeneity, and microfuge 1 min at maximum speed. Analyze an 8 µl sample of the aqueous phase by DNA electrophoresis through a 1% agarose / TAE gel (Figure 14). DNA samples are prepared by diluting to 20 µl total volume with 4 µl of 5x dye.[a] The samples are heated to 65°C for 10 min immediately before being loaded on the gel.

DNA Electrophoresis through Agarose Gels. A solution of 1% agarose in 1x TAE (40 mM Tris•acetate, 1 mM EDTA)[179] is prepared by autoclaving the solid agarose in TAE buffer. The resulting solution is swirled to homogeneity and maintained as a sterile solution stored at 65°C. A 20 ml sample of 1% agarose solution is poured onto a clean, level $3^1/_4$ x 4 inch glass plate.[c] Surface tension keeps the fluid on the plate. A comb is suspended to mold the desired wells into the agarose which hardens in a few minutes at room temperature. The comb is removed with care not to rip the agarose matrix, and the plate and agarose gel are moved as a unit through electrophoresis, staining and destaining to photography. The glass plate / agarose gel

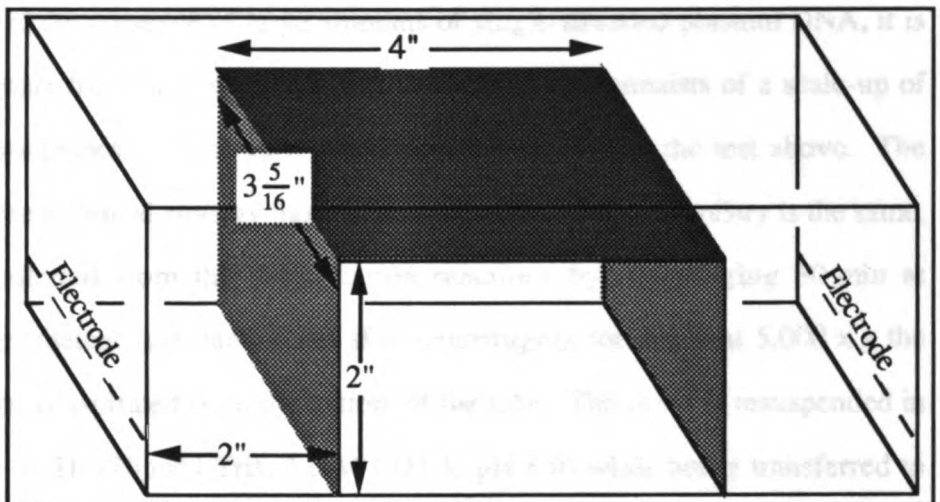


Figure 13. Dimensions of a DNA agarose gel electrophoresis box.[c] A potential of 200V can be applied across the electrodes to fractionate the samples in 30 min through a 1% agarose gel in TAE buffer.

[a] 25% ficoll, 0.5% sodium dodecylsulfate (SDS), 25mM, pH 8.0 ethylenediamine tetraacetic acid (EDTA), 0.05% bromphenol blue; stable at room temperature for years.

[c] The gel box system of DNA electrophoresis used in the Craik lab was engineered by R. Shadel, Inc., 1684 Hudson Avenue, San Francisco, CA 94142. Other gel box systems cannot be run at these settings so this one is described in detail.

[c] "projector slide cover glass" Eastman Kodak Company, Rochester, NY, 14650, catalog #140 2130

assembly is suspended on the stage of the gel box (colored black in Figure 13) in the minimal volume of TAE buffer needed to cover the mouths of the wells. The DNA samples are loaded in the wells from a pipettor and a 200V potential is applied across the gel for 30 minutes. The DNA in the gel is stained in a 500 ml solution of 1.4 μg/ml ethidium bromide (Sigma Chemical Company No. E-8751) for 10 min. The gel is then destained for 10 min in a 500 ml solution of TAE buffer and photographed. The gel is photographed off of the glass plate using a FotoPhoresis UV illumination apparatus and a Polaroid FCR–IO CRT camera with UV and orange filters.[a] Polaroid 667 coaterless black and white film is exposed 1 sec @ f5.6 and developed 30 sec (Figure 14).

Purification of Large Amounts of Uracil-laden Single-stranded DNA. After verification of the presence of large amounts of single-stranded plasmid DNA, it is worthwhile to prepare the entire lot of uracil-laden DNA. This consists of a scale-up of size and purity requirements of the small scale preparation used in the test above. The larger volumes make technical strategy more important, although the chemistry is the same. The phage are pelletted from the precipitation reactions by centrifuging 30 min at 5,000 x *g*. The supernatant is decanted and after centrifuging for 2 min at 5,000 x *g* the residual supernatant is aspirated from the bottom of the tube. The pellet is resuspended in 4 ml total volume of TE (10 mM Tris, 1 mM EDTA, pH 8.0) while being transferred to 15 ml tubes for convenient manipulation. The protein is then exhaustively extracted away from the DNA using phenol/chloroform as follows: (1) 4 ml of buffered phenol is vortexed to homogeneity with the aqueous sample; (2) the phenol / sample mixture is then vortexed to homogeneity with 4 ml of 1:24 IAA:chloroform; finally, (3) the extraction is centrifuged at 5,000 x *g* for 5 min to separate the phases and pack the denatured protein at the interface. The aqueous phase is conveniently aspirated and transferred to a fresh tube with a 5 ml pipet, and the extraction is repeated again until *no* white denatured protein is

[a] The apparatus is commercially available from Fotodyne, Inc., 16700 W. Victor Road, P.O. Box 183, New Berlin, Wisconsin 53151-0183, (414) 786-9550

visible at the interface. This procedure is repetitious and uses a lot of buffered phenol.[a] Repetition at this scale is necessary to remove the protein with as few manipulations as possible. The volume of the aqueous phase is reduced to 300 µl by repeated extraction with *n*-butanol. A butanol extraction is performed by adding an equal volume of *n*-butanol to the aqueous phase, centrifuging in a clinical table-top centrifuge to separate the phases, aspirating, and discarding the upper butanol phase. A volume of butanol 2–3x the volume of the aqueous phase can be used, but care must be taken not to extract *all* of the aqueous phase into the butanol layer.[b] At the 700 µl volume, the aqueous phase can be transferred to a 1.5 ml Eppendorf tube and extractions carried out in these tubes using a microfuge. No significant loss of DNA has ever been noted due

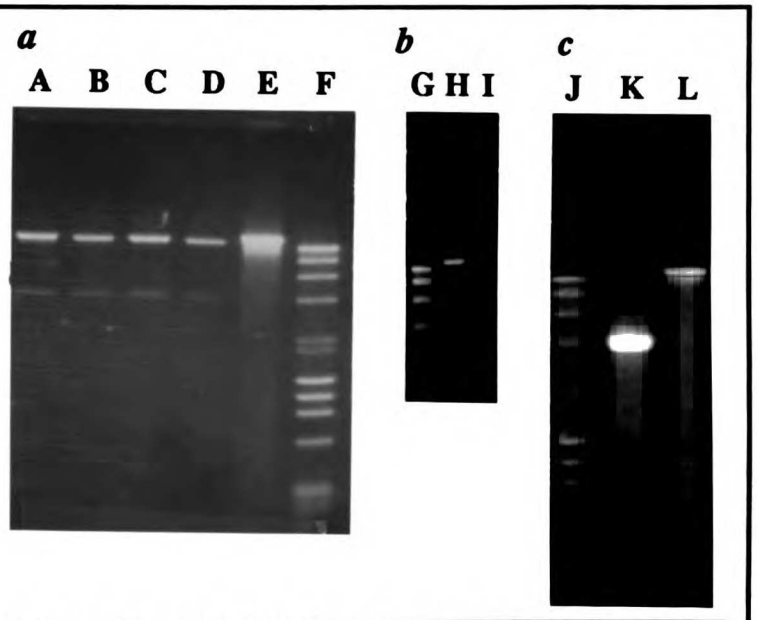


Figure 14. Stages in the Preparation of Single-stranded Plasmid DNA (***a***) Initial screens were conducted on 4 clones of p^mT_3, LE112 transformants (**A–D**). The ratio of helper phage DNA to plasmid DNA is high because the cultures were superinfected with too much helper phage. The preparation in lane **A**, however, was mutated with 80% efficiency to make p^mT_C (Figure 12). (***b***) A saturated overnight culture inoculated from the glycerol stock of clone **A**, was diluted 0.1x into 25 mls of 2x YT which was inoculated with 5×10^9 plaque forming units of helper phage immediately (**I**) or after 30 min (**H**) and grown for 10 hours. The samples were prepared according to the Small Scale Test protocol in the text. (***c***) The uracil-laden single-stranded p^mT_3 DNA from 27.5 ml of culture tested in I was purified according to the Large Scale protocol in the text (**K**). Uracil-laden single-stranded M13 DNA (**E,L**); λ/ϕ linear double-stranded DNA markers (**F,G,J**)

---

[a] It has been reported that after one or two extractions, an equal volume of 7.5M ammonium acetate can be used to precipitate protein, followed by two volumes of ethanol to precipitate the DNA. This procedure was never attempted by the author, but would be very useful.

[b] When this happens, a few microliters of 5M NaCl wan be used to extract an aqueous phase back out of the butanol with no loss of DNA.

to this butanol extraction technique which relieves the problems of ethanol precipitating DNA, drying and resuspending the pellet in a 15 ml centrifuge tube. When the volume of the aqueous phase is 300 µl, the phenol/chloroform extractions are repeated on the smaller scale; the quality of the first interface is ignored. When the interface of the small scale extractions is *clean*, the DNA is precipitated. DNA precipitations are carried out by mixing in a $^1/_{10}$ volume of 3M sodium acetate (NaOAc) and 2.5 volumes of 5°C 100% ethanol. The precipitation is carried out at 0°C for 10 min and the DNA pelleted by microfuging at maximum speed for 30 min. The supernatant is aspirated using a sterile tipped pipettor. The pellet is rinsed by gently, but quickly, adding and removing ~100 µl of 5°C, 70% ethanol. The pellet is dried under vacuum in about 5 minutes in a "Speed Vac Concentrator."[a] The pellet is resuspended in 100 µl TE and a 2 µl sample of this solution is analyzed on a 1% agarose gel in TAE. This preparation is the template used in mutagenesis reactions.

Design of Mutagenic Oligonucleotides

The design of unique oligonucleotides has been reviewed.[152,249] Operationally, 22 – 28mers will prime at a unique site, although sequences may be duplicated in the template. It is desirable to have G's and C's at the end of the oligo, and care must be taken that the oligos are not palindromic so that they dimerize. Some strategy in the design of the mismatched base pairs may facilitate the later steps of the mutagenesis. In particular, the introduction of a restriction site which is present if and only if the mutation is present facilitates screening of the putative mutants. Such mutations were identified using the computer program "Snipper"[157] to identify base changes which introduce restriction sites without changing the amino acid sequence encoded by the resulting RNA. All of the various stringencies are possible for the coupling of the restriction site to the mutation, and

[a] Savant Instruments Co., Farmingdale, NY.

occasionally, no restriction site can be introduced at all. The user must then rely upon the advantages of the Kunkel method.

Mutagenesis Reactions

The mutagenic oligonucleotide is phosphorylated at the 5′ position to provide a substrate for ligase after the polymerization reactions. The phosphorylated mutagenic primer is annealed to the template and the second, non-uracil-containing strand is synthesized by the polymerization of free nucleotides by T_4 DNA polymerase. The advantages of T_4 DNA polymerase are processivity and lack of a strand displacement activity. Other polymerases have been successfully used but will falter occasionally. The T_7 DNA polymerase is highly processive but, under the conditions of the protocol in this thesis, apparently displaces the mutagenic oligonucleotide resulting in no mutants. The polymerization reaction is carried out in the presence of T_4 DNA ligase to seal the nicks in the DNA and improve the efficiency of mutagenesis.[134]

Oligonucleotide-directed Mutagenesis Protocol

Preparation:

1. **Uracil Laden DNA Template** should be at least 100 ng/μl or 100 ng per reaction.
2. **Mutagenic Oligonucleotide** should be 50-100μM for the following kinasing reaction.

Kinasing:

1. In a 0.5 ml eppendorf tube combine
   a. 100 pmol of mutagenic oligonucleotide.
   b. 1.5 μl 10x Kinase Buffer[a] (III-227C)
   c. 0.5 μl 10 mM ATP
   d. 1.0 μl T_4 polynucleotide kinase
   e. Sterile water to 15 μl
2. Incubate at 37°C for 50 minutes.
3. Heat to 65°C for 10 minutes.

Annealing:

1. In a 0.5 μl eppendorf tube combine

---

[a] 10x Kinase Buffer = 0.5 M Tris•HCl, pH 8.0; 0.1 M $MgCl_2$; 0.1 M DTT; 0.01 M Spermidine

a. 100 ng of Uracil-laden DNA Template
b. 1.5 μl of phosphorylated mutagenic oligonucleotide.
c. 1.0 μl of 10x annealing buffer (III-285A)[a]
d. Sterile water to 10 μl.

2. Boil 3 minutes.
3. Incubate at 37°C for 20 minutes.
4. Plunge into ice.

Synthesis:

1. Prepare Synthesis Mix in a 0.5 μl eppendorf tube
   a. 1.0 μl of 0.1 M dithiothreitol (DTT) per reaction
   b. 1.0 μl of T_4 DNA ligase per reaction
   c. 1.0 μl of T_4 DNA polymerase per reaction
   d. 2.0 μl of 10x Synthesis Buffer[b] per reaction
   e. 2.0 μl of 4 mM dNTP's (III-285C) per reaction
   f. 1.5 μl of 10 mM ATP (III-227A) per reaction
   g. 1.5 μl of sterile water per reaction
2. Add 10 μl of Synthesis Mix to each tube of Annealing Mix.
3. Incubate the Synthesis Reactions for 90 minutes at 37°C.

Transformation

Transform a dut+ ung+ strain of the highest available competency with 3 μl of Synthesis Reaction.

Transformation

One of the major advantages of using *E. coli* as a biological system is that the cells are readily transformed.[90,91] The chimeric DNA, (half uracil-laden) is transformed into a *dut*$^+$*ung*$^+$ strain. The *dut*$^+$*ung*$^+$ host must tolerate the gene and gene product, and so, for trypsin, it is convenient to transform X90 at this point. The number of transformants may be encouraging or discouraging but is never diagnostic for the success of the mutagensis reaction.

Transformation of Ca^{2+} Competent *E. coli*

Purpose:

Frozen Ca^{2+} competent *E. coli* were routinely transformed using the following protocol.

---

[a] 10x = 200mM Tris, pH 7.4; 100mM $CaCl_2$
[b] 10x = 200mM Tris, pH 7.4; 350mM NaCl

Set Up:

1. Thaw frozen Ca^{++} competent cells on ice.
2. Lable and chill 17x100mm snap-capped culture tubes on ice.
3. Lable selective plates.

Procedure:

1. Add 100 µl of thawed competent cells to ice-cold, 17x100mm snap-capped culture tubes.[a]
2. a. Add 5 µl of the indicated DNA preparation to each tube.
   b. Incubate the transformations on ice for 1 hour.
3. Heat shock the transformations at 37°C for 5 minutes.
4. a. Add 400 µl of LB medium to each tube.
   b. Outgrow at 300rpm at 37°C for 1 hour.
5. a. Plate 167 µl of each outgrowth on selective plates (L-amp).
   b. Incubate plates at 37°C overnight.

Preparation of Frozen Ca^{2+} Competent *E. coli*[b]

Frozen Ca^{2+} competent *E. coli* were routinely prepared using the following protocol.

Culture:

1. Pick and grow a single colony from a fresh streak in a 2.5ml overnight.
2. a. Autoclave 200mls of LB in a 2L flask.
   b. Store this preparation at 37°C with shaking, if possible.

Set Up:

Solutions can be prepared during growth of the 2L cultures:

0.1M $MgCl_2$:
Dilute 5mls of sterile 1.0M $MgCl_2$ to 50mls with sterile water and store the solution on ice.

0.1M $CaCl_2$: ***Make fresh***
a. Dissolve 1.47g of solid $CaCl_2 \cdot 2H_2O$ in 100mls of gd.[c]
b. Sterile filter the solution and store it on ice.

15% glycerol in 0.1M $CaCl_2$:
Dilute 1.9ml of sterile 80% glycerol with 8.3mls of sterile 0.1M $CaCl_2$.

As time permits, label ~60 sterile 1.5ml eppendorf tubes with the strain and date.

---

[a] To avoid shear forces, cut the tip off of any pipet tip used to transfer competent cells.
[b] By Jennifer Turner and Daniel Wandres
[c] glass-distilled water

Procedure: **Keep the preparation on ice at all times!!!**

1. a. Add 2mls of the overnight culture to 200 mls of LB in a 2L flask.
   b. Grow at 37°C at 300rpm to A_{600} = 0.5 (~2.5-3hours for JM-101).
2. a. Chill the flask in an ice-water bath 10 minutes, also
   b. Chill 4x 50 ml sterile orange capped centrifuge tubes on ice.
3. a. Pellet cells 5 minutes at 6,500 rpm (5096 g) in an SS-34 rotor.
   b. Decant, bleach, and discard the supernatants.
4. a. Resuspend each pellet gently with a 10 ml pipette in 10 mls 0.1 M $MgCl_2$.
   b. Combine the resuspended cells in a single tube.
   c. Rinse the three empty tubes with 10 mls of 0.1 M $MgCl_2$.
   d. Keep the preparation (Total volume = 50mls) on ice for 5 minutes.
5. a. Centrifuge the preparation at 6,500 rpm (5096 g) for 5 minutes.
   b. Decant, bleach, and discard the supernatant.
6. a. Resuspend the cells gently but thoroughly in 10 mls of 0.1 M $CaCl_2$.
   b. Add 40 mls of 0.1 M $CaCl_2$.
   c. Incubate the preparation on ice for 2 hours.
7. a. Centrifuge the competent cells at 6,500 rpm (5096 g) for 5 minutes.
   b. Decant, bleach, and discard the supernatant.
8. a. Resuspend the cells very gently in 10 mls of 15% glycerol in 0.1 M $CaCl_2$.
   b. Incubate the cells on ice for 10 minutes.
9. a. Chill ~60 closed, 1.5 ml eppendorf tubes on powdered dry ice.
   b. Open the tubes, but keep them on dry ice.
   c. Aliquot 205μl of the competent cell preparation into each of the tubes.

Store the competent cells at -70°C.

Detection of Mutants

The trypsin protein and gene produced by *E. coli* clones which have been transformed with the mutagensis reaction can each be rapidly characterized. Small scale cultures provide sufficient samples to allow an initial analysis of the putative mutants based on one or more of the following: periplasmic trypsin activity, periplasmic trypsin protein, and restriction or sequence analysis of the plasmid DNA. Trypsin activity can be quickly assayed in incubations of periplasmic extracts with substrates which are monitored using a microplate reader.[63] Alternatively, the periplasm is fractionated by polyacrylamide gel electrophoresis (PAGE) and activity-stained using a gel overlay containing phenol red pH indicator and tosyl-arginine methyl ester (TAME) (Figure 36).[62] Trypsin protein production is monitored specifically by immunological analysis of PAGE-fractionated

periplasmic extracts transferred to nitrocellulose paper (western blot) as described.[227] Miniprep plasmid DNA is prepared by a slight modfication of published methods which yields DNA of sufficient purity for restriction analysis and DNA sequencing.

Trypsin Activity Gel Protocol

Trypsin activity in periplasmic extracts was routinely assayed using the following protocol.

Cultures:

1. Grow a 2ml culture under conditions to select for the expression plasmid overnight (in L-amp [IPTG if necessary] in 17x100mm culture tubes ~8 hours @ 37°C x 300rpm).
2. Don't forget to run a positive control such as pTn2 in SM-138.

Set Up:

1. a. Pour a 12.5% PAGE-SDS separating gel.
   b. Pour a 4.75% stacking gel with 40 μl wells.
2. Prepare a 50mg/ml lysozyme solution in 20mM EDTA, pH 8.0 (300 μg of lysozyme per sample).
3. Pour an Activity Stain Gel:
   in a 125ml filter flask combine:
   -4.3mls of 30% acrylamide (29.2% acrylamide + 0.8% bis-acrylamide)
   -6.0mls of 2mM phenol red
   -8.0 μl of 10M NaOH
   -5.0ml of 233mM TAME
   -125 μl of 10% ammonium persulfate
   -degass
   -add 13.5 μl TEMED and pour the gel
   -Layer the top of the polymerizing gel solution with 70% EtOH.

Protocol:

1. a. Pellet a 1.5ml lot of the culture in a microfuge: 1 minute at maximum speed (~12,000rpm).
   b. Decant the supernatant. Re-microfuge the sample briefly.
   c. Aspirate and discard the remaining supernatant.
2. a. Resuspend the pellet in 20 μl of 25% sucrose, 10mM Tris, pH 8.0.
   b. Add 6 μl of a fresh 50mg/ml lysozyme, 20mM EDTA, pH 8.0 solution.
   c. Incubate the suspension at room temperature for ~45 minutes.
3. Pellet the spheroplasts by microfuging the sample for 5 minutes at the maximum possible speed.
4. a. Remove 20 μl of the supernatant to a fresh tube containing 20 μl of 2x-Laemmli buffer which contains no β-mercaptoethanol.

b. Heat the sample at 65°C for 3 minutes.
c. Load the whole sample on a 12.5% PAGE-SDS gel.
d. Electrophorese at 250V for 30 - 90 minutes.[a]

5. Mark the orientation of the gel by cutting off one of the bottom corners.

6. a. Soak the gel in 300mls of 2.5% Triton X-100 for 40 minutes. (30mls of a 25% solution + 270mls of glass-distilled water)
   b. Soak the gel in 300mls of 5mM $CaCl_2$. (1.5mls of 1M $CaCl_2$ + 300mls of glass-distilled water)

7. a. Open the Activity Stain Gel only immediately before using it.
   b. Overlay the separation gel on the Activity Stain Gel.
   c. Wrap the developing activity gel with plastic wrap and incubate at 37°C.

8. Photograph the gel using the Foto/Phoresis I camera. (Lay a ruler under the field to be photographed.)
   Transmitted white light illumination (no filters)
   Polaroid 668 Color Film
   2 seconds @ f32

Western Blot Protocol

Trypsin protein content in periplasmic extracts was routinely determined using the following protocol.

PAGE-SDS Gels:

1. Pour two gels. -12.5% Lowers -4.75% Uppers (40 μl wells).

2. Prepare the samples (if this has not already been done).
   -Mix 20 μl of 4x Laemmli[b] with 60 μl of each sample.
   -Boil 10 minutes.

3. a. Load half of the sample on each gel.
   b. Electrophorese the samples for 1.5hours at 10watts (constant power) per gel.[c]

Western Blot:

1. a. Blot one gel to nitrocellulose for 30 minutes at 15 Volts[d] in the Genie Blot.[e]
   b. Stain the other gel with Coomassie Blue stain.

---

[a] 30 minutes is sufficient to separate the proteins enough to give a signal which can easily be assigned to the well from which it came. Runs of 90 minutes give better resolution of size, but the identification of each signal becomes more difficult, especially for samples near the edge of the gel.

[b] 4x = 4ml β-mercaptoethanol, 9.2ml 20% SDS, 5ml 1M Tris, pH 7, 1.32ml 0.3% bromphenol blue in ethanol, 2g ficoll (Dr. Lilia Maria Babé, *personal communication*)

[c] 2 hours at 12.5watts for gradient gels

[d] An EC-420 power supply from E-C Apparatus Corp, 3831 Tyrone Boulevard N, St. Petersburg, FL (813) 344-1644 was used to maintain this potential.

[e] This apparatus is available from Idea Scientific Company, Box 2078, Corvallis, OR 97330, (503) 758-0999

2. Block the nitrocellulose blot with 10% Dry Milk in 1xTBS,[f] 0.1 % Triton X-100 30min @ 37°C.
3. Probe the blot with rabbit anti-trypsin antibodies[g] 30min @ 37°C.
4. Wash the blot three times with 50mls of 1x TBS, 0.1% Triton X-100.
5. Probe the blot with goat anti-rabbit IgG-HRP[h] (30min @ 37°C).
6. Wash the blot twice with 50mls of 1x TBS, 0.1% Triton X-100, then once with 1x TBS.
7. Develop the Western with the following solution[i]
   In a 150ml beaker combine
   -60mgs of 4-chloro-1-napthol with
   -20mls of MeOH.
   -Adjust to 100mls with 1x TBS, and add
   -35 μl of 30% H_2O_2.

Miniprep DNA Protocol

Miniprep DNA was routinely preapred for sequencing by this modification of the modification by D. Ish-Horowicz[141] of the method of Birnboim and Doly[26])

Cultures:

Grow 2.5ml cultures overnight or ~8 hours in 17x100mm culture tubes at 37°C x 300 rpm under conditions to select for the plasmid.

Set Up:

1. Chill **Solution III**[a] on ice.
2. Prepare **Solution II**[b] during Solution I incubation.

Protocol:

1. a. Pellet 1.5ml of the culture in 1.5ml eppendorf tube in microfuge, 1 min at the maximum available speed (~12,000rpm).
   b. Decant the supernatant; microfuge the sample quickly.
   c. Aspirate and discard the supernatant.
2. a. Resuspend the cell pellet in 100 μls of **Solution I.**[c]
   b. Incubate the suspension for 5 minutes at room temperature.

---

[f] 10x TBS is 1.54M NaCl, 100mM Tris, pH 7.4

[g] 100μl of stock antibodies to 10mls of 3%BSA, 0.1% Triton X-100 , 1xTBS

[h] 10μl of stock antibodies to 10mls of 3%BSA, 0.1% Triton X-100 , 1xTBS

[i] The following reagent is 50x more sensitive: In 90ml 50mM Tris, pH 7.6 dissolve 0.06g 3,3'-diaminobenzidine tetrachloride; add 10ml 0.3% $NiCl_2 \cdot H_2O$. Immerse the blot in this solution and add 100μl H_2O_2. Develop for 1-2 minutes.

[a] 3M K^+ 5M ^-OAc: to 60ml of 5M KOAc add 11.5ml of glacial acetic acid and 28.5ml of H_2O

[b] Add 0.2ml 10M NaOH to ~8mls gd, then add 1.0ml 10% SDS, and adjust the volume to 10mls with glass-distilled water.

[c] 50mM Glucose, 10mM EDTA, 25mM Tris pH 8.0

3. a. Add 200 µls of **Solution II** and mix gently by inverting the tube.
   b. Incubate the preparation on ice 5 minutes **(less than 10 minutes!!!)**
4. a. Add 200 µls of ice cold **Solution III** and mix gently by inverting the tube.
   b. Incubate the preparation on ice 15 minutes.
5. a. Microfuge the preparation for 5 minutes at the maximum speed.
   b. Transfer the supernatant to a fresh tube.
6. a. Precipitate the DNA by adding 1000 µls of room temperature 100% EtOH.
   b. Incubate the preparation at ambient temperature for 2 minutes.
7. a. Microfuge the preparation at maximum speed for 2 minutes.
   b. Aspirate and discard the supernatant.
8. a. Resuspend the pellet in 300 µls of TE (10mM TrisHCl, 1mM EDTA, pH 8.0).
   b. Add 150 µls of 7.5M ammonium acetate.
   c. Incubate on ice 20 minutes.
   d. Pellet protein by microfuging at maximum speed for 5 minutes.
9. a. Transfer supernatant to a fresh tube.
   b. Add 900µl of -20°C ethanol.
   c. Incubate on ice 10 minutes
   d. Pellet the plasmid DNA by centrifuging at maximum speed for 5 minutes.
   e. Aspirate and discard the supernatant.
10. a. Resuspend the pellet in 300 µls TE containing 20 µg of RNAse (boiled, DNAse-free).
    b. Incubate at 37°C for 30 minutes.
11. Phenol/chloroform extract:
    a. Add 300 µls of buffered phenol and vortex to homogeneity.
    b. Add 300 µls of chloroform:isoamyl alcohol (24:1) and vortex to homogeneity.
    c. Microfuge the extraction 5 minutes.
    d. Aspirate the upper aqueous layer to a fresh tube.
    e. Repeat steps a-d until the interface is clean.
12. Chloroform extract the solution of DNA as in steps a-d above, except microfuge only 30seconds.
13. Ethanol precipitate the DNA as follows:
    a. Add 30 µl of 3M NaOAc (0.1 volume).
    b. Add 670 µls of -5 °C 100% EtOH (~2.5 volumes).
    c. Mix, and incubate the preparation on ice 10 minutes.
    d. Microfuge 30 minutes at max. speed.
    e. Aspirate and discard the supernatant.
    f. Rinse the pellet with ~100 µls of -5°C, 70% EtOH (add and aspirate).
14. Dry the DNA in a speed vac.[d]
15. Resuspend the resulting pellet in 30 µl TE.

---

[d] 5 minutes under a reasonable vaccuum (1-2 Torr) is usually sufficient.

16. Analyze a 3 μl sample on a 1% agarose gel in TAE.[e]
    a. Pour a minigel.
    b. Add the 3 μl sample to 2 μl of 5x dye[f] and 5 μl glass distilled water.
    c. Heat the sample at 65°C for 10 minutes.
    d. Load the whole sample and electrophorese at 200V for 30 minutes.
    e. Stain the gel in a 500ml solution of 1.4 μg/ml ethidium bromide in 1x TAE for 10 minutes.
    f. Destain the gel in a 500ml solution of 1x TAE for 10 minutes.
    g. Photograph the gel using the Foto/Phoresis unit.
       Conditions: Polaroid 667 Film, UV illumination, UV filter, orange filter: 1 sec @ f5.6; Develop 30 seconds

Genetic Purification

Mutant trypsin genes were detected in clones of *E. coli* transformed with the mutagenesis reactions by either trypsin activity or protein production and always by the preparation and analysis of the miniprep DNA from these clones. To ensure the genetic purity of the mutant, the plasmid DNA from primary isolates was re-transformed into an expression host. This was either done directly or ideally (from a genetic purity point of view) involved subcloning of the mutation into another plasmid which was transformed into an expression host. Therefore, either the ligation reactions used to prepare pT_3S214E and pT_3S214K mutant plasmids or miniprep mutated p^mT_3 plasmids were transformed into the *lac* i^q *E. coli* strain X90.[6] Six transformants of each plasmid were screened for trypsin protein production by antibody hybridization to SDS-PAGE-fractionated periplasmic fractions which had been blotted to nitrocellulose (western blot) as described.[227] An expression clone was selected and plasmid DNA was prepared from 100 ml of saturated culture which had been inoculated from the 40% glycerol stock of this clone. The trypsin gene in this plasmid DNA was sequenced as described below to verify the presence of the mutation and the integrity of the trypsin gene. All of the experiments reported here use protein prepared from the glycerol stocks of the original expression clones.

---

[e] TAE (40mM Tris acetate, 1mM EDTA) 242g Tris Base, 57.10ml Glacial Acetic Acid, 100ml 0.5M EDTA, pH 8.0/1L =50X

[f] 25% Ficoll, 0.5% SDS, 25mM EDTA, 0.05% bromphenol blue.

DNA Sequencing

The complete DNA sequence of each mutant gene was determined from the alkaline-denatured double-stranded expression plasmid using dideoxynucleotide chain termination[180] of replicative polymerization of deoxynucleotide triphosphates catalyzed by either chemically or genetically modified T7 polymerase.[217]

Midiprep DNA Protocol

To obtain enough plasmid DNA to sequence the trypsin gene, it was necessary to scale up the miniprep DNA protocol above.

Cultures

Inoculate 100 ml of selective medium in a 250 ml flask and grow overnight with aeration at 37°C.

Set Up:

1. Chill **Solution III**[a] on ice.
2. Prepare **Solution II** during Solution I incubation.

Protocol:

1. a. Pellet cultures in 50ml centrifuge tubes in SS-34 rotor for 5 minutes at 6,500 rpm (5096 x *g*).[b]
   b. Decant, bleach, and discard the supernatant.
2. a. Resuspend the cell pellets in 5 mls of **Solution I.**
   b. Incubate the suspension for 5 minutes at room temperature.
3. a. Add 10 mls of **Solution II** and mix gently by inverting the tube.
   b. Incubate the preparation on ice 5 minutes **(less than 10 minutes!!!)**
4. a. Add 10 mls of ice cold **Solution III** and mix gently by inverting the tube.
   b. Incubate the preparation on ice 15 minutes.
5. a. Centrifuge the preparation for 15 minutes in an SS-34 rotor at 6,500 rpm (5096 x *g*).
   b. Transfer the supernatant to a fresh tube by filtering through cheesecloth if necessary [c]
6. a. Precipitate the DNA by adding an equal volume of isopropanol.
   b. Incubate the preparation at ambient temperature for 20 minutes.

---

[a] Solutions I, II, and III are the same as used in the miniprep protocol above.
[b] Use Corning 25330 (orange capped tubes).
[c] Filtering through cheesecloth reduces the yield of plasmid (Jeff Higaki, personal communication).

c. Centrifuge the preparation in an SS-34 rotor for 15 minutes at 6,500 rpm (5096 x *g*).
d. Decant and discard the supernatant.

7. a. Resuspend the pellet in 8 mls of TE (10mM TrisHCl, 1mM EDTA, pH 8.0).
b. Add 4 mls of 7.5M ammonium acetate.
c. Incubate on ice 20 minutes.
d. Pellet protein by centrifuging 15 minutes at 6,500rpm in SS-34 rotor (5096 x *g*).

8. a. Transfer the supernatant to a fresh tube.
b. Add two volumes of -20°C ethanol.
c. Incubate on ice 20 minutes
d. Pellet the plasmid DNA by centrifuging 15 minutes at 6,500rpm in SS-34 rotor (5096 x *g*).
e. Decant and discard the supernatant.

9. Resuspend the pellet in 5 mls of TE containing 75 μg of boiled RNase at 37°C for 30'.

10. Phenol/chloroform extract the solution of DNA as follows:
a. Add 5 mls of buffered phenol and vortex to homogeneity.
b. Add 5 mls of chloroform:isoamyl alcohol (24:1) and vortex to homogeneity.
c. Centrifuge the extraction to speed the separation of the phases and pack debris at the interface of the phases.[d]
d. Aspirate the upper aqueous layer to a fresh tube.
e. Repeat steps i-iv until the interface is clean.

10' a. Combine aqueous layers into a single 15ml conical centrifuge tube.
b. Butanol extract as in **Step 11** to reduce the volume to 3 mls.
c. Continue phenol:chloroform extractions on 3ml scale.

10" a. Butanol extract as in **Step 11** to reduce the volume to 300 μl.
b. Continue phenol:chloroform extractions on 300 μl scale in 1.5ml eppendorf tubes.

11. Reduce the volume of the aqueous phase by extracting with butanol as follows:
a. Add an equal volume of butanol and vortex to homogeneity.
b. Centrifuge in a table-top clinical centrifuge at maximum speed for 5 minutes.
c. Aspirate and discard the upper, saturated butanol layer.
d. Repeat steps i-iii to a final volume of 300 μl in a 1.5ml eppendorf tube.[e]

12. Chloroform extract the solution of DNA as in step 10 above except centrifuge 30 seconds.

13. Ethanol precipitate the DNA as follows:
a. Add 30 μl of 3M NaOAc (0.1 volume).

---

[d] This may be done in a clinical table-top centrifuge at maximum speed for ~10 minutes, although better packing is accomplished in the SS-34 rotor at 6,500 rpm (5096 g) for the same amount of time.

[e] Larger fractional volumes of butanol may be used to speed the process. Adding too much butanol, however, may result in a single phase which can be separated by adding 5 M NaCl.

b. Add 670 μls of -5°C 100% EtOH (2.5 volumes).[f]
c. Mix and incubate the preparation on ice 10 minutes.
d. Microfuge 30 minutes at maximum speed.
e. Aspirate and discard the supernatant.
f. Rinse the pellet with ~100 μls of -5°C, 70% EtOH.

14. Dry the DNA in a speed vac.[g]
15. Resuspend the resulting pellet in 500 μl TE.
16. Analyze a 2 μl sample on a 1% agarose gel (TAE).[h]

DNA Sequencing Reactions

Sequencing reactions were carried out in microtiter plates using the following protocol, except that reagents in the microplate wells could also be mixed after being spotted on the sides of wells by tapping the plate on the bench top.

^{35}S-Double-stranded Sequencing with Sequenase™[a],[226]

Sequencing Solutions

1. **Double-stranded template DNA**: 2 μg in 27.5 μl H_2O
2. **Primer**: 10 μl of 3 ng/μl in H_2O (0.5 - 2 μM of a 24mer).
3. **5M NaOH**
4. **5x Sequenase Buffer:**

   200 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 250 mM NaCl
   -2 ml 1M Tris-HCl, pH 7.5
   -0.5 ml 1M $MgCl_2$
   -0.5 ml 5M NaCl
   -7.0 mls glass-distilled water

5. **0.1M DTT** : 15.4mg DTT to final volume of 1ml with H_2O
6. **5x (dATP) Labeling Mix**: (7.5 μM dGTP, 7.5 μM dTTP, 7.5 μM dCTP)
   -Dilute 1:5 with H_2O (i.e. 4 μl 5X labeling mix + 16 μl H_2O for 6 template rxns)
7. Termination Mixes:
   ddGTP : 80 μM deaza-dGTP, 80 μM dATP, 80 μM dTTP, 80 μM dCTP, 8 μM ddGTP, 50mM NaCl
   ddATP : 80 μM deaza-dGTP, 80 μM dATP, 80 μM dTTP, 80 μM dCTP, 8 μM ddATP, 50mM NaCl

---

[f] This volume has been slightly adjusted to facilitate removal of the supernatant from the resulting precipitated DNA.

[g] 5 minutes under a reasonable vacuum (1-2 Torr) is usually sufficient.

[h] TAE (40 mM Tris acetate, 1 mM EDTA) 242 g Tris Base, 57.10 ml Glacial Acetic Acid, 100 ml 0.5 M EDTA, pH 8.0/1L =50X

[a] From the Master Protocol Book of the Craik Lab by Jeff Higaki after the Sequenase Booklet and a protocol by Michelle Browner of the Fletterick Lab

ddTTP : 80 μM deaza-dGTP, 80 μM dATP, 80 μM dTTP, 80 μM dCTP, 8 μM ddTTP, 50mM NaCl
ddCTP : 80 μM deaza-dGTP, 80 μM dATP, 80 μM dTTP, 80 μM dCTP, 8 μM ddCTP, 50mM NaCl

8. **TE** (10mM Tris-HCl, pH 7.5,1.0mM EDTA) : Ice cold
9. **6x Sequenase** (14U/ μl) : Dilute 1:6 with cold TE, pH 7.5.
10. **^{35}S-dATP** (1 μCi/ μl) : Dilute 10 μCi/μl stock 1:10 with di.

Gel Stock Solution[b]

-30ml 38% Acrylamide, 2% Bis-acrylamide
-20ml 10x TBE[c,d]
-70ml gd
-84g Urea
-Adjust to 200mls with gd

Per Gel:

-Degas 70ml of Stock per gel.
-Add 105 μl 10% ammonium persulfate
-Add 105 μl TEMED

-Pour the gel.

Procedure

1. Dissolve or dilute in 1.5ml Epp. tube, 2 μg ds Template in final volume of 27.5 μl with H_2O.
2. a. Add 2.5 μl 5M NaOH.
   b. Incubate at 37°C for 15 minutes
3. a. Add
      -10 μl (~60ng) primer
      -9 μl 3M NaOAc, pH 5.2
      -200 μl 100% EtOH (-5°C)
   b. Incubate on ice for 10 minutes.
   c. Microfuge 15 minutes at the maximum available speed.
   d. Aspirate and discard the supernatant, and wash the pellet with 70% EtOH.
   e. Speed Vac to dryness (5 minutes at a reasonable vacuum, e.g. 1-2 Torr is sufficient).
4. a. Resuspend the pellet in 7.9 μl H_2O.
   b. Add 3.6 μl 5x Sequenase buffer.
   c. Add 1.5 μl 1 μCi/μl ^{35}S-dATP.
   d. Incubate at 37°C (or room temperature) for 15' (as long as it takes to prepare the extention reaction mix.)
5. a. **Extension reactions:**
   To each reaction add

---

[b] This solution is good for seven days at 4°C.

[c] 10xTBE = 60.55g Tris Base; 27.5g Boric Acid; 4.65g $Na_2EDTA \cdot 2H_2O$ (25mls 0.5M pH 8.0 Stock)/500mls

[d] 10ml of 10x TBE for tall, narrow gels (~24" x 8")

-1.0 μl 0.1 M DTT
-2.0 μl 1x Labelling mix (Dilute 5X Labeling mix with H_2O).
-2.0 μl 1X Sequenase (Dilute stock sequenase 1:6 with TE)

b. Incubate at room temperature, 8-10 minutes.

6. Spot 2.5 μl termination mixes on one side of the microtiter plate wells.
7. **Termination reactions.** (After extension reactions are complete)
   a. Spot 3.5 μl extension reactions on one side of the microtiter plate wells.
   b. Spin @ 1K.
   c. Incubate at 37°C for 20 minutes.
8. Add 4 μl stop mix and freeze until the sequencing gel can be loaded and run.
9. Heat the samples at 95°C, 4 minutes without the microtiter dish cover to reduce the volume.
10. a. Load 2 μl per well on gel.
    b. Run gel @ 55W, constant power (50°C) with ~7.5x TBE in the lower reservoir.
11. a. Dry gel onto Whatman 3mm paper.
    b. Autoradiograph dried gel.

Reading Sequencing Gels

The DNA gels were read using a personal computer similarly to a published method.[164] Four adjacent buttons (Q, W, E, and R) on a computer keyboard were used to represent the G, A, T, or C lanes of the gel. This substitution allows the gel reader to concentrate on the gel with one hand typing blindly on the keyboard. The band pattern was read by typing the keyboard letter corresponding to the gel band into a new file of a text editor such as Microsoft WORD 4.0. After the DNA sequence was transcribed, the "**Change...**" feature of the text editor was used to replace all "q" characters with "G", "w" with "A", *etc.* In particular, headings could be typed in the ALL CAPS font of WORD and would not be altered if the check box "**Ignore Words in All Caps**" in the "**Change...**" command set-up box is used. Thus the several sequencing reactions for each gene could be typed into one computer file with ALL CAPS headings indicating the putative gene, template preparation, sequencing primer, gel, and date. The typed sequence is then compared with the known sequence. This method allowed rapid, unbiased gel reading.

Protein Purification

Cultures

Heterologous Expression System. Mutant rat anionic trypsins were expressed in *Escherichia coli* by subcloning the 327 bp Xho I/Sty I fragment of the trypsin gene pBSTn2-S214E or -S214K plasmids[62] into the trypsin gene in the expression plasmid pT_3[101] or were expressed directly from mutated p^mT_3 plasmids.[63] The expression plasmids pT_3 and p^mT_3 encode a *his* J signal peptide-trypsin fusion protein under control of the *tac* promoter (Figure 12). The p^mT_3 plasmid is a derivative of pT_3 which contains an M13 origin of replication to facilitate the recovery of single-stranded DNA (*see* Directed Mutagenesis: Practice). The as for the *phoA* signal peptide,[227] the *his* J signal peptide is cleaved by *E. coli* signal peptidase to liberate the proper amino terminus of trypsin which is critical for activity.

High Density Culture of *E. coli*. Expression of trypsin from the pT_3 expression plasmid was found to be most successful under constitutive derepression of the *tac* promoter.[101] A 2 ml aliquot of LB-amp was inoculated from a -20°C glycerol stock of the *E. coli* expression clone. The culture was propagated at 37°C with vigorous shaking and expanded as follows: in 2 mls for 2-4 hours in 17x100 mm culture tubes, in 60 mls for 1-2 hours in a 250 ml culture flask, in 400 mls for 1-2 hours in a 2 l culture flask. The 400 mls of LB-amp culture was then used to inoculate a fermentor vessel containing 10 l of sterile LB-amp containing 1mM IPTG. The 10 l culture was grown for approximately 14 hours with vigorous stirring at 37°C, pH = 7.4, and pO_2 = 30% where pO_2 = 100% when the media is saturated with air at 1 atmos pressure. A solution of 50% glucose containing 0.5 g/l ampicillin was added at 0.7 ml/min during the growth of the fermentor culture. The cultures reached final concentrations of OD_{600} = 28 – 30.

Large Scale Culture of *E. coli* Expressing Rat Anionic Trypsins

E. coli cells harboring plasmids which encode rat anionic trypsin mutants were cultured aerobically in a BIOSTAT E[a] fermentor according to the following protocol.

Media:

1. Prepare 5X LB media. in the fermentor
   a. To 8.5L of stirring tap-distilled water, add
      - 500g of Tryptone
      - 250g of Yeast Extract
      - 100g of NaCl
   b. Adjust to 9L with tap distilled water.
2. Sterile filter 1 L of 50% glucose containing 500 mg of ampicillin
3. Autoclave the fermentor and media for 30 minutes at 120°C.
4. Prepare an approximately 10 ml solution of primary additions as necessary:
   Ampicillin: - 500mg
   IPTG: - 2.38g
5. After the media has cooled, sterile filter the primary additions solution into the fermentation chamber.

Cultures:

1. a. Grow a starter culture in 2.5ml L-amp in 17x100mm blue capped culture tubes.
   b. Also start a **No Inoculum Control.**
   c. Grow 6-8 hours at 37°C, 300rpm.
2. a. Inoculate 50mls of L-amp in a 250ml culture flask with 2 mls of the starter.
   b. Grow 2-3 hours at 37°C, 300rpm.
3. Inoculate 500mls of LB containing 50μg/ml ampicillin in a 2 L flask.
4. a. Inoculate 10L of media containing 50μg/ml ampicillin, 1mM IPTG.
   b. Grow overnight at 37°C, 800rpm, pH 7.4, 0.7ml/min glucose additions.
5. Isolate periplasmic fractions.

Periplasm Preparation

The cells were pelleted at 9,000 x *g* for 6 minutes and then resuspended with a total volume of 750 mls of 25% sucrose, 10 mM Tris, pH 8.0. The periplasms were prepared by adding 80 ml of 15 mg/ml lysozyme, 20 mM EDTA and shaking the preparation gently at 37°C for 30 minutes. The spheroplasts were pelleted by centrifugation at 6,000 x *g* for 30 minutes.

---

[a] B. Braun Diessel Biotech GMBH, P. O. Box 120, D-3508 Melsungen, Germany

Preparation of Periplasmic Fractions

Periplasmic fractions of fermentor cultures of *E. coli* were routinely prepared according to the following protocol.

Set Up: (Standard Preparation for 300OD-L[a])

1. Prepare 1000 ml of a solution of 25% Sucrose in 10mM Tris pH 8.0.
   -Dissolve 250g of sucrose in ~800 mls gd.[b]
   -Add 10 mls of 1.0M Tris, pH 8.0.
   -Adjust the volume to 1000 mls by the mark on the beaker.
   -Store on ice. -Sterile filter any not used.
2. Prepare a solution of 2 mg/ml lysozyme in 40mM EDTA, pH 8.0.[c]
   -Dissolve 0.3g of lysozyme (Sigma L-6876) in 146mls gd.
   -Add 12.8ml of 0.5M EDTA, pH 8.0.
   -Store on ice.
3. Prepare a stock solution of 0.5M Na Citrate, pH 2.00[d]
   -Dissolve solid Na citrate to make 0.5M of the desired volume.
   -Add < the theoretical number of HCl equivalents.
   -pH to 2.00 with concentrated HCl.
   -Autoclave 30 minutes and store at 4°C.

Protocol:

1. a. Pellet the cultures in several lots in 6 x 500ml bottles in a GS-3 rotor, 6 minutes at 7,500 rpm (9,000g).
   b. Decant, bleach, and discard the supernatants.
2. At 0°C, resuspend each of the final pellets in 150 mls of 25% sucrose, 10mM Tris, pH 8.0.
3. a. Add 26ml of 2 mg/ml lysozyme in 40mM EDTA, pH 8.0 to each bottle.
   b. Incubate the preparation with gently shaking at room temperature for 1 hour.
4. a. Pellet the spheroplasts by centrifuging the preparation in the GS-3 rotor for 30 minutes at 6,000rpm (6089g).
   b. Decant and save the supernatants.
5. Start a CM column.
6. Dialyze the supernatant (~1000 mls) against 12L of 10mM Na Citrate, pH 2.8 (ambient pH) at 4°C overnight.[e]

---

[a] 30 OD_{600} x 10L = 300OD-L)

[b] glass-distilled water

[c] This solution may be prepared immediately while the cultures are being pelleted.

[d] Dilutes to pH 2.7 at 10 mM.

[e] 12L x 10mM Na Citrate, pH 2.8 = 240mls of 0.5M Na Citrate, pH 2.00 in 12.0L

Carboxymethyl (CM) Cellulose Cation Exchange Chromatography

The periplasm prepared above was dialyzed overnight at 4°C against 12 l of 10 mM sodium citrate, pH 2.8 buffer, which is prepared from an aqueous solution of trisodium citrate dihydrate and concentrated hydrochloric acid. A white precipitate was pelleted from the dialyzed retentate at 13,700 x *g* for 20 minutes and the supernatant loaded at room temperature at ~10 ml/min onto a 2.5 x 20 cm CM Sepharose column equilibrated in the sodium citrate buffer. The column was washed with 2 l of sodium citrate buffer and eluted with a 500 ml 0 - 0.5 M NaCl gradient in sodium citrate buffer. Trypsin protein and the endogenous *E. coli* trypsin inhibitor, ecotin,[a,44,147] were localized by western blot analysis as described.[227] Fractions which contain trypsin, but not ecotin, were pooled and dialyzed against 2 x 12 L of 10 mM MES, pH 6.0 buffer.

CM Column Protocol

Periplasms were fractionated using carboxymethyl (CM) cellulose chromatography according to the following protocol.

Sample Preparation:

1. Dialyze the sample 3x to pH 2.8 in 10mM Na Citrate, pH 2.8.
   12L of dialysis buffer is 240mls 0.5M Na Citrate, pH 2.00 to 12L with tap-distilled water

Column Set Up:

1. Pour a 2.5 x 20cm CM cellulose column (if necessary).
2. Equilibrate a CM column at ambient temperature. Sparge all buffers.
   a. Equilibrate the column with 1L of 0.5M NaCl, 10mM Na Citrate, pH 2.8 .
   b. Equilibrate the column with 1L of 10mM Na Citrate, pH 2.8.
3. As time permits, pour two (2) 12.5% PAGE-SDS separation gels with 4.75% stacking gels with 16-18 wells.

Protocol

1. Clarify the sample by centrifuging in the GS-3 rotor 10 minutes at 9,000rpm (13,700g).
2. a. Load the CM column at ~10ml/min. (Save a sample of the load).

---

[a] The original preparation of recombinant rat anionic trypsin which was used for the induction of polyclonal antibodies contained trace amounts of ecotin. Consequently, we have preparations of polyclonal antibodies for rat anionic trypsin which cross-react with ecotin (unpublished results).

b. Collect the flow through at ambient temperature.

3. a. Wash the column with 2.0L of 10mM Na Citrate, pH 2.8 at ~10ml/min.
   b. Monitor the A_{280} of the effluent. The column may be eluted at an A_{280} ≤0.010. (This may take 0.8-2.0L.)
4. a. Elute the column with a 500ml gradient: 10mM Na Citrate, pH 2.8; 0-0.5M NaCl
   b. Collect 270 drop fractions (~9mls) at ~1.5ml/min.
5. Determine the A_{280}'s of the fractions.
6. Regenerate the resin in the column.
   a. Wash with 0.5M NaOH for 30 minutes
   b. Wash with deionized water for 30 minutes.
   c. Wash with 0.5M HCl for 30 minutes.
   d. Wash with deionized water for 30 minutes.
   e. Wash with 1L of 10mM Na Citrate, pH 2.8.
   f. Wash with 500mls of sparged 10mM Na Citrate, pH 2.8 containing 20% ethanol.

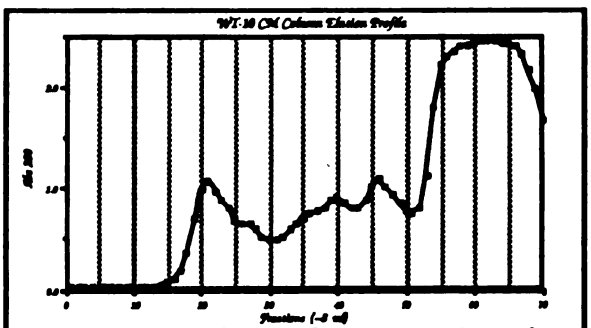


Figure 15. CM Column Elution Profile of Periplasmic Extracts of *E. coli* Expressing Trypsin. Acid precipitated periplasmic extracts were clarified by centrifugation and loaded as indicated in the text. The column was washed and proteins were eluted using a NaCl gradient. A variety of profiles are obtained and western analysis should be carried out to test the whole range of fractions.

5. Select fractions spanning the total elution of the column, but select fractions to help define the anticipated ends of the trypsin peak.
   a. Prepare 80 µl samples in 4X Laemmli buffer for Coomassie and Western analysis. (10 minutes at 110°C)
   b. Prepare 40 µl samples in 4X Laemmli buffer without β-mercaptoethanol for Activity analysis. (3 minutes at 65°C)
6. a. Analyze the chosen fractions for Activity,
   b. Trypsin and Ecotin Content (Western), and
   c. Protein Content (Coomassie).

Data Transfer between the HP 8451A Spectrophotometer and a Macintosh Computer

The amount of protein in each fraction from both CM-Sepharose and BPTI-affinity columns was routinely determined by measuring the absorbance at 280 nm (A_{280}) using a Hewlett-Packard[a] HP 8451A Diode Array Spectrophotometer equipped with an HP-85 microprocessor, dual 3.5" disk drives, and an RS-232C port. The program "DATACOLLECT"

[a] Hewlett-Packard, 1000 N. E. Circle Blvd., Corvallis, OR 97330

was used to write the A_{280} data to the file "ABSDATA" on a disk. The file "ABSDATA" can accommodate two hundred A_{280} determinations. The program "DATACOLLECT" starts the program "DATATRANSFER" which reads the file "ABSDATA" and outputs it through the RS-232C port. The terminal emulator program Versaterm PRO[a] is used to drive the Macintosh to act as a VT100 terminal and to store the data stream from the spectrophotometer as a Cricket Graph file using the **"Save stream..."** protocol. Data was routinely transferred from the HP Spectrophotometer at 2400 baud, 7 data bits, odd parity, 1.0 stop bit using the Xon/Xoff handshake protocol.

```
100   ASSIGN# 1 TO "ABSDATA:D701"
110   DISP "You should already have REFERENCED your cuvette at
      280 nm"
130   DISP "Press EXECUTE to continue"
150   DISP "Press STOP to halt program"
170   INPUT A$
180   PRINT A$
190   PRINT "ABSORBANCES >10 ARE NOT STORED"
210   PURGE "ABSDATA:D701"
220   CREATE "ABSDATA:D701", 200, 8
230   ASSIGN# 1 TO "ABSDATA:D701"
240   DISP "Press EXECUTE to MEASURE and store the
      measurement"
260   DISP "Type DONE (carriage return)"
270   DISP "when you are finished"
290   INPUT A$
300   IF A$ = "DONE" THEN 360
310   MEASURE
320   IF NMEAS = 0 THEN 320
330   IF VALUE(280) >10 THEN GOTO 240
340   PRINT# 1 ; VALUE(280)
350   GOTO 240
360   ASSIGN# 1 TO *
370   DISP "FILE -- ABSDATA:D701 -- CLOSED"
380   DISP "PROGRAM HALTED AT YOUR REQUEST"
400   CHAIN "DATATRANSFER:D701"
410   END
```

Figure 16. "DATACOLLECT" for Acquiring A_{280} Data.

```
100   DISP "Be sure that"
110   DISP "Versaterm PRO"
120   DISP "is on John's Mac"
130   DISP "and that the cable is connected"
150   DISP "Press EXECUTE to continue"
160   INPUT A$
170   DISP "Versaterm SETTINGS SHOULD BE"
180   DISP "2400 BAUD"
190   DISP "Xon/Xoff"
200   DISP "ODD PARITY"
210   DISP "7 BIT CHARACTERS"
220   DISP "1.0 STOP BIT"
230   DISP "Hit EXECUTE to continue"
240   INPUT A$
250   REM  SET THE BAUD RATE
260   REM  IN REGISTER #3
270   REM  11 = 2400 BAUD
280   CONTROL 10, 3 ; 11
290   REM  SET XON/XOFF HANDSHAKE
300   CONTROL 10, 11 ; 192
310   CONTROL 10, 14 ; 19
320   CONTROL 10, 15 ; 17
330   CONTROL 10, 5 ; 16
340   REM TRANSFER DATA
350   ASSIGN# 1 TO "ABSDATA:D701"
360   DISP "TRANSFERRING DATA"
370   DISP "Just a note:"
380   DISP "ERROR 71 ON LINE XXX : EOF"
390   DISP "means END OF FILE"
400   FOR I = 1 TO 200
410   READ# 1, I ; J
420   OUTPUT 10 ; J
430   NEXT I
440   DISP "DATA TRANSFER SHOULD BE COMPLETE"
450   END
```

Figure 17. "DATATRANSFER" for Downloading A_{280} Data

Affinity Chromatography

Affinity Column Operation

Trypsin-containing CM fractions were pooled and dialyzed in 10mM MES, pH 6.0. The retentate from the dialysis was sterile filtered through a 0.22 μm filter and recycled at 4°C at 1.0 ml/min onto a 0.5 x 8 cm BPTI affinity column for at least 10 load volumes. The column was washed with 10

[a] Abelbeck, L.R. *Versaterm PRO* (Abelbeck Software, 1988)

mls of MES buffer. Then, while collecting 2 ml fractions, the column was washed with 35 mls of 0.5 M NaCl in MES buffer, then with 25 mls of MES buffer. The trypsin was eluted at ambient pH with 0.1M citric acid or, in later preparations, 0.5M NaCl, 0.1M citric acid, pH ~2 by loading 20 ml onto the column. The flow was then stopped for at least 30 minutes before 10 ml were eluted from the column. For some mutants, repeated cycles of stopping the column and eluting 10 ml volumes releases a significant amount of protein after the first elution. Each column is used for only one mutant.

Affinity Chromatography: BPTI Column Protocol

Sample Preparation:

1. Pool the CM fractions which contain trypsin and
2. Dialyze them against 12L of 10mM MES, pH 6.0. 12L of dialysis buffer is 240mls 0.5M MES to 12L with tap-distilled water.
3. Sterile filter the dialysate.

Protocol:

1. Equilibrate a small BPTI column with ~75mls of 10mM MES, pH 6.0.
2. Load the column by using a peristaltic pump to recycle the sample through the column.
3. Chase the load through with 20mls of 10mM MES, pH 6.0.
4. a. Begin collecting 62-drop (2ml) fractions at ~1ml/min.
   b. Wash the column with 35mls of 0.5M NaCl, 10mM MES, pH 6.0.
5. Chase the wash through with 25mls of 10mM MES, pH 6.0.
6. Elute the trypsin in 10ml steps with 0.5M NaCl., 0.1M citric acid, pH 2.0 buffer at ~1ml/min.
   a. Load 20mls of 0.5M NaCl., 0.1M citric acid on the column while collecting fractions.
   b. Stop the flow for 30 minutes.
   c. Run 10mls of 0.5M NaCl., 0.1M citric acid through the column.
   d. Repeat steps b and c, as desired.
   e. Wash through with 20mls of 10mM MES, pH 6.0.
7. Record and plot the A_{280}'s, and pH's of each fraction.
8. Wash the BPTI column with 50mls of 10mM Tris, pH 7.0, 0.2% NaN_3.
9. Locate the pure trypsin protein by western and Coomasie analysis.
10. a. Pool the Eluted Peak fractions.
    b. Sterile filter the pooled fractions. Rinse the filter thoroughly.
11. Concentrate the Eluted Trypsin pool in a Centriprep-10 to 200 μl.

12. Exchange the buffer for 1mM HCl by several washes in the Centriprep-10.
    -1st dilution (200 μl 0.1 M Citric acid to 15 mls = 1.3 mM)
    -2nd dilution (200 μl 1.3 mM Citric acid to 15 mls = 18 μM)
13. Transfer the Centriprep concentrate to a Centricon for final concentration.
14. Store the protein in 1 mM HCl at 4°C.

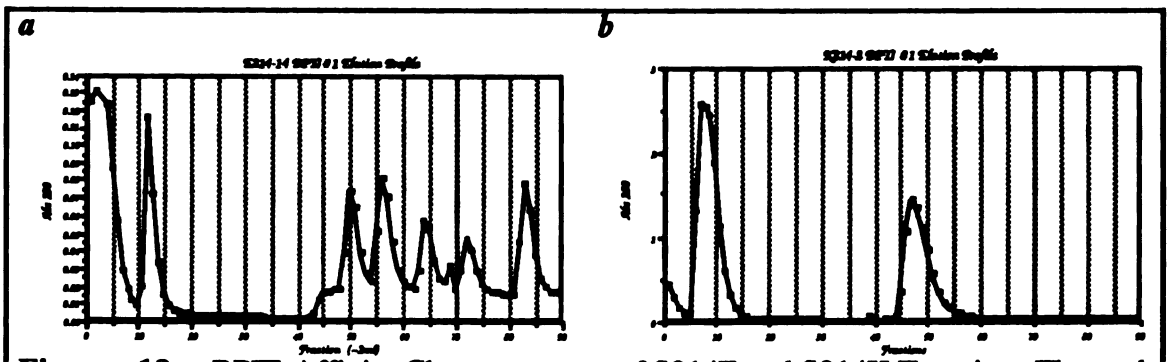


Figure 18. BPTI-Affinity Chromatograms of S214E and S214K Trypsins. The peak at ~12 ml is nonspecifically bound protein stripped with 0.5M NaCl, 10mM MES, pH 6.0. The column is washed free of NaCl during which time no protein elutes. Then repeated cylces of washing with 10ml lots of 0.1M citric acid and stopping the flow for 30 minutes results in the elution profiles shown. ***a.*** S214E Trypsin ***b.*** S214K Trypsin

Preparation of BPTI-Resin

BPTI was coupled to Affi-Gel 10 according to the manufacturer's instructions.[a] White crystalline BPTI (Trasylol™) (0.8 g) was dissolved in 12 mls of 0.1M MES, pH 6.0, in a 15 ml polypropylene centrifuge tube. Isopropanol was removed from a fresh 25 ml bottle of Affigel-10 by vacuum filtration. The resin was washed similarly three times with ice cold deionized distilled water with care not to dry and crack the resin bed. The resin was then transferred to the tube containing the solution of BPTI in MES buffer and incubated with gentle rotation at 4°C for ~3 hours. Ethanolamine•HCl (1M, 100 µl) was added to the reaction for ~2 hours to neutralize unreacted resin sites. The BPTI-resin is stored at neutral pH in 0.02% NaN_3 at 4°C.

Concentration

The protein content of the fractions was determined by measuring the optical absorbance at 280 nm of each fraction. The fractions containing purified trypsin were

[a] Biorad, Inc., Chemical Division, 1414 Harbour Way South, Richmond, CA 94804 (415) 232-7000

pooled, sterile filtered, and concentrated in a Centriprep-10 to ~200 μl. The buffer was exchanged for 1mM HCl by twice diluting the ~200 μl sample to ~15mls and reconcentrating in the Centriprep-10. The process of washing the sample from the Centriprep into the Centricon results in a further dilution of the citric acid buffer. The theoretical concentration of the citric acid buffer in the final preparation was approximately 1 μM. The protein was stored in 1 mM HCl at 4°C. This preparation was diluted with 1 mM HCl to prepare samples for kinetic analysis.

HPLC Purification of Affinity Purified Trypsin

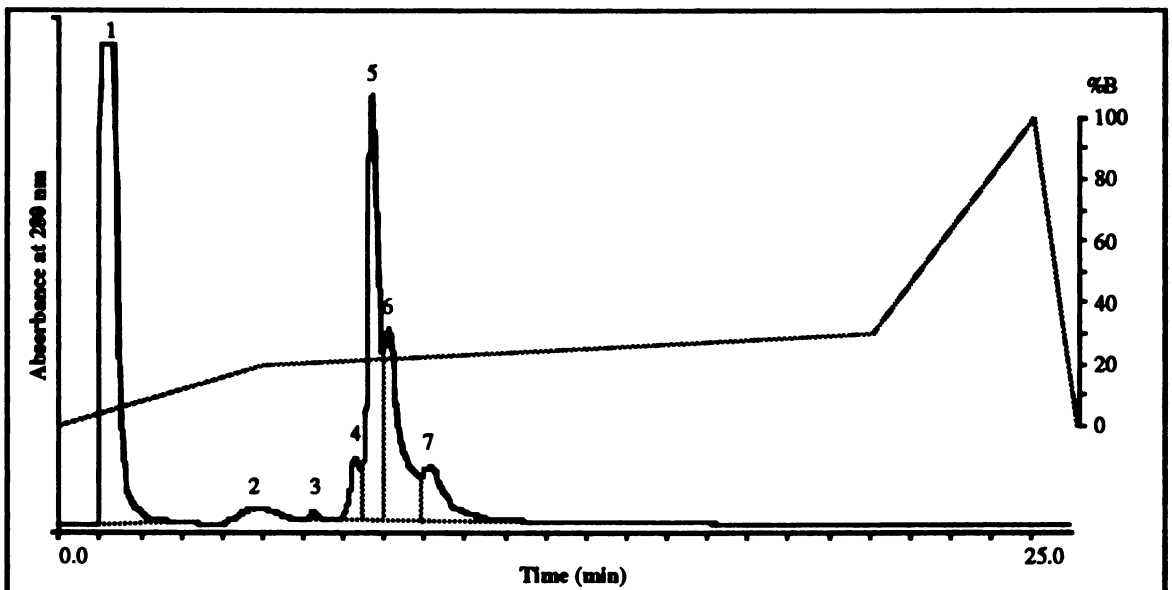


Figure 19. Strong Anion Exchange Chromatography of Affinity Purified Trypsin. Affinity-purified S214Q trypsin was further purified by strong anion exchange HPLC on a Dupont Bio Series Zorbax® SAX resin (6.2 x 80 mm) packed in a Dupont Reliance Cartridge. HPLC samples were prepared in 10 mg/ml benzamidine, 20 mM Tris, pH 7.0 and microfuged through nonsterile 0.2 μ Nylon-66 membranes.[a] Two 200 μl samples of ~850 μg of the protein were loaded and eluted at 1 ml/min with a 15 min 20–30% linear gradient of 1.0 M NaCl, 20 mM Tris, pH 7.0 (Solvent B) in 20 mM Tris, pH 7.0 (Solvent A). The column effluent was monitored at 280nm. Peak 1 contains benzamidine. Coomassie-stained PAGE gels (of S214E and wild type trypsin) indicate the following compositions: Peak 4, β-trypsin (uncut); Peaks 5 and 6, autolyzed trypsin; Peak 7, BPTI-trypsin complex. S214Q trypsin Peak 5 was crystallized (Figure 18).

HPLC purification[227] does not affect the kinetic parameters determined for either mutant or wild type trypsins but is used to generate protein for crystallization.[149] Samples

[a] Rainin Instrument Co., Inc., Mack Road, Woburn, MA 01801, (617) 935-3050 cat. #39-402

are prepared at neutral pH with benzamidine as an inhibitor. Fractions were collected by hand and made 10 mg/ml in benzamidine immediately.

Crystallization of Trypsin

Diffraction quality trypsin crystals were prepared essentially as described.[149] In the case of the HPLC-purified S214Q trypsin above, ~1.3 ml of the sample was concentrated and diluted twice at least 1:10 with 10 mg/ml benzamidine hydrochloride, 10 mM $CaCl_2$, 20 mM Tris hydrochloride, pH 7.0, using Centricon-10 microconcentrators at 5,000 x *g*.[146] Diffraction quality crystals were obtained reproducibly in ten weeks from hanging drops prepared as follows: On a silanized glass cover slip 1 µl of concentrated S214Q trypsin (~10 mg/ml in the exchange buffer described above) was added to 1 µl of 34% $MgSO_4$, and then 3, 4, or 5 µl of 25 mg/ml benzamidine hydrochloride was added and mixed well. The drop was suspended in an airtight system over a well containing 936 µl of 34% $MgSO_4$ and 64 µl of 156 mg/ml benzamidine hydrochloride. The crystal trays were incubated at 17°C. One or two crystals per drop appeared at approximately 8 weeks. When Dr. Mary McGrath broke the airtight seal, $MgSO_4$ rapidly crystallized in the drops. It was therefore necessary for her to quickly add several microliters of 10 mg/ml benzamidine hydrochloride to prevent $MgSO_4$ precipitation before mounting the crystals for collecting diffraction data.

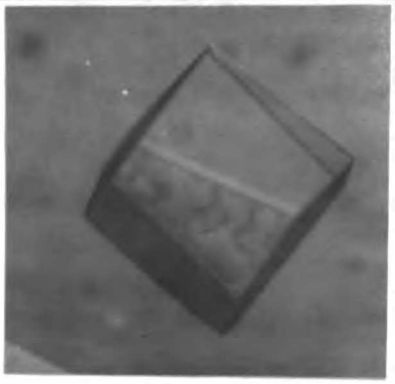


photo provided by Dr. Mary E. McGrath

Figure 20. Crystal of S214Q Trypsin The crystal is 300µm on the longest side.

Characterization of Mutant Trypsins

Biophysical Characterization of the Protein

Analytical Isoelectric Focusing

Approximately 10 µg total of protein standards and 5 µg per lane of each trypsin sample were electrophoretically focused at the isoelectric point at 10°C under nitrogen at 25

watts, constant power, for 50 minutes from a position 5 cm from the anode on a 14 cm LKB Ampholine® PAGplate, pH 4.0 - 6.5. Glutamic acid (0.1M), 0.5M H_3PO_4, was used as the anode buffer, and 0.1M β-alanine was used as the cathode buffer. The potential varied from 650 V to 2,370 V and the current from 35 mAmps to 9.5 mAmps. The gel was fixed 1 hour in ~100 ml 11.5% trichloroacetic acid, 3.45% sulphosalycylic acid and stained in 100 ml of 60°C Coomassie Blue R solution[a] for 10 minutes and destained[b] exhaustively.

Mass Spectral Characterization

Samples of affinity-purified protein in 1 mM HCl (~30 – 100 μg of sample which were ~60μM by active site titration) were submitted for mass spectrometry analysis. The relative molecular mass of each of the S214X mutants were determined to within ±2 atomic mass units by electrospray mass spectrometry (B. Gibson, S. Kaur, and B. Green, *personal communication*).

Structural Characterization

Crystal Structure Determination. The crystal structures of S214K, S214E, and S214Q trypsins were determined by Dr. Mary McGrath in Prof. Robert Fletterick's laboratory. X-ray diffraction data were measured as described[148] from single crystals. The position-214 and Trp-215 side chains were each fitted to difference electron density ($|F_O–F_C|$ maps are described below), and the structure was refined using the Hendrickson-Konnert algorithm as implemented in the program HKREF.[100]

Interpretation of Crystallographic Results[c]

Crystallographic data is an essential component of the complete analysis of the effects of mutations on the function of a protein. Two particular types of crystallographic

---

[a] Stain is a filtered solution of 1 g of Coomassie Brilliant Blue R solid in 225 ml methanol, 50 ml acetic acid, and 225 ml deionized distilled water.

[b] Strong destain is 45% methanol, 45% water, and 10% acetic acid; weak destain is 5% methanol, 7.5% acetic acid in water.

[c] The author is indebted to Dr. Chris Bystroff, Virginia Rath, and particularly Dr. Mary McGrath for any accurate statements in this section.

results have served as the basis for the conclusions in this thesis and so are described below.

X-ray Diffraction

The primary mechanism by which electrons scatter X-rays was first described by J. J. Thomson and is called Thomson scattering.[244] The periodic, electromagnetic nature of X-rays induces a periodic motion of electrons. In turn, the periodic motion of electrons produces electromagnetic radiation. In the diffraction experiment, an electron essentially absorbs an X-ray and re-emits it in all directions, and so the intensity of the scattered X-rays is proportional to the electron density.

In the X-ray crystallography experiment, X-rays are shined into the regular lattice of a protein crystal. The electrons of the protein molecules in the lattice scatter the X-rays as described above. The electromagnetic waves from such a system of point scatterers constructively and destructively interfere with each other giving rise to a particular, absolutely discrete distribution of emitted waves, the diffraction pattern.[150] Because the diffraction pattern depends on the number and the position of point scatterers, large, well-ordered crystals give rise to more useful diffraction patterns. Crystallography entails the collection, analysis, and deconvolution of the diffraction pattern to calculate the the electron density from which the X-rays were scattered.

Crystallographic Results

The philosophy of crystallography is simple: propose a model which accounts for the experimental observations. The "model" is the electron density for the proposed chemical structure of the protein, and the experimental observations are the spots in the diffraction pattern. The structure factors, ($\vec{F}$), are used to calculate the electron density map and are a function of the amplitude (F) and phase of the scattered x-ray waves.[a] A Fourier

[a] Each spot in the diffraction pattern is due to the wave which results from the constructive and destructive interference of all the scattered waves. And since interference, in general, is a function of the amplitudes and phases of the interfering waves, it follows that the amplitude and phase of each scattered wave contributes to each spot in the diffraction pattern.

transform converts the structure factors into an electron density map and vice versa.[83,150] Thus, a model structure can be used to calculate both amplitudes, F_c's,[a] and phases, ϕ_c's, of the waves which would be scattered. The square root of the intensity of a spot of the diffraction pattern is the observed amplitude of the resultant scattered wave, F_o. The phase of the resultant wave, ϕ_o, is not discernable from the observed data, so calculated phases, ϕ_c's, are used to calculate the structure factors, ($\vec{F}$), for any electron density map. The electron density maps presented in this thesis were calculated by Dr. Mary McGrath using information from a known rat anionic trypsin structure and the observed amplitudes derived from her own experimentally determined diffraction patterns.

Omit maps: $|F_o-F_c|$

The $|F_o-F_c|$ map, a difference map, is a direct comparison of the experimental data and the model. These are particularly useful for evaluating the position of a mutant side chain in an otherwise known structure. For example, the model may be constructed with the mutant side chain omitted; hence the name, "omit map." The positive contours of the omit map indicate electron density which is present in the observed structure but is missing from the model. Thus the position of the side chain for Lys-214 was derived from a model which had alanine at position 214. The negative contours of the omit map indicate electron density in the model which is absent in the observed structure. Difference maps can be used to highlight discrepancies between the model and observed data and therefore aid in refining the model.

Electron density maps: $|2F_o-F_c|$

Since phases are not generally observed, the structure factors, $\vec{F_o}$'s, used to calculate the electron density map based on the observed amplitudes, F_o's, would contain calculated phase data, ϕ_c's. Therefore, in order to calculate the "observed" electron density map, $|2F_o-F_c|$ amplitudes are used to increase the amount of observed data in the map. The

---

[a] Subscript "c" denotes a calculated parameter; subscript "o," an observed one.

$|2F_o-F_c|$ map can be thought of as an $[|F_o| + |F_o-F_c|]$ map, an $|F_o|$ map plus a difference map to correct for the calculated phases used in the "observed" structure factors, $\vec{F}_o$'s. The $|2F_o-F_c|$ electron density map can be superimposed on the model structure and visually inspected to see how well the model fits the data.

Quality of the Structure

The quality of the fit of the model to the data the *R*-factor calculated in (24).

$$R_{cryst} = \frac{\Sigma \ |F_o-F_c|}{\Sigma \ F_o} \qquad (24)$$

The lower the better. Petsko argues that ambiguity in the structure is best relieved by optimizing the *resolution* of the map (Gregory Petsko, *personal communication*). The selection for lower *R*-factors results in the publication of maps with lower *R*-factors, not better resolution. Higher resolution maps show more detail and are less subject to the subjective interpretations of the scientist interpreting the electron density map in terms of chemical structure.

Steady-State Kinetics Analysis

Determination of Reaction Progress

Reaction rates of thioester and amide substrates were followed as described under "**Enzyme Assays**" below. A Macintosh SE/30 computer was used to drive a Perkin-Elmer LS-5B Luminescence Spectrometer. This system allowed raw data to be recorded, stored, and used to calculate the Michaelis-Menten parameters directly and so is described below.

Interface of the LS-5B Luminescence Spectrometer with a Macintosh

Hardware

The LS-5B Luminescence Spectrometer, the "fluorimeter," is equipped with an internal microprocessor with an RS-232C communications port. A symbolic language is provided by the manufacturer[163] which allows control of the fluorimeter functions by input

through the RS-232C port. Such input can be provided to the fluorimeter by any computer terminal (Ralph Brooker, *personal communication*) provided the electronic language (baud rate, stop bits, *etc.*) and handshake protocols are compatible. A Macintosh computer contains the essential electronic message sending and receiving capabilities, and a standard Electronics Industry Association RS-232C – 8-pin mini din Macintosh modem cable[a] is properly wired for connecting the fluorimeter to the Macintosh.

Software

Terminal emulator programs drive the Macintosh to act as a terminal with some important enhancements. Three of these enhancements of the terminal emulator program Versaterm PRO make this system useful for data acquisition: First, the program contains a macro programming language which allows fluorimeter instructions to be recorded and played back in fluorimeter language when necessary. Second, the "**Save stream...**" protocol of Versaterm PRO allows data and keystrokes to be stored as text files. Third, the programming language includes a command which allows the standard Macintosh "Minifinder" to operate. The Minifinder allows files to be named and placed in specified directories. These features allow the collected data to be filed and sorted usefully according to the needs of the user. Pre-

Table 4. *Some useful commands in LS-5B Luminescence Spectrometer and Versaterm PRO Macro languages.*

Command	Meaning	Language
$RE 0	free response mode	Spectrometer
^M	send string	Versaterm
\	instruction set delimiter	Versaterm
^J	wait for stream	Versaterm
$GXNNN	set excitation wavelenth to NNN	Spectrometer
$GMNNN	set emission wavelenth to NNN	Spectrometer
$TI N	set time interval to N seconds	Spectrometer
^!Q	toggle **Save stream...** command	Versaterm
$TD N	collect N data points $TI apart	Spectrometer

[a] Available from Inmac®, 2465 Augustine Drive, Santa Clara, CA 95052, (800) 547-5444 (Cat. #74416)

packaged, commercial fluorimeter systems often lack flexibility or convenience by either naming data files by default, or by requiring naming at awkward moments.

The format for fluorimeter commands is "$AA N" where $ is $, A is an alphabetic character and N is a numeric character. Thus, two letters and a number define each command. All strings used in the collection of enzymological data begin with "$RE 0" which puts the fluorimeter in "free response" mode so that data is reported out of the RS-232C port as it is obtained. All commands sent by the Macintosh receive a reply from the fluorimeter, "0000" indicating that the command was received and executed. Similarly, "$GX360" sets the excitation monochrometer to 360 nm, and "$GM450" sets the emmision wavelength to 450 nm. The program in the combination fluorimeter and macro languages for this operation is

```
$RE 0^M\^J$GX360^M\^J$GM450^M\
```

which is named "4-MU Settings". The characters are listed in Table 4, and the screen resulting from the "4-MU Settings" macro is shown in Figure 21.

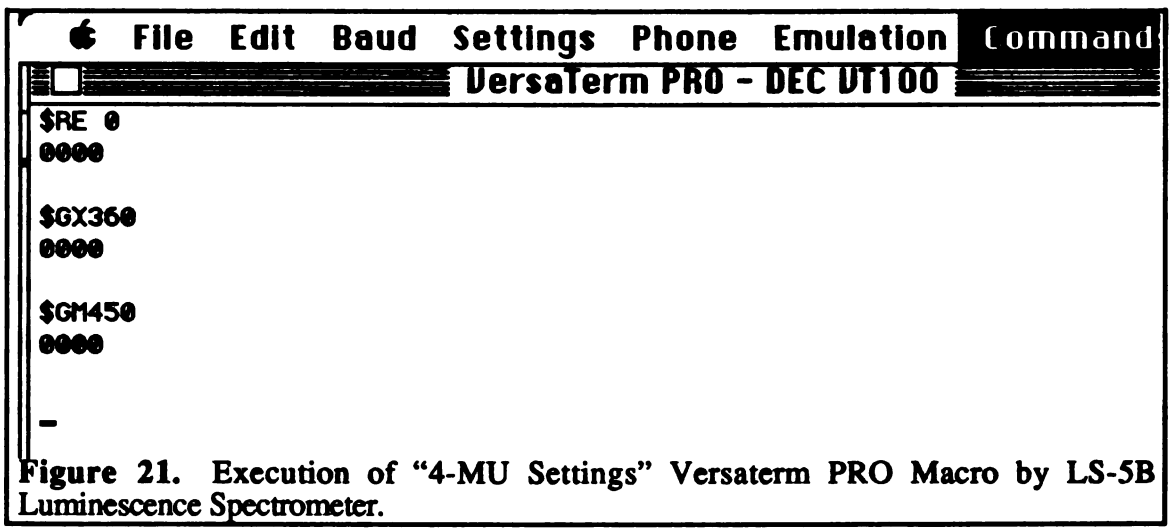


Figure 21. Execution of "4-MU Settings" Versaterm PRO Macro by LS-5B Luminescence Spectrometer.

Again "$RE 0" puts the fluorimeter in free response mode. "^M" (control-M) is the carriage return command which sends the previous string "$RE 0". The backslash

character is used to separate the commands and functions in the Versaterm PRO macro language so that "$RE 0^M" is one operation for the Macintosh. "^J" is the "wait for stream" command observed by the Versaterm PRO program; again, the prompt from the fluorimeter is "0000". The Macintosh then sends the command "set the excitation wavelength to 360 nm" as the string "$GX360" followed by "^M" to send the string. Again, the computer waits, "^J", for the "0000" response, then sends the command for the emmision wavelength and stops.

The fluorimeter may also send data streams of luminescence as a function of time (for observing kinetic rates) or as a function of excitation or emission wavelength (for recording spectra). The fluorimeter output buffer is able to accommodate ten data points of five characters each, thus outputting arbitrary luminescence units on a scale from -9999 to 19999 where 19999 is the level of saturation for the photodetector and zero is the null reading of the photodetector. The output buffer fills with ten numbers before the string is sent out the RS-232C port.

Two additional conditions exist during the acquistion of rate data. The data is to be stored using the "**Save stream...**" command protocol of Versaterm PRO, and the initial parameter line for "time drive" mode must be accommodated. The "**Save stream...**" subroutine is normally operated by the user as follows. The "**Save stream...**" command is selected prompting the computer to open a file on the hard disk. The Macintosh tool box provides the "minifinder" (Figure 22) which allows the user to name the data file and decide in what directory (in what "folder") the data file will be written. A good thorough name can be entered and the computer and fluorimeter will wait until the "*return*" key is pressed. When the "*return*" key is pressed, all characters typed in the window by either the fluorimeter or the keyboard operator are recorded and stored in a text file[a] until the "**Save stream...**" command is deselected. Thus, a new command "^!1Q" is introduced from the

[a] The creator of the file containing the saved stream can be designated as WORD using the "Extras" subroutine of Versaterm PRO. Thus, these data files will open from the Finder as WORD documents.

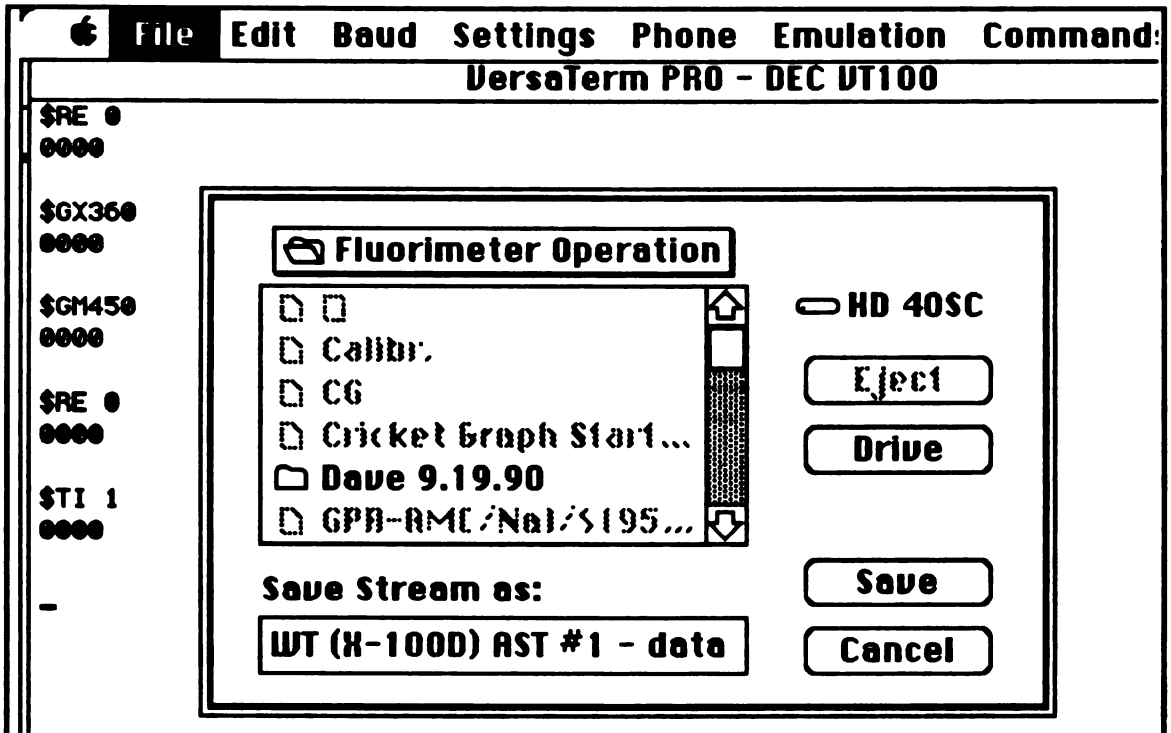


Figure 22. The Minifinder Dialog box of the the Apple Macintosh used with the **Save stream...** protocol of Versaterm PRO. Higher levels of the directory are accessible by clicking the box **Fluorimeter Operation**. Lower levels are accessed by double clicking on the folders (*e.g.* **Dave 9.19.90**) revealed using the scroll bar. Other files are indicated in grey print. The **Drive** button is used to move to the directory of other disks. Clicking the **Save** button or hitting the "*return*" key on the keyboard begins storage of the stream to the file **WT (X-100D) AST #1 - data** in this case. **Cancel** will also start storage of the stream, but when the **Save stream...** command is deselected, Minifinder will reappear.

Versaterm PRO macro language for "**Save stream...**", and a "^J" (wait for stream) for the parameter line must be added to the data collection part of the program in addition to one "^J" for every ten data points to be aquired. The resulting program for the aquisition of sixty data points is

$RE 0^M\^J$TI 1^M\^J^!1Q$TD 60^M\^J\^J\^J\^J\^J\^J\^J^!1Q

Again, "$RE 0" puts the fluorimeter in free response mode. On prompt, the string "$TI 1" sets the time interval between data points to 1 second and on prompt "^!1Q" activates the "**Save stream...**" subroutine of Versaterm PRO and stops the program for the ensuing data stream to be named in the minifinder. The "**Save stream...**" menu manipulations

described above are not actually seen when the "Time Drive 60" macro is run. The strings for putting the fluorimeter in free response mode and setting the time interval are sent and the minifinder dialog box appears (Figure 22). When reagent and substrate have been mixed and luminescence data is to be acquired, the "*return*" key is pressed. The fluorimeter

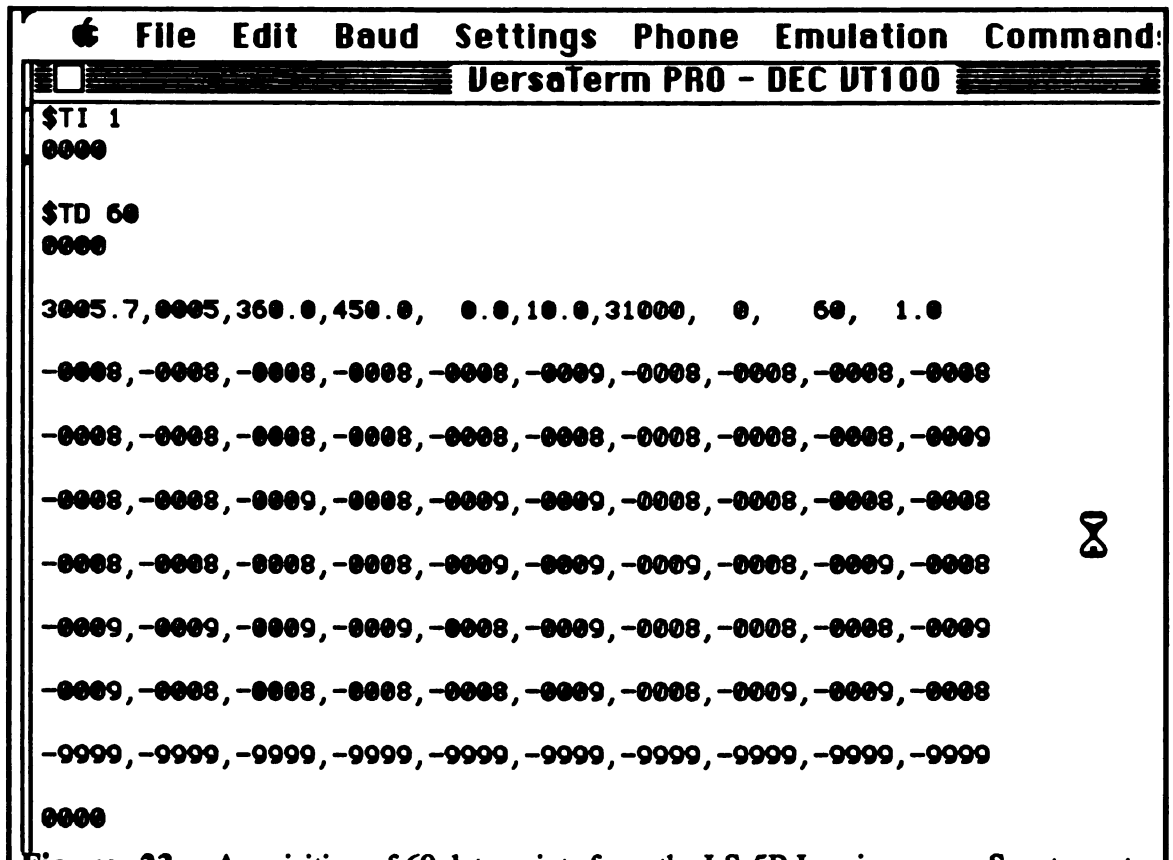


Figure 23. Acquisition of 60 data points from the LS-5B Luminescence Spectrometer into a Macintosh SE/30 using Versaterm PRO. In the next instant, the hourglass icon which is shown during all "^J" (wait for stream) commands would revert to the I-bar or arrow.

sends a parameter line indicating the status of various settings of the fluorimeter. The fluorimeter bundles and sends data in ten data point sets as described above. While Versaterm PRO is waiting under a "^J" command, the cursor appears as the hourglass icon (Figure 23).[a] When the last data point (as specified by the $TD command) is collected, the

[a] Running the Macintosh under "Multifinder," it is possible to leave Versaterm PRO using the hourglass icon to click on something on the screen not associated with the Versaterm window. Other programs such as WORD, EXCEL, CRICKET GRAPH or PRINT MONITOR can be used, and the data will be

Footnote continued on the next page

fluorimeter purges the output buffer with a string of ten "-9999" virtual data points. It is not necessary, therefore, to collect data points in multiples of ten; however, the fluorimeter will still output in multiples of ten data points making up the balance with "-9999" strings.

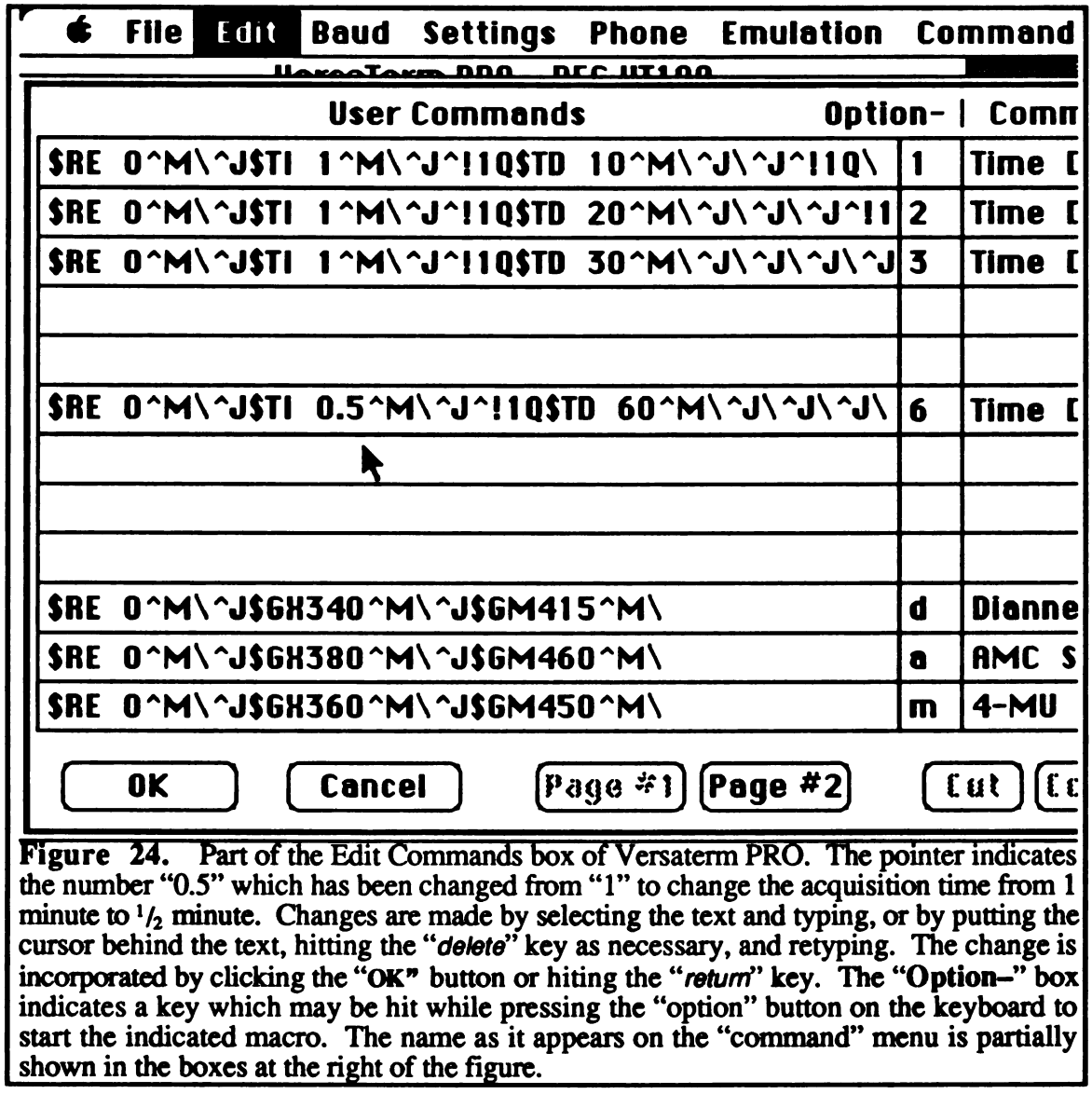


Figure 24. Part of the Edit Commands box of Versaterm PRO. The pointer indicates the number "0.5" which has been changed from "1" to change the acquisition time from 1 minute to ½ minute. Changes are made by selecting the text and typing, or by putting the cursor behind the text, hitting the "*delete*" key as necessary, and retyping. The change is incorporated by clicking the "OK" button or hiting the "*return*" key. The "**Option–**" box indicates a key which may be hit while pressing the "option" button on the keyboard to start the indicated macro. The name as it appears on the "command" menu is partially shown in the boxes at the right of the figure.

Sixty is a useful number of data points because the time interval setting can be easily changed by using the **Edit Commands** subroutine in Versaterm PRO (Figure 24). Since the time interval setting is expressed in seconds, the real time for a "$TD 60"

collected in the background as normal. Versaterm PRO functions are completely inactivated while the hourglass icon is visible, however.

acquisition is the "$TI" setting in minutes. Thus in "$TD 60" macro, "$TI 10" would be a ten minute acquisition, and "$TI 0.5" would be a half minute, 30 second, acquisition.

Finally, macros are recorded by selecting the "**Auto Macro Define**" function on the "**Commands**" menu and then typing in the $-sign two letter commands understood by the fluorimeter. The appropriate number of "^J" characters must be added in the Edit Commands dialog box since a "wait for stream" command is satisfied as soon as the stream is interrupted, as it is while the fluorimeter accumulates ten more data points.

Recording Data

Acquisition

Coupled with the manipulations of substrate and protein, the acquisition of rate data for a single assay includes the following steps. (1) The "option 6" command is used to start the "Time Drive 60" macro. (2) The substrate is mixed into the buffer in the cuvette in the fluorimeter sample holder. (3) The name of the file is typed into the minifinder. (4) The enzyme is mixed into the contents of the cuvette, and (5) the "*return*" key is hit to initiate data acquisition. Active site titration data is acquired similarly except that data for the background hydrolysis rate for the titrant is recorded for $^1/_{10}$ to $^1/_6$ of the total acquisition time before the enzyme is added. For standard curves, data are recorded for a few minutes while aliquots of a standard solution are intermittently mixed into the cuvette.

Reprocessing

Large amounts of enzymatic data can be collected and processed in a short time using this system. Twenty five determinations of k_{cat} and K_m using five substrate concentrations (as for a pH profile, for example) can easily be recorded in a single day. The data can be reprocessed (including the calculations) in half a day using an overall strategy for naming, storing, and manipulating data files in the Macintosh environment. Versaterm PRO is used to collect and store the data as WORD files. Microsoft WORD[a] is

[a] Microsoft Corporation, Redmond, WA.

used to extract and reformat the data for use in Microsoft EXCEL,[a] and also to print a hard copy of the data. EXCEL is used to calculate the Michaelis-Menten parameters. In addition, the macro languages of these programs or Macromaker[a] is used to assign keyboard commands to facilitate the process. A Macintosh with 4 kilobytes of RAM and the speed of a Macintosh SE/30 is minimal for the manipulations needed to reprocess data using this system. A trackball also facilitates the selection of text, speeds moving the cursor through the text windows, and leaves the cursor positioned appropriately for many of the operations described below.

Some strategy in saving and storing files should be employed. Separate folders were set up for each combination of enzyme and substrate to be tested. The data for rates of hydrolysis for five substrate concentrations were used for each determination of k_{cat} and K_m and these five acquisitions were placed in one folder. The data files were named with the numerical value of the substrate concentration first, so that they could be quickly ordered from lowest to highest concentration in WORD. The folders containing the raw data were stored in a single folder which was kept open in "**View by Name**" mode during data reduction.[b]

To begin reprocessing data under Multifinder, the folder containing the five data files for a given combination of enzyme and substrate was opened and positioned hanging from the bottom of the desktop. Note that the Desktop is visible below a normal WORD window and the folder was positioned to be visible there. The key combination "command-A"[c] selects all of the files in the active folder; "command-O" opens them. Since the creator for text files written by Versaterm PRO was set to WORD in the "**Extras...**" subroutine, this operation opens the raw data files in WORD Version 3.0.[d]

---

[a] Denman, D. *Macromaker* (Apple Computer Corporation, Cupertino, CA).

[b] If the "**View by date**" mode is used and the data are recorded in the order of decreasing substrate concentration, WORD will open the files in this order (Christopher Tsu, *personal communication*).

[c] The "command" key on a Macintosh keyboard has the "daisy" character imprinted on it.

[d] WORD version 3.0 is used because the macros written here do not work with version 4.0.

The files are ordered in WORD using the "**Window**" menu. By selecting in this menu the lowest substrate concentration file, then the next lowest, up to the highest, the files are progressively ordered so that closing the one on top reveals the data file for the next highest substrate concentration.

Versaterm PRO writes "escape" characters at the beginning of every line of text. These show up in WORD as back brackets (]) in the "Times" font. With data files ordered and the cursor in the default "home" position, a macro (keystroke "control-]") was written in Macromaker to do the following: move the cursor to a position behind each "]" at the front of a data line and delete it. At the end of the macro, *the cursor is left at the start of the first line of data*. The concentration of substrate for that data stream is typed by hand followed by a comma. The data in this window from the concentration to the last data point is selected by double-clicking on the concentration and dragging to the last data point. Double-clicking simplifies the selection of the last data point without selecting the carriage return character behind the last data point (made visible using the "**Show ¶**" setting [command-Y] in WORD). The selected text is copied to the clipboard (command-C) and the window is closed (command-W) without saving changes (n). The keystrokes for doing this are in parentheses and can be typed as fast as humanly possible without hanging the computer. When the WORD window with the next highest substrate concentration appears, the macro for stripping out the escape characters (control-]) is run again and the clipboard pasted in (command-V). Again, the key commands used are in parentheses and can be typed in quick succession to good effect. At this point, the "*return*" key is pressed to separate the first stream of data from the second stream, and the substrate concentration for the second stream is typed followed by a comma. The first and second data streams are selected by double clicking from the name of the first data stream to the last data point of the second data stream taking care to avoid the last carriage return as before. The data is

copied to the clipboard and the window closed without changes as before.[a] The escape characters are removed from the data stream for the third highest substrate concentration, and the data for the first two data streams is pasted in as before (hold the "control" key down with the little finger of your left hand and type a "]" with your right; the thumb of your left hand falls naturally over the "command" key. Hold the command key down with your left thumb and type a "v" with the index finger of your left hand). Type a carriage return and the name of the third substrate concentration. This time the window has scrolled the substrate concentration for the first data stream off the screen. Select the first three data streams by double-clicking on the last data point of the third stream and dragging to the name of the first data stream. Copy, close, no, control-], paste, *return, etc.* until all five data streams are copied on the clipboard and all data file windows are closed.

Paste (command-V) the five labeled data streams into a new text file. To use the data in Microsoft EXCEL, it must be tab-delimited. In addition, there are *two* carriage return characters at the end of each data line within a given stream, but only *one* carriage return between data streams. The double carriage returns are selectively removed using the find-and-replace text feature of WORD called "**Change...**" Manually, this would be accomplished by typing "^p^p" into the "**Find What:**" box and "^p" into the "**Change To:**" box but a macro was recorded for this purpose. Unfortunately, the shortcomings of the software and/or the author become evident here. The macro will not close the "**Change...**" box. Using a trackball, the close window hit box is clicked and the pointer is not moved because the next step involves a second "**Change...**" subroutine that will not close the box either. The commas are changed to tabs ("^t" in WORD language) using a "control-tab" initiated macro and the hit box can be triggered by hitting the trackball click

---

[a] Hold down the command key with your left thumb and type "cw" with convenient fingers on your left hand; then type an "n" naturally with your right hand removing your left thumb from the keyboard

button provided the pointer was not moved from its last position.[a] The data are now ready to be pasted into an appropriate EXCEL spreadsheet. The tabulated file, however, should be saved and a heading added describing the observed incubations. (There is a macro.....) Once this file is saved, it is convenient to use a "Save as..." command to name the tabulated file for the next set of data. The old data is then deleted from this file, and the finder accessed by clicking on the thin strip of the original folder which is visible under the WORD window. The folder is closed (command-W), the next folder opened by double-clicking, and positioned at the bottom of the screen as before. The process is repeated for each folder of data streams.

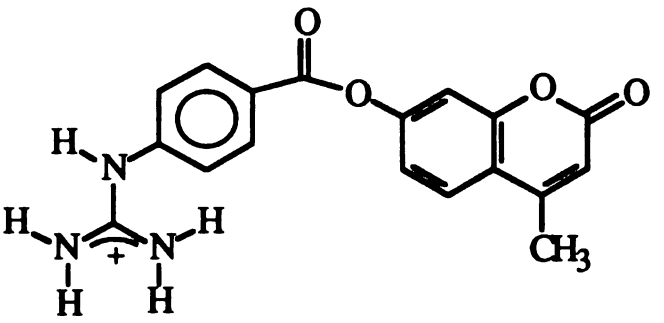
When dealing with a large number of Michaelis-Menten determinations, it is more efficient to accumulate and store the tabulated data files before using EXCEL. The minifinder of WORD was kept set to a folder for tabulated data by opening the raw data files using the Finder as described above. Once accumulated, a template spreadsheet was opened in EXCEL such that the folder containing the tabulated data files was visible on the desktop. The WORD files were opened by double-clicking in the finder. The tab-delimitted data was copied to the clipboard,[b] the window closed (command-W), and the EXCEL spreadsheet clicked to activate the program. The data were pasted at the appropriate place in the spreadsheet which could be saved, then saved as the next spreadsheet. The next tabulated data file was obtained by double clicking on the file name in the open folder in the finder, and the process repeated, *ad nauseum*. Happily, the data have been much enjoyed (by some people).

---

[a] To be sure, a supermacro could be written using the latest versions of the applications with better macro software to do all of this provided some means of encoding the identity of the data streams is worked out. This should be done by a post-doc, however, to avoid this type of documentation in another thesis.

[b] Put the cursor in the left margin of the WORD window where it becomes an arrow pointer. Hold down the "command" key, and click.

Active Site Titrations



MUGB

The concentration of trypsin active sites was determined using 4-methylumbelliferyl *p*-guanidinobenzoate (MUGB).[109] Wild type, S190G, S214A, and S214Q rat anionic trypsins and bovine cationic trypsin reacted with MUGB in the dead time of mixing so that the fluorescence intensity of 4-methylumbelliferone was recorded for 3 minutes. In order to re-establish linear hydrolysis rates, data were recorded for 10 minutes for the S214E mutant, for 30 minutes for the S214K mutant, and for 90 minutes for the S214R mutant.

Initial Velocity Studies

Choice of Substrates. Amide hydrolysis rates were determined using 7-amino-4-methylcoumarin (AMC) substrates. These compounds (1) are stable in water, (2) are highly specific, (3) have excitation and emission wavelengths which are different for substrates and products, (4) are sensitive substrates (assays require short incubation times), (5) can be used spectrophotometrically if necessary ($\Delta\varepsilon_{360}$ = 7,800 M^{-1} cm^{-1}),[11] (6) can be readily developed based on any peptide, and (7) are commercially available. Ester hydrolysis was followed using thiobenzyl ester substrates

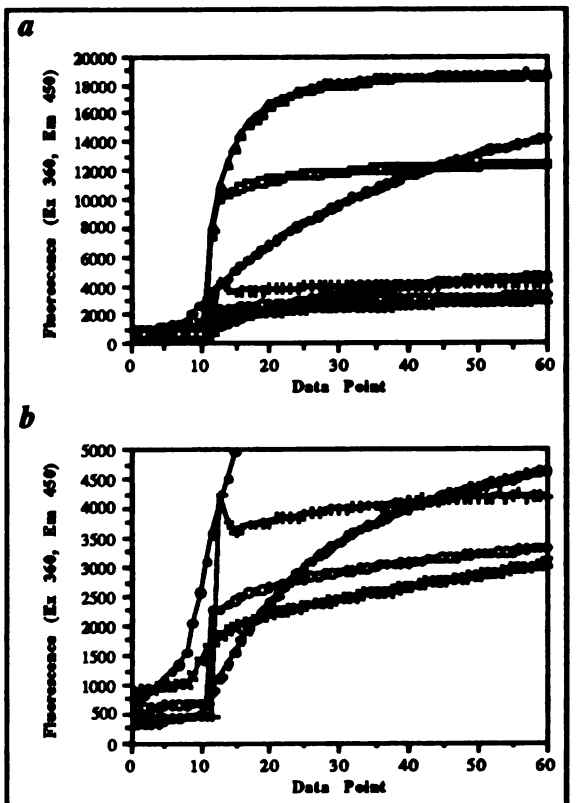


Figure 25. Active Site Titrations. *a* Observed release of 4-methylumbelliferone for the reaction of MUGB with trypsin(+) and S190G(○), S214(A[□], Q[×], E[▲], K[◆], and R[●]) trypsins. *b* Expanded scale.

(Figure 3). These ester substrates allow the chracterization of ester hydrolysis rates at pH > 7 where oxygen esters react too rapidly with hydroxide.

Enzyme Assays. Initial velocities for the hydrolysis of AMC substrates[84] and thiobenzyl ester substrates[49] were determined as described. The fluoresence of AMC was followed in the LS-5B Luminescence Spectrometer. The hydrolysis of thiobenzyl ester substrates was monitored by following the generation of 4-thiopyridone (Figure 3) in a UVICON 860 Spectrophotometer.[a]

The buffer system for the pH-rate profiles has been described previously.[49] A set of 0.5M stock MES (pH 4.5 – 6.9), MOPS (pH 6.1 – 8.1), TAPS (pH 7.3 – 9.3), and glycine (pH 8.5 – 10.5) buffer solutions of various pH values were prepared, sterile filtered through 0.22 μm filters, and kept sterile. To make buffer solutions for the pH-rate profiles, the stock buffer was diluted (to 50 mM) along with stock solutions of $CaCl_2$ (to 20 mM) and NaCl (to 100 mM) in 50 ml orange-capped tubes. The pH of each solution was measured and recorded. It was experimentally verified that the enzyme incubations would not exceed the capacity of the buffers during the time of the observed rates.

Enzymes were stored in 100x stocks in 1 mM HCl and were diluted in thermostatted cuvettes just prior to reaction. Enzymatic hydrolysis rates were determined for 2.5 nM or 250 nM enzyme in 100 mM NaCl, 20 mM $CaCl_2$, 50 mM Tris, pH 8.0, with 1% DMF for amides or 2% DMF and 250 μM 4,4'-dithiodipyridine for esters. Solvolysis rates were subtracted for Z-R-S-Bzl substrates[49] and were otherwise insignificant. An effort was made to optimize enzyme and substrate concentrations[5] so that initial velocities could be determined at concentrations above and below K_m. Unfortunately, the solubility of many substrates limited the maximum available concentrations for kinetic determinations. Kinetic parameters k_{cat} and K_m were calculated from Eadie-Hofstee plots[68] which were found to give k_{cat} and K_m values which did not

[a] Kontron Instruments

deviate significantly from non-linear regression of the same data to the Michaelis-Menten equation or from statistically weighted equations for k_{cat} and K_m.[45] Values of the apparent second order rate constant k_{cat}/K_m were calculated using the statistically weighted equations of Cornish-Bowden and Wharton.[45] Theoretical curves for the pH profiles were determined by nonlinear regression of the data to the equations indicated using the program Eureka: The Solver. The advantage of this program is that it will perform nonlinear, least squares fit of the data to any equation and additionally has a convergence test so that it will run until a good fit is obtained. This feature relieves the operator of guessing the number of iterations necessary to achieve convergence.

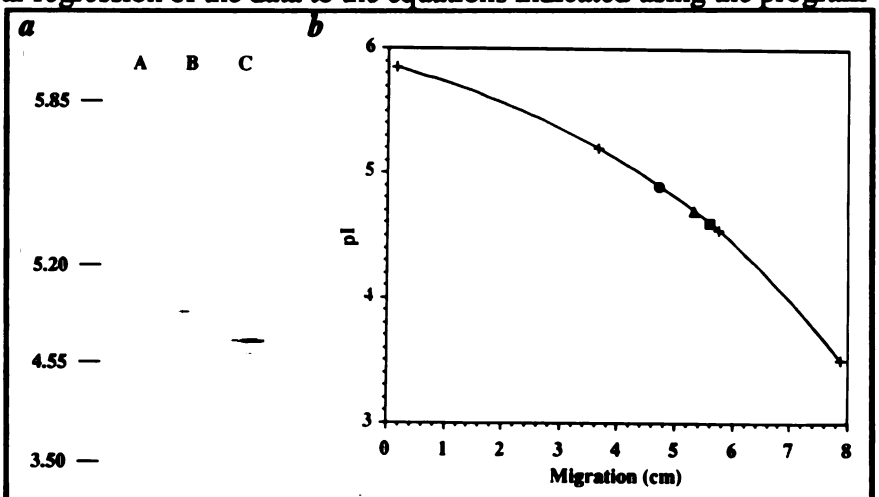


Figure 26. Analytical Isoelectric Focusing of Rat Anionic Trypsin Mutants. Approximately 7.5 μg of each protein were focused in an electric field under nitrogen at 25 watts constant power in a commercially prepared 14 cm polyacrylamide gel containing ampholines for the pH range 4 – 6.5. (*a*) Coomassie-stained isoelectric focusing gel of (A) standards, (B) S214K trypsin, and (C) trypsin. Standards are bovine carbonic anhydrase (pI 5.85), β-lactoglobulin (pI 5.20), soybean trypsin inhibitor (pI 4.55), amyloglucosidase (pI 3.50). (*b*) Interpolation of the isoelectric point of (●) S214K (4.90) and (▲, ■) wild type (4.70 and 4.60) rat anionic trypsins. (*c*) Coomassie-stained isoelectric focusing gel of (A) standards (as above), (B) S214R, (C) S214K, (D) wild type, (E) S214A, (F) S214Q, and (G) S214E trypsins

PERTURBATION OF THE ELECTROSTATIC FIELD IN THE ACTIVE SITE OF TRYPSIN – RESULTS

ANALYTICAL ISOELECTRIC FOCUSING

Mutant trypsins with Lys- or Arg-214 side chains focus at an isoelectric point 0.2 pI units more basic than rat anionic trypsin which has a Ser-214 side chain (Figure 26). By comparison, mutants with Ala-, Gln-, or Glu-214 side chains focus at a slightly more basic

pI relative to trypsin (Figure 26*c*). Trypsins with Arg-, Lys-, Ala-, and Gln-214 side chains focus as single bands within the limits of the experiment (Figure 26*c*). The trypsins with Glu-214 and Ser-214 (wild type) side chains focus as doublets indicating some heterogeneity in the sample.

Structural Characteristics of S214(K,Q,E) Trypsins

The crystal structure of rat anionic trypsin has not been reported. The crystal structure of D102N rat anionic trypsin,[203] however, is only subtly different from bovine trypsin.[29,213] In trypsin, the Ser-214 side chain is buried and normally forms hydrogen bonds with the catalytic Asp-102 Oδ2 and buried water molecule Wat-408 (Figure 27). Wat-408, in turn, forms hydrogen bonds with buried Wat-241 (a.k.a. Wat-703),[29] with the backbone carbonyl of Leu-99, and with the Thr-229 hydroxyl group (not shown). In the Michaelis complex, the carbonyl of Ser-214 forms a hydrogen bond with the backbone amide of the P_1 residue of the substrate.

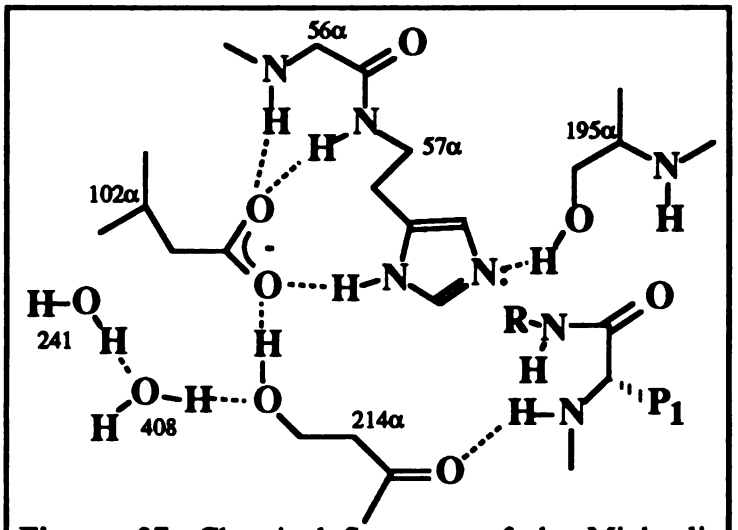


Figure 27. Chemical Structure of the Michaelis Complex of Trypsin.

S214K Trypsin

X-ray diffraction data from a single crystal of S214K trypsin were recorded on a Xentronics area detector. Using phases and amplitudes computed from the D102N:S195C rat anionic trypsin structure (2.0Å, $R = 0.2$),[149] an $|F_O-F_C|$ map was calculated for a model in which water molecules were excluded and His-57, Asp-102, Lys-214, and Trp-215 were replaced with alanine side chains (Figure 28*a*). The $|2F_O-F_C|$ map of S214K

trypsin (Figure 28*b*) has been calculated at 2.2Å resolution ($R = 0.164$) after 80 cycles[148] of Hendrickson-Konnert refinement using the program HKREF.[100]

The Lys-214 side chain coils to fit into the space formerly occupied by the Ser-214 side chain and Wat-408. The Lys-214 Nζ replaces Wat-408 and forms the corresponding hydrogen bonds with buried Wat-241, with the Leu-99 O, and with the Thr-229 Oγ (Figure 28*b*). The Lys-214 side chain is sequestered from bulk solvent. As shown in Figure 29, the solvent accessible surface of S214K trypsin is cut away to reveal the Lys-214 side chain which is highlighted with a mauve van der Waals' surface. Also visible are the Asp-102

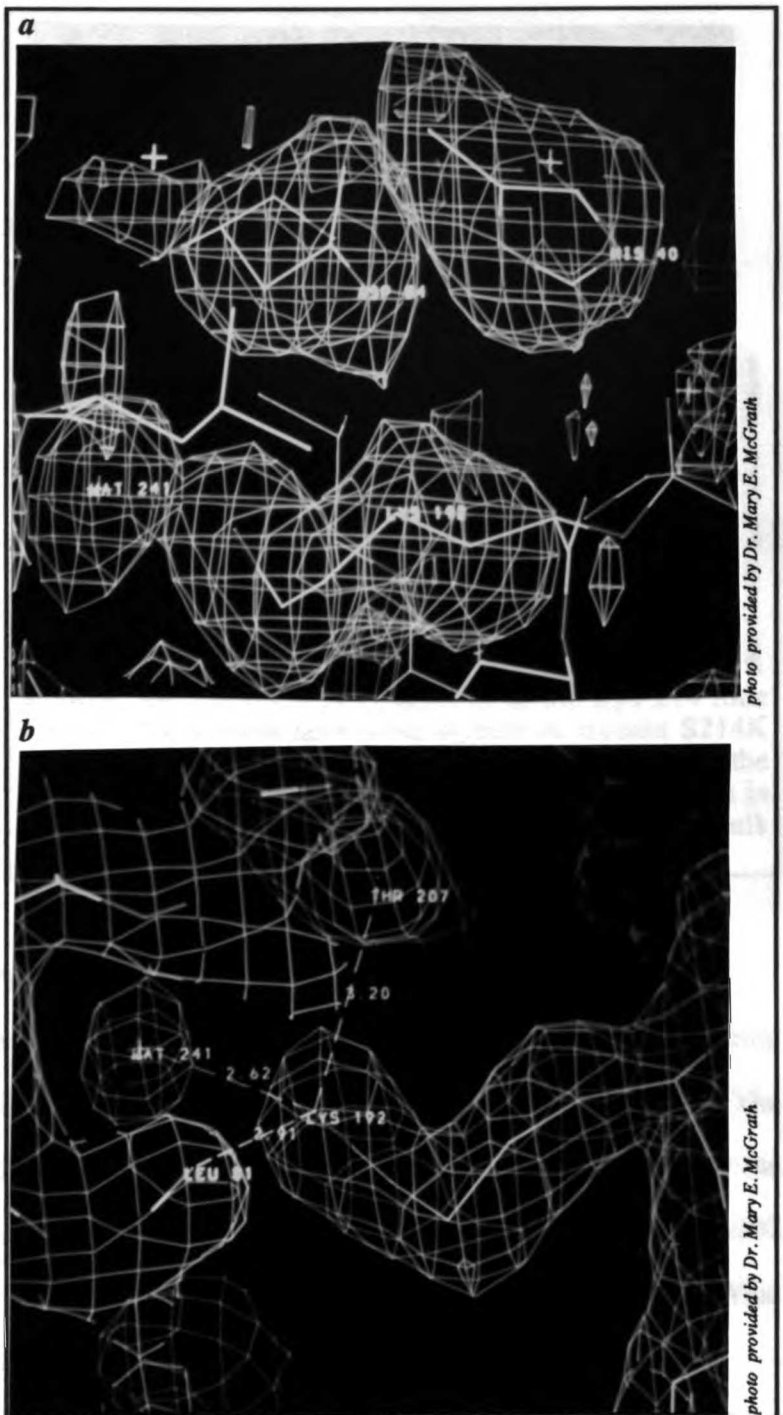


Figure 28. Crystal structure of S214K trypsin. *a.* $|F_o-F_c|$ map superimposed on the model structure. *b.* $|2F_o-F_c|$ 2.2Å map superimposed on the model structure. The crystallographic numbering maps to the standard chymotrypsinogen numbrering as follows: 40 = 57, 81 = 99, 84 = 102, 192 = 214, 207 = 223.

side chain (between the "1" and "0" digits), the His-57 imidazole ring (below the "57"), and benzamidine which stands in the P_1 site. Not shown is the Leu-99 side chain which lays over the Asp-102 side chain but which has been cut away in this slice through the protein. The position of the Leu-99 side chain appears in white below and parrallel to the Asp-102 side chain in Figure 30 and is clearly visible in the $|F_o - F_c|$ map of S214E trypsin (Figure 32).

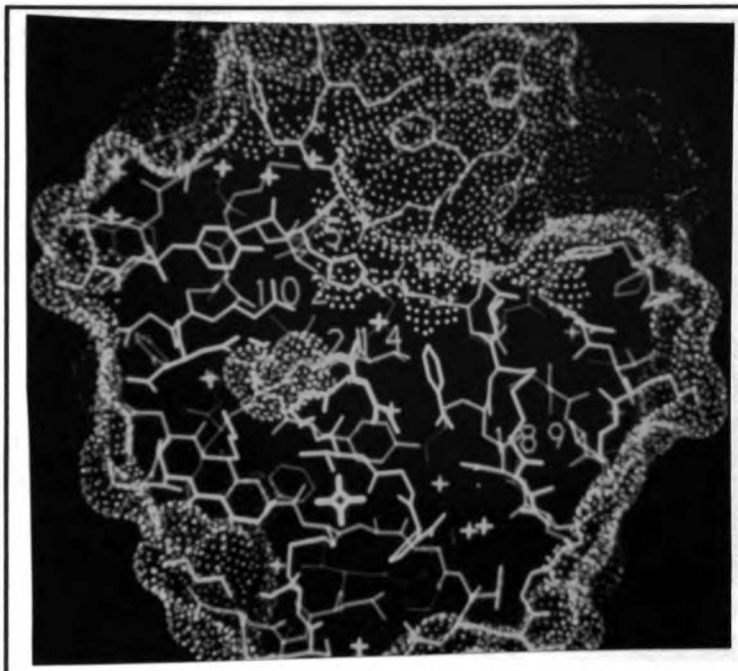


Figure 29. Solvent Accessibility of the Lys-214 Side Chain. The solvent accessible surface of trypsin S214K represented by turquoise dots is sliced away to reveal the Van der Waal values surface of the Lys-214 side chain in mauve. The Lys-214 side chain is sequestered from bulk solvent.

The solvent accessible surface can be seen stretching over the His-57 imidazole ring and dipping into the P_1 binding pocket to touch the top of the benzamidine ring. The Asp-102 side chain is buried[27] so that the solvent accessible surface does not touch the Asp-102 carboxylate. The Lys-214 side chain is roughly in the same plane as the Asp-102 side chain. Since the solvent accessible surface can be cut away but not the van der Waal values surface, the Lys-214 side chain is clearly in the interior of the protein.

The introduction of a lysine side chain into the position formerly occupied by a serine side chain and a water molecule causes two significant side chains to move: (1) Trp-215 and (2) Ser-195. The Trp-215 side chain is rotated 134° around the Cα–Cβ bond

so that the indole ring rests against the main chain trypsin backbone from Ser-214 to Gly-216 (Figure 30). This conformation is called "*out*." The Ser-195 Oγ is shifted to a position 3.7Å from the His-57 Nε from a normal position[28,108,200] ~3.0Å from the His-57 Nε. Val-227 appears to be rotated in S214K trypsin, but it is not known if this is a difference between the proteins, or a difference due to the altered positon of the Trp-215 side chain. The electron density of the solvent accessible Gln-192 side chain is poorly defined and so appears to move in the mutant.

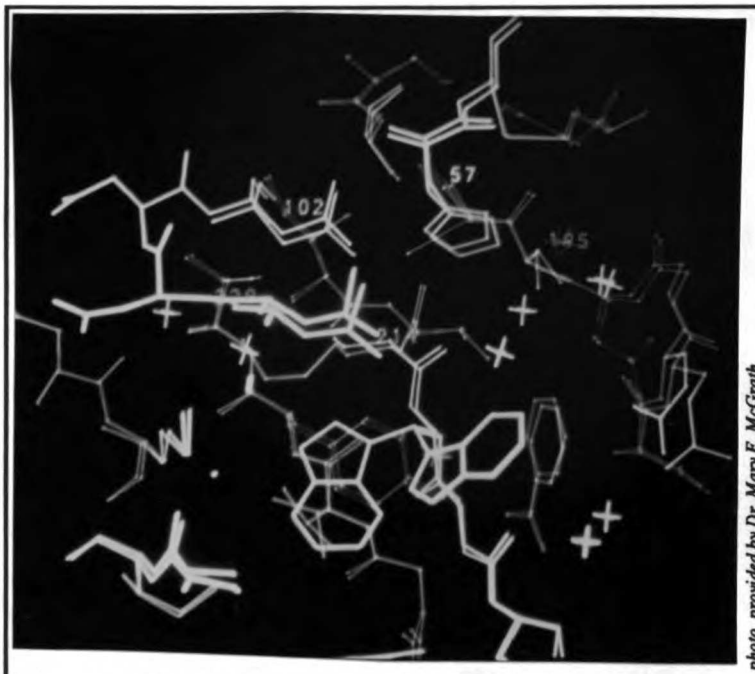


Figure 30. Active site of S214K trypsin compared to cow trypsin. The S214K Trp-215 "out" position can be seen blocking the backbone above benzamidine which sits in the P_1 site. The other Trp side chain is in cow trypsin.

S214Q Trypsin

The crystal structure of S214Q trypsin has been determined to 2.4Å resolution ($R = 0.196$) (Dr. Mary E. McGrath, *personal communication*) after 40 cycles of refinement using HKREF.[228] Like the Lys-214 Nζ, the Gln-214 side chain also replaces buried Wat-408. By taking account of hydrogen bond donor and acceptor interactions, the orientation of the side chain amide carbonyl has been assigned as shown (Figure 31). The Gln-214 Nε primarily forms a hydrogen bond with Thr-229 Oγ (2.5Å) and a somewhat longer contact with Wat-241 (3.3Å). Given this assignment of the amide bond, the

Gln-214 Oε is too far to form hydrogen bonds with Wat-241 (3.5Å) or Thr-229 Oγ (3.8Å)

the only neighbors which potentially could form hydrogen bonds.

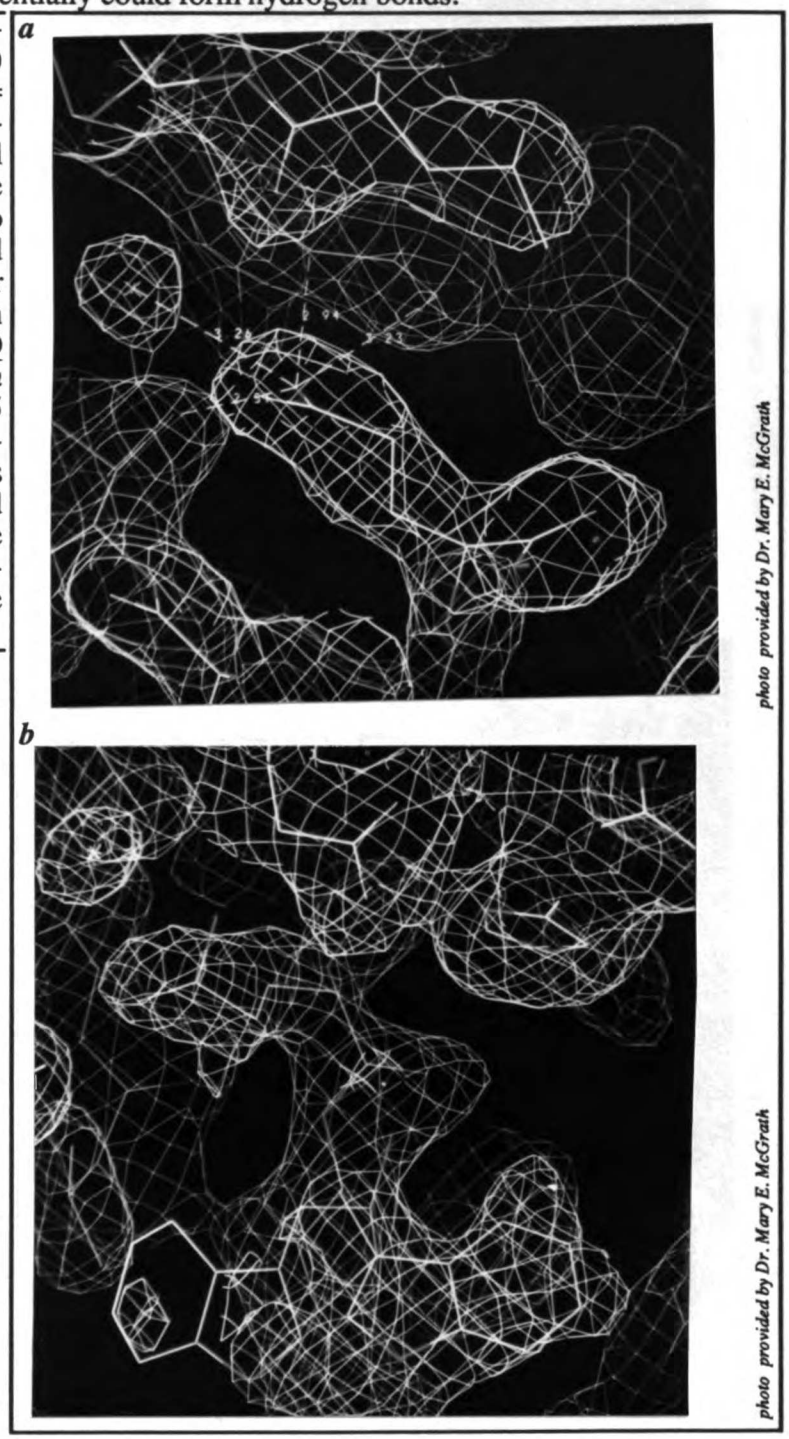


Figure 31. Crystal Structure of S214Q Trypsin. (*a*) 2.4Å $|2F_o-F_c|$ map (R = 0.168) of S214Q trypsin superimposed on the model structure. Gln-214 side chain with close contacts to Asp-102 Oδ2, Wat-241, and Leu-99 O. Not shown for clarity, 2.5Å contact with Thr-229 Oγ. (*b*) S214Q map with both Glu-214 *in* and *out* positions of Trp-215 found in the Glu-214 structure. The *out* position is identical in both S214E and S214Q trypsins, but the density for the *in* conformation is found only in the S214E trypsin. (Figure 32)

The glutamine side chain causes the same two significant changes in the structure as the Lys-214 side chain: Trp-215 is displaced and the catalytic Ser-195 Oγ is shifted. The Trp-215 side chain is rotated around the Cα–Cβ bond to a more severe "out" conformation than is observed in S214K trypsin (Figure 30). The **Ser-195 Oγ is shifted to a position 4.0Å from the His-57 Nε from a normal position ~3.0Å from the His-57 Nε.**

S214E Trypsin

The crystal structure of S214E trypsin has been determined to 2.8Å resolution (R = 0.164) (Dr. Mary McGrath, *personal communication*) after 51 cycles of HKREF.[229] The Glu-214 side chain apparently adopts two conformations denoted "*up*" and "*down*" (Figure 32). In the up conformation, the Glu-214 carboxylate interacts directly with the carboxylate of Asp-102. In the *down* conformation, the Glu-214 carboxylate replaces Wat-408 and interacts primarily with buried Wat-241 (a.k.a. Wat-703).[29] In the *up* conformation, the Oε1 of Glu-214 forms a strong hydrogen bond (2.4Å) with the Oδ2 of Asp-102 while the Oε2 of Glu-214 is 2.9Å from

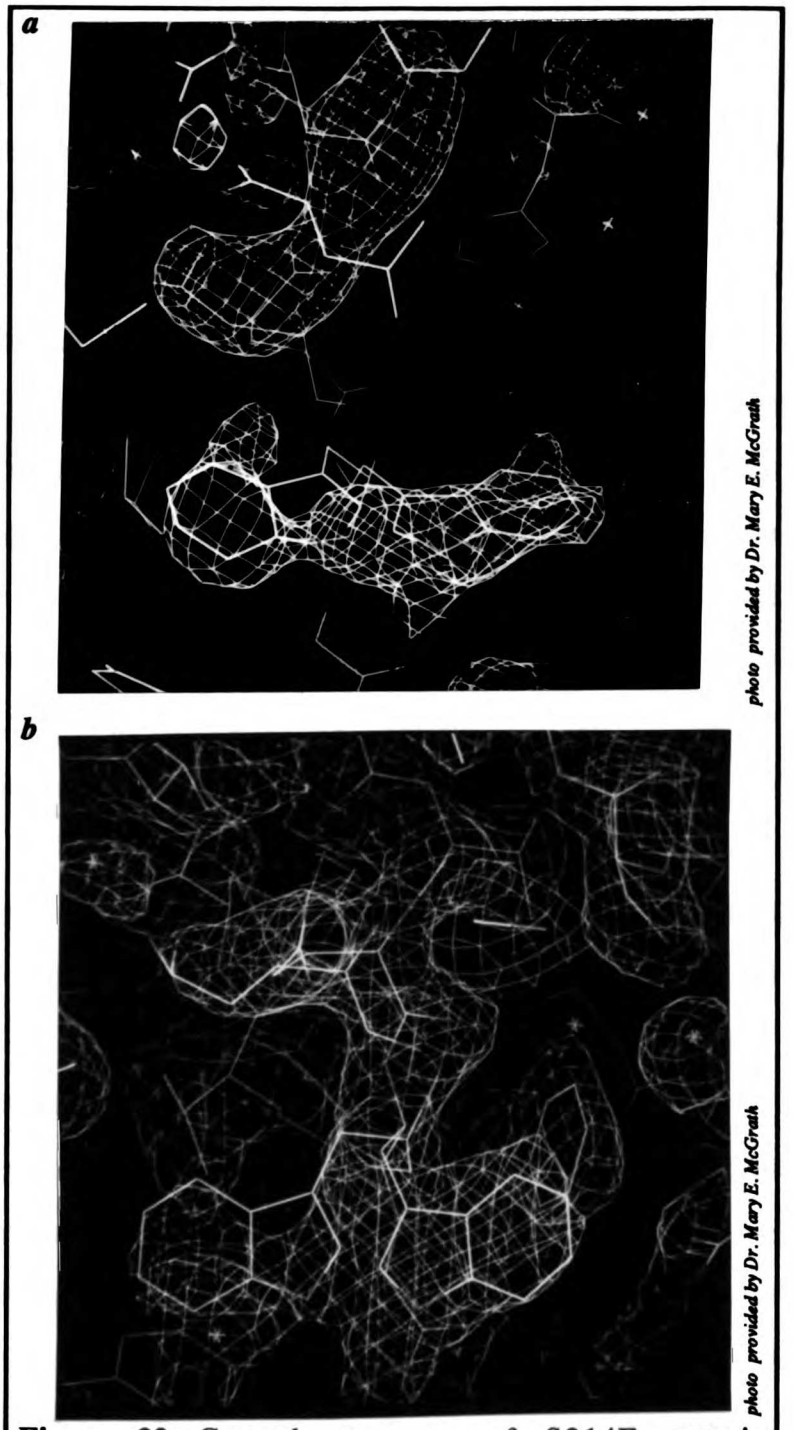


Figure 32. Crystal structure of S214E trypsin (*a*) Omit map superimposed on the model structure which has the final pairs of Glu-214 and Trp-215 side chain positions. (*b*) $|2F_o–F_c|$ map superimposed on the model structure. Note that the Trp-215 *in* position is at a 90° angle to the position of the Trp-215 rings in trypsin.

the Oδ2 of Asp-102 and is stabilized by interaction with the Thr-229 hydroxyl group (3.0Å). In the *down* conformation, the Glu-214 side chain primarily interacts with buried water molecule Wat-241. The Glu-214 Oε2 is 2.7Å from Wat-241 while Oε1 is 3.3Å from Wat-241 and 3.2Å from the Thr-229 hydroxyl group. The relative intensities of the electron density in the $|2F_o-F_c|$ map (Figure 32*b*) are approximately 60% and 40% for the *up* and *down* conformations, respectively.

The introduction of a glutamic acid side chain into the position formerly occupied by a serine side chain and a water molecule causes two other significant changes in the structure: Trp-215 is displaced and the catalytic Ser-195 Oγ is shifted. The electron density for the Trp-215 side chain is split between two conformations denoted "*out*" and "*in*." The Trp-215 *in* conformation is approximately in the position observed in trypsin; however, the indole ring is apparently flipped 180° and rotated -45° with appropriate adjustment of the side chain between the indole ring and the α-carbon. The *out* conformation for Trp-215 in S214E trypsin is identical to that found in S214Q trypsin. Integration of the electron density in the $|2F_o-F_c|$ maps indicates that the Trp-215 side chain is displaced to the out position in one third of the molecules. The Ser-195 Oγ is shifted to a position 4.0Å from the His-57 Nε from a normal position ~3.0Å from the His-57 Nε.

KINETIC CHARACTERISTICS AT PH 8.0

Steady-State Kinetics. The Michaelis-Menten parameters for the hydrolysis of various substrates by each enzyme are summarized in Table 5. In general, duplicate determinations agree well, although for Suc-AFK-AMC the scatter in the determination of K_m is appreciably larger than for the other substrates. The substrates Z-GPR-AMC and Z-GPR-S-Bzl were used to test the effect of the chemistry of the substrates on the observed rate of reaction. When these substrates were hydrolyzed at the same rates by rat anionic trypsin, single residue Z-R-AMC and Z-R-S-Bzl substrates were used to probe this aspect of the chemistry of the active sites of trypsin and mutant trypsins. Suc-AFK-AMC and

MeOSuc-AFK-AMC were used to test the effect of charges on the interaction of the substrate with the mutant enzymes. Suc-AFK-AMC is electrostatically neutral while MeOSuc-AFK-AMC is positively charged at neutral pH. Finally, Suc-AAPF-S-Bzl was used to control for the effect of having no positive charge on the substrate in the P_1 binding pocket. In addition, ester substrates easily acylate the chemically activated Ser-195 hydroxyl group. Therefore, the rate of reaction for esters and serine proteases is generally limited by the rate of the deacylation reaction, *i.e.*, by the activation of a water molecule to attack the acyl-enzyme intermediate. Amides, however, are chemically less active than esters, and the rates of reactions of amides with serine proteases are generally limited by the rate of acylation of the enzyme; the acyl enzyme, being an ester, is chemically more labile than the substrate amide and therefore reacts faster.

Using Z-R-S-Bzl and Z-R-AMC substrates, the k_{cat} for the thioester substrate was found to be greater than the k_{cat} for the amide substrate for S214Q and wild type rat anionic trypsin. Similarly, using Z-GPR-S-Bzl and Z-GPR-AMC substrates, the k_{cat} for the thioester substrate was found to be greater than the k_{cat} for the amide substrate for the S214D, S214Q, S214E, S214K, and S190G mutants. Amide substrates which do not have arginine or lysine as the P_1 residue were not cleaved by S214K trypsin at a measurable rate (unpublished results). No substrates were found to have a K_m low enough to allow determination of the Michaelis-Menten parameters k_{cat} or K_m for the S214R mutant.

The values of the kinetic parameters k_{cat}, K_m, and k_{cat}/K_m normalized to the values for rat anionic trypsin are listed in Table 6. Bovine cationic trypsin hydrolyzes the MeOSuc-AFK-AMC and Suc-AFK-AMC substrates with a catalytic efficiency (k_{cat}/K_m) of 35-50% that of rat anionic trypsin. The S214A and S190G mutations have little effect on k_{cat}/K_m relative to wild type. The S214D and S190G mutations increase the K_m of the MeOSuc- and Suc-AFK-AMC substrates relative to all other substrates. Therefore, for

Table 5. *Kinetic constants for the hydrolysis of various substrates by S214X and S190G trypsins at 25°C in 50 mM Tris HCl pH 8.0, 100 mM NaCl, 20 mM $CaCl_2$, containing 1% DMF for amides and 2% DMF and 250 μM 4,4'-dithiodipyridine for esters. The values are the means of two determinations ± the standard deviations for each pair of determinations. Each determination included the observed hydrolysis rates for reactions at five concentrations of substrate. It was not always possible to obtain data for substrate concentrations greater than K_m due to the limit of the solubility of the substrate.*

Enzyme	Substrate	kcat (min^{-1})	Km (μM)	kcat/Km ($min^{-1}\mu M^{-1}$)
Cationic Cow WT	Suc-AFK-AMC	800 ± 30	600 ± 20	1.3 ± 0.01
	MeOSuc-AFK-AMC	580 ± 20	220 ± 10	2.7 ± 0.01
Anionic Rat WT (Ser-214)	Suc-AFK-AMC	490 ± 20	190 ± 10	2.6 ± 0.06
	MeOSuc-AFK-AMC	440 ± 0.04	58 ± 0.8	7.7 ± 0.1
	Suc-AAPF-S-Bzl	3,200 ± 80	240 ± 10	14 ± 0.5
	Z-GPR-S-Bzl	5,000 ± 400	12 ± 0.6	430 ± 20
	Z-GPR-AMC	5,300 ± 100	20 ± 0.6	270 ± 10
	Z-R-S-Bzl	5,300 ± 100	3.6 ± 0.2	1,500 ± 60
	Z-R-AMC	39 ± 0.6	62 ± 10	0.63 ± 0.1
Rat S214A	Suc-AFK-AMC	900 ± 7	260 ± 2	3.4 ± 0.01
	MeOSuc-AFK-AMC	780 ± 10	100 ± 1	7.4 ± 0.04
	Suc-AAPF-S-Bzl	2,300 ± 50	150 ± 10	15 ± 1
	Z-GPR-S-Bzl	3,100 ± 60	88 ± 8	35 ± 4
	Z-GPR-AMC	4,600 ± 70	16 ± 0.9	290 ± 10
	Z-R-S-Bzl	6,100 ± 40	4.3 ± 0.4	1,400 ± 200
	Z-R-AMC	110 ± 1	68 ± 4	1.6 ± 0.08
Rat S214D	Suc-AFK-AMC	87 ± 5	3,500 ± 300	0.024 ± 0.0001
	MeOSuc-AFK-AMC	25 ± 2	550 ± 70	0.047 ± 0.001
	Suc-AAPF-S-Bzl	1,300 ± 40	270 ± 20	4.9 ± 0.4
	Z-GPR-S-Bzl	2,300 ± 40	12 ± 0.9	190 ± 10
	Z-GPR-AMC	2,000 ± 10	76 ± 1	26 ± 0.3
	Z-R-S-Bzl	4,500 ± 400	11 ± 3	440 ± 90
	Z-R-AMC	7.3 ± 3	130 ± 20	0.054 ± 0.01
Rat S214Q	Suc-AFK-AMC	-	-	0.0040
	MeOSuc-AFK-AMC	-	-	0.0058
	Suc-AAPF-S-Bzl	1,200 ± 300	4,600 ± 1,500	0.26 ± 0.01
	Z-GPR-S-Bzl	3,200 ± 20	120 ± 1	26 ± 0.4
	Z-GPR-AMC	470 ± 40	320 ± 30	1.5 ± 0.01
	Z-R-S-Bzl	2,400 ± 200	60 ± 5	40 ± 0.4
	Z-R-AMC	12 ± 0.1	2,700 ± 10	0.0044 ± 0.0004
Rat S214E	Suc-AFK-AMC	220 ± 70	6,700 ± 2,000	0.033 ± 0.0003
	MeOSuc-AFK-AMC	160 ± 8	1,900 ± 100	0.08 ± 0.01
	Suc-AAPF-S-Bzl	1,300 ± 200	2,400 ± 400	0.55 ± 0.01
	Z-GPR-S-Bzl	5,200 ± 50	94 ± 1	55 ± 0.5
	Z-GPR-AMC	2,000 ± 50	390 ± 20	5.1 ± 0.09
	Z-R-S-Bzl	4,700 ± 20	43 ± 1	110 ± 4
	Z-R-AMC	16 ± 1	730 ± 60	0.022 ± 0.0005
Rat S214K	Suc-AFK-AMC	4.2 ± 0.8	6,500 ± 1,000	0.00064 ± 0.00006
	MeOSuc-AFK-AMC	7.2 ± 0.2	4,400 ± 200	0.0016 ± 0.00001
	Suc-AAPF-S-Bzl	62 ± 5	3,400 ± 400	0.018 ± 0.0006
	Z-GPR-S-Bzl	7,900 ± 400	1,300 ± 40	6.3 ± 0.1
	Z-GPR-AMC	630 ± 30	4,200 ± 60	0.15 ± 0.005
	Z-R-S-Bzl	46 ± 1	370 ± 8	7.5 ± 0.06
	Z-R-AMC	-	-	0.0014 ± 0.0001
Rat S214R	Z-GPR-S-Bzl	-	-	0.70
	Z-GPR-AMC	-	-	0.0096
Rat S190G	Suc-AFK-AMC	140 ± 4	3,000 ± 50	0.046 ± 0.001
	MeOSuc-AFK-AMC	67 ± 20	700 ± 200	0.097 ± 0.008
	Suc-AAPF-S-Bzl	1,700 ± 30	190 ± 3	9.2 ± 0.01
	Z-GPR-S-Bzl	10,000 ± 1,000	34 ± 9	290 ± 40
	Z-GPR-AMC	6,700 ± 60	100 ± 1	64 ± 0.1
	Z-R-S-Bzl	10,000 ± 200	9.4 ± 0.6	1,100 ± 40
	Z-R-AMC	58 ± 0.1	340 ± 1	0.17 ± 0.001

Table 6. *Normalized kinetic constants for the hydrolysis of various substrates by S214X and S190G trypsins*

Enzyme	Substrate	k_{cat}	K_m	k_{cat}/K_m
Cationic Cow WT	Suc-AFK-AMC	160%	3.1x	52%
	MeOSuc-AFK-AMC	130%	3.4x	35%
Anionic Rat WT (Ser-214)	Suc-AFK-AMC	100%	1x	100%
	MeOSuc-AFK-AMC	100%	1x	100%
	Suc-AAPF-S-Bzl	100%	1x	100%
	Z-GPR-S-Bzl	100%	1x	100%
	Z-GPR-AMC	100%	1x	100%
	Z-R-S-Bzl	100%	1x	100%
	Z-R-AMC	100%	1x	100%
Rat S214A	Suc-AFK-AMC	180%	1.4x	130%
	MeOSuc-AFK-AMC	180%	1.8x	96%
	Suc-AAPF-S-Bzl	72%	0.63x	110%
	Z-GPR-S-Bzl	62%	7.6x	8.2%
	Z-GPR-AMC	92%	0.8x	110%
	Z-R-S-Bzl	120%	1.2x	95%
	Z-R-AMC	280%	1.1x	260%
Rat S214D	Suc-AFK-AMC	18%	18x	0.92%
	MeOSuc-AFK-AMC	5.7%	9.5x	0.61%
	Suc-AAPF-S-Bzl	41%	1.1x	35%
	Z-GPR-S-Bzl	43%	1.0x	44%
	Z-GPR-AMC	38%	3.8x	9.6%
	Z-R-S-Bzl	85%	3.1x	29%
	Z-R-AMC	19%	2.1x	8.6%
Rat S214Q	Suc-AFK-AMC	-	-	0.15%
	MeOSuc-AFK-AMC	-	-	0.075%
	Suc-AAPF-S-Bzl	38%	19x	1.9%
	Z-GPR-S-Bzl	64%	10x	6.0%
	Z-GPR-AMC	8.9%	16x	0.55%
	Z-R-S-Bzl	45%	17x	2.7%
	Z-R-AMC	31%	44x	0.70%
Rat S214E	Suc-AFK-AMC	44%	35x	1.3%
	MeOSuc-AFK-AMC	37%	33x	1.0%
	Suc-AAPF-S-Bzl	42%	10x	4.1%
	Z-GPR-S-Bzl	110%	8.1x	13%
	Z-GPR-AMC	37%	20x	1.9%
	Z-R-S-Bzl	88%	12x	7.4%
	Z-R-AMC	41%	12x	3.4%
Rat S214K	Suc-AFK-AMC	0.86%	34x	0.025%
	MeOSuc-AFK-AMC	1.6%	76x	0.021%
	Suc-AAPF-S-Bzl	1.9%	14x	0.14%
	Z-GPR-S-Bzl	160%	110x	1.5%
	Z-GPR-AMC	12%	210x	0.055%
	Z-R-S-Bzl	0.86%	100x	0.51%
	Z-R-AMC	-	-	0.21%
Rat S214R	Z-GPR-S-Bzl	-	-	0.16%
	Z-GPR-AMC	-	-	0.0036%
Rat S190G	Suc-AFK-AMC	28%	16x	1.8%
	MeOSuc-AFK-AMC	15%	12x	1.3%
	Suc-AAPF-S-Bzl	54%	0.80x	68%
	Z-GPR-S-Bzl	200%	2.9x	68%
	Z-GPR-AMC	120%	5.3x	24%
	Z-R-S-Bzl	190%	2.6x	72%
	Z-R-AMC	150%	5.5x	27%

these substrates k_{cat}/K_m is decreased to approximately 1% of trypsin by the S214D and S190G mutations, while for the variety of other substrates k_{cat}/K_m is only decreased to 10% of the rate for trypsin. The values of k_{cat}/K_m for the S214Q and S214E enzymes are reduced to 10^{-1} to 10^{-2} of the k_{cat}/K_m for trypsin while the catalytic efficiency of the S214K trypsin drops to 10^{-3} to 10^{-4} of trypsin.

pH Profiles

The Michaelis-Menten parameters were determined between pH 4.0 and pH 11.0 for hydrolysis of Z-GPR-AMC

by trypsin, S214Q, S214E, and S214K trypsins. The observed values of the parameters are plotted in Figure 33 along with theoretical curves which were derived from nonlinear regression analysis. The limiting values, and pK values for the theoretical curves in Figure 33 are listed in Table 7.

Table 7. *Nonlinear regression of pH-rate data to various theoretical kinetic models for wild type, S214Q, S214E, and S214K trypsins. See text for models.*

Trypsin	Model	Parameter	Limiting Values	acidic pK	basic pK
WT	(24)	k_{cat}	4,400 min^{-1}	6.9	–
WT	(25)	k_{cat}	2,000 min^{-1}	6.0	–
			2,600 min^{-1}	7.6	–
S214E	(26)	k_{cat}	2,100 min^{-1}	6.8	9.2
S214Q	(26)	k_{cat}	619 min^{-1}	6.2	10.5
S214K	(26)	k_{cat}	460 min^{-1}	7.8	10.8
S214Q	(27)	K_m	400 μM	6.7	9.9
S214Q	(28)	K_s	400 μM	7.0	9.1
S214Q	(28)	k	0.77	5.8	9.8
S214K	(27)	K_m	510 μM	8.6	10.7
WT	(24)	k_{cat}/K_m	340 μM^{-1} min^{-1}	6.8	–
S214E	(26)	k_{cat}/K_m	10 μM^{-1} min^{-1}	6.9	8.7
S214Q	(26)	k_{cat}/K_m	2.0 μM^{-1} min^{-1}	7.1	9.3
S214K	(26)	k_{cat}/K_m	1.0 μM^{-1} min^{-1}	8.9	10.3

The pH-k_{cat} profile for rat anionic trypsin showed no reduction of activity at high pH as has been reported for the hydrolysis of Z-Lys-S-Bzl.[49] The theoretical curve, derived from nonlinear regression of the data to the single protic model,

$$k_{obs} = \frac{k_{lim}}{1 + 10^{pK_a - pH}} \qquad (24)$$

gave a pK_a of 6.9 with a limiting rate of 4,400 min^{-1}, but does not fit the data well. As can be seen in Figure 33*a*, the theoretical curve for the diprotic model with two limiting rates

$$k_{obs} = \frac{k}{1 + 10^{pK_1 - pH}} + \frac{k'}{1 + 10^{pK_2 - pH}} \qquad (25)$$

fits the data.

The S214E mutation causes a loss of trypsin activity at high pH resulting in a bell shaped pH-k_{cat} curve which fits the diprotic model with a single limiting rate

$$k_{obs} = \frac{k_{lim}}{1 + 10^{pK_a - pH} + 10^{pH - pK_b}} \cdot \qquad (26)$$

Aliquots of S214E trypsin incubated at pH 10 for several hours at 37°C retain 80% of the active sites when assayed with MUGB at pH 8.0 indicating the loss of activity is not due to irreversible denaturation of the protein. Similarly, the S214K mutation inactivates the enzyme significantly so that kinetic data is not easily recorded at neutral pH. There is a modest diminution of activity at high pH which also fits the diprotic model (26). The loss of activity at low pH is not due to irreversible denaturation of the protein since the enzyme was stored at pH 3.0 and diluted to higher pH for assay.

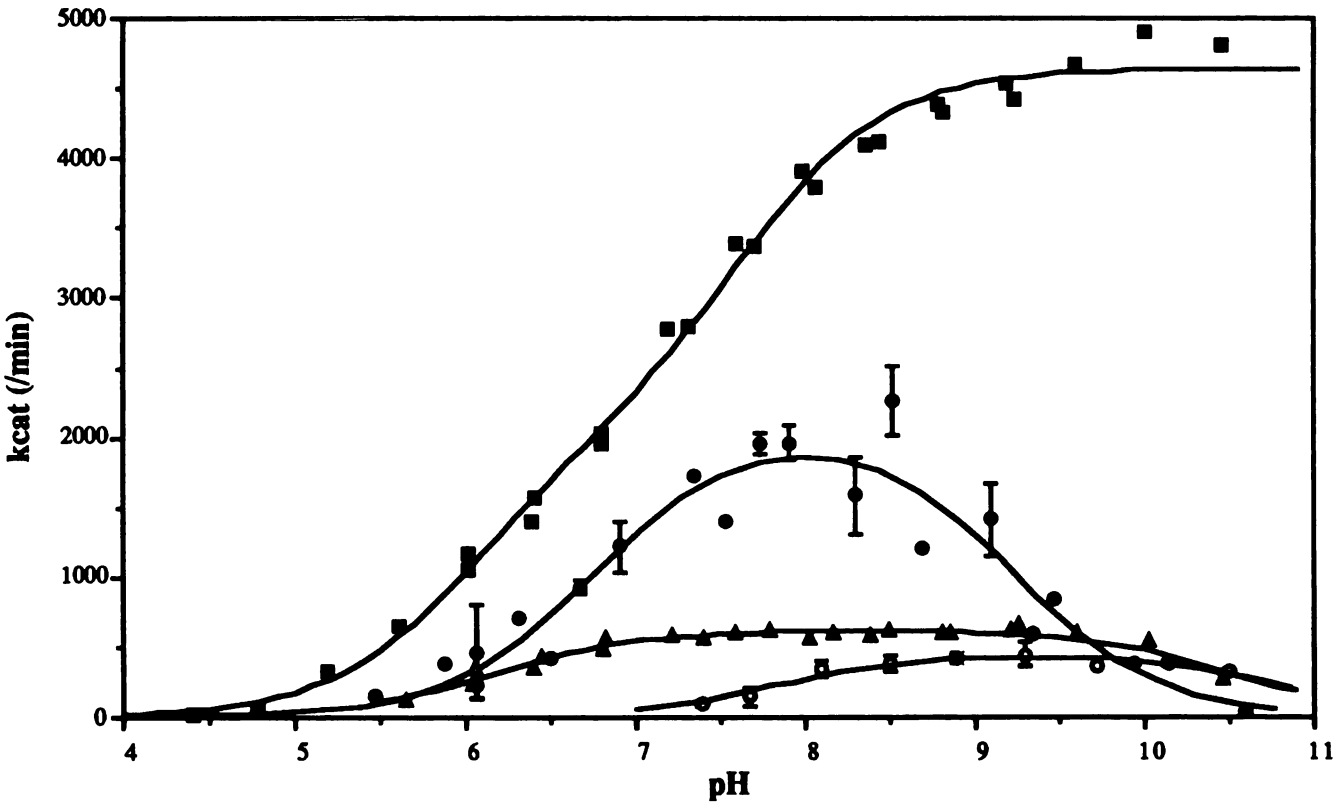
The pH-k_{cat} profile of the S214Q mutation is a flattened, bell-shaped curve. Evidence for an unusual interaction of the Z-GPR-AMC substrate with the S214Q protein can be found by comparing the results between S214E and S214Q trypsin (Table 6). At pH 8.0, the k_{cat} values for the hydrolysis of other substrates by S214Q trypsin are approximately the same as those for S214E trypsin. The reaction of Z-GPR-AMC with S214Q trypsin results in an anomalously low k_{cat} among those substrates for which k_{cat} can be determined (Table 6). Accordingly, the pH-k_{cat} plot reaches a maximum by pH 7.0 and remains level until pH 9.5. In contrast, the pH-k_{cat} plot for trypsin or S214E trypsin increases significnatly in the pH range 7 – 8. Moreover, the kinetic models for the pH-K_m plots for this substrate are extremely complicated.

The pH-K_m plots for wild type and S214E trypsin indicate that pH has little effect

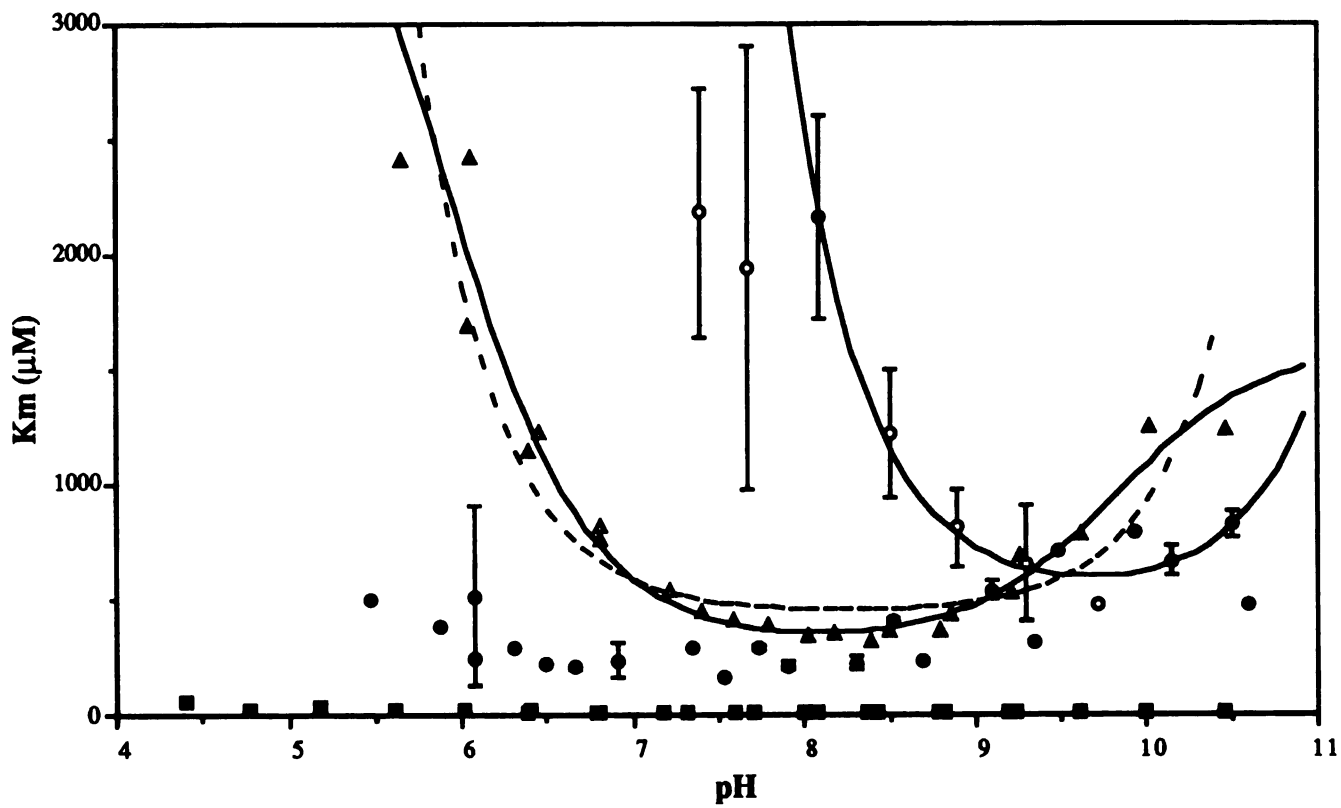
---

Figure 33. pH-rate profiles for the hydrolysis of Z-GPR-AMC by (■) trypsin and (○) K214, (●) E214, and (▲) S214Q trypsins in 50mM buffer, 1mM $CaCl_2$, 100mM NaCl, 1% DMF at 25°C. MES buffer was used from pH 4.5-6.9; MOPS from pH 6.1-8.1; TAPS from pH 7.3-9.3; and glycine from pH 8.5-10.5. (*a*) pH vs k_{cat}, (*b*) pH vs K_m, (*c*) pH vs k_{cat}/K_m. Solid lines are theoretical curves defined in Table 7. The dashed line for the S214Q pH-K_m data is model (27).

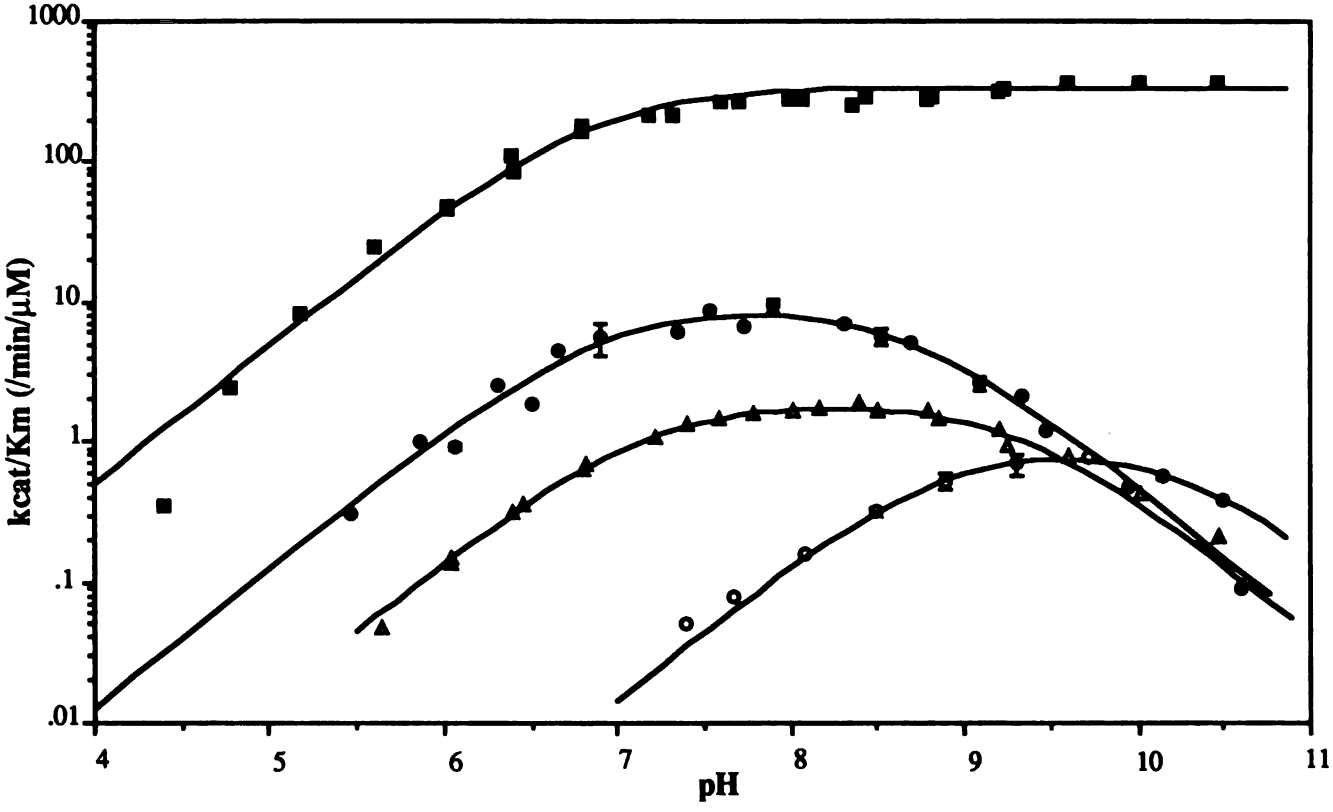
a



b



c



on K_m. However, the S214Q and S214K mutation introduces a dramatic dependence of K_m on pH. The data for both mutants can be fit to model (27).

$$K_{m\text{-}obs} = K_{m\text{-}lim} (1 + 10^{pK_a - pH} + 10^{pH - pK_b}) \quad (27)$$

For S214K trypsin, data below pH 8.0 do not fit the model probably due to errors inherent to attempts to measure K_m values using substrate concentrations well below K_m. The observed pH-K_m profile for S214Q trypsin hydrolysis of Z-GPR-AMC fit better with model (28) (see DISCUSSION).

$$K_{m\text{-}obs} = K_{s\text{-}lim}(1 + 10^{pK_a - pH} + 10^{pH - pK_b})\left(\frac{k_{lim}}{1 + 10^{pK_c - pH} + 10^{pH - pK_d}}\right) \quad (28)$$

The single protic model (24) fits the pH-k_{cat}/K_m data for rat anionic trypsin reasonably well. The S214Q, S214E, and S214K mutants all exhibit diprotic, bell-shaped dependence on pH (26) where the curve for the S214K trypsin is shifted significantly to the basic half of the pH range. Preliminary results for the S214R mutant indicate that in the range pH 8.0 to 10.5, the pH-k_{cat}/K_m data fit the single protic model

$$k_{obs} = \frac{k_{lim}}{1 + 10^{pH - pK_b}} \quad (29)$$

where $k_{lim} = 0.0092$ and $pK_b = 9.6$.

PERTURBATION OF THE ELECTROSTATIC FIELD IN THE ACTIVE SITE OF TRYPSIN – DISCUSSION

Evidence is presented that the kinetic constants represent chemical reaction rates followed by a brief discussion of the need to compare normalized numbers to discern the effects of the mutations on the function of the enzyme. Such an analysis is then used to argue that the S214Q, S214E, and S214K mutations cause similar perturbations of the steric interactions of the mutants with substrates. Evidence for the introduction of a positive charge and a description of the stabilization of the buried charge in the trypsin

structure are discussed. Finally, the electrostatic effects of these mutations on substrate binding and the function of the catalytic triad are discussed and contrasted with the effects of surface charges.

RATE-DETERMINING STEPS IN TRYPSIN CATALYSIS

Pairs of cognate amide and thioester substrates were hydrolyzed to determine whether kinetic parameters reflected chemical reactions or physical processes. In general, the k_{cat} for esters is greater than the k_{cat} for amides which is consistent with the usual assumptions about the kinetic mechanism of the serine proteases (for a discussion, see KINETIC MECHANISM, Chapter 1) and so, except as noted below, the data will be interpreted using these assumptions.

Despite the difference in chemical reactivity between Z-GPR-AMC and Z-GPR-S-Bzl, the turnover numbers (k_{cat}) are the same for wild type, S214D, and S214A trypsins (Table 5). This suggests that the k_{cat} values for the hydrolyses of these substrates by wild type, S214D, and S214A trypsins do not reflect a chemical step on the reaction coordinate. It is possible that a physical step, product dissociation, is the rate-limiting step for the hydrolysis of Z-GPR-S-Bzl in these reactions. Consistent with this hypothesis, the pH-k_{cat} profile of Z-GPR-AMC hydrolysis by rat anionic trypsin exhibits two pK_a values (Table 7) which probably result from the limiting rates due to (1) the protonation state of the His-57 imidazole and (2) product dissociation. The Z-GPR moiety apparently forms an extremely tight complex with rat anionic trypsin which is not disturbed by the S214A or S214D mutations.

Mutations which destabilize the enzyme/product complex apparently accelerate the observed rate of reaction of this substrate while slowing the reactions of other substrates. Indeed, S214K trypsin actually has a higher turnover number for Z-GPR-S-Bzl and an anomalously high k_{cat} for Z-GPR-AMC relative to wild type (Table 6). Similarly, the S214E trypsin has an anomalously high k_{cat} for the hydrolysis of Z-GPR-S-Bzl (Table 6)

while S214Q trypsin shows a normal 10x difference in k_{cat} for these substrates (Table 5). The accelerated turnover of Z-GPR-S-Bzl by S214K trypsin is probably due to destabilization of the trypsin/Z-GPR complex by the steric effects of the Trp-215 side chain in S214E and S214K[148] trypsins and by the electrostatic effects of the S214K mutation discussed below. In support of the electrostatic basis of this rate acceleration, the more severe structural perturbation of S214Q trypsin (Figure 31) is not sufficient to accelerate the reaction as is seen for S214K trypsin. In support of the hypothesis that destabilization of the Z-GPR/trypsin complex by steric effects can accelerate the reaction, the S190G mutation accelerates the hydrolysis of arginine substrates while the rates for lysine substrates and Suc-AAPF-S-Bzl are attenuated (Table 6). The Ser-190 Oγ forms a hydrogen bond with the substrate Arg Nη[129] and is a major determinant of lysine substrate specificity.[63] Evidently, removing this hydroxyl group by the S190G mutation permits efficient catalysis for Arg substrates but destabilizes the trypsin/arginyl-product complexes.

Analysis of Normalized Kinetic Data

The kinetic analysis of substrate/protein interactions involves more than the simple comparison of kinetic constants. For example, the equilibrium constant K_s describes the concentrations of free substrate, free enzyme, and the enzyme-substrate complex in a reaction which is grossly affected by the interactions of these species with the solvent. The K_m values for for the enzymatic hydrolysis of Suc-AFK-AMC and MeOSuc-AFK-AMC provide an excellent example of the magnitude and nature of these solvent effects. The Suc-AFK-AMC substrate binds to rat anionic trypsin with a K_m approximately 3x higher than MeOSuc-AFK-AMC. The less soluble methoxysuccinyl substrate is apparently driven into the active site of trypsin by the hydrophobic effect.[166] In general, this ~3x difference in K_m is observed for each of the trypsins tested despite the differences in absolute binding energy among the various enzymes and these two substrates. Yet, the normalized K_m values are approximately the same for each enzyme and the magnitudes vary depending on

the mutation. The primary data is analyzed whenever possible, but comparison of the normalized values of the kinetic parameters mitigates the effects of forces such as the hydrophobic effect, and the effects of the mutations on the function of the enzyme are therefore revealed independently of the chemistry of the substrates.

CHARGE AND BULK TOLERANCE AT POSITION 214

Steric Effects of the Mutations on Substrate Binding

The S214A mutation reduces the bulk of the 214 side chain and appropriately has almost no effect on K_m. The S214D mutation has a small effect on the K_m values of most substrates but has a larger effect on the K_m values of MeOSuc-AFK-AMC and Suc-AFK-AMC. The intrusion of a lysine,[148] glutamine, or glutamic acid side chain into space previously occupied by a serine side chain and a water molecule causes the local disruption of the position of the neighboring Trp-215 side chain while the remainder of the structure is remarkably similar to the structure of trypsin. The new position of Trp-215 interferes with the anti-parallel β-sheet structure formed between the peptide backbone of trypsin residues Ser-214, Trp-215, and Gly-216 and the backbone of the P_1, P_2, and P_3 residues[a,183] of substrate.[176] In particular, the hydrogen bond formed between the amide proton of Gly-216 and the carbonyl carbon of the P_3 residue is blocked by the Trp-215 side chain in one third of the molecules of S214E trypsin and in S214Q trypsin as described for S214K trypsin.[148]

Results for the S214Q structure further indicate that the Gln-214 side chain is *down* and the Trp-215 side chain is *out* identically to the *out* position in S214E trypsin. Consistent with these results, the Trp-215 side chain is *up* in the S214K structure and *up* proportionately to the occupancy of the Glu-214 down position. Evidently, occupancy of

---

[a] In the Suc-AFK-AMC substrates, alanine, phenylalanine, and lysine are respectively the P_3, P_2, and P_1 residues.[183]

the *down* position by the Glu-, Gln-, or Lys-214 side chains causes the Trp-215 side chain to flip to the *up* position.

This structural result should cause the S214E, S214Q, and S214K mutants to bind tripeptide substrates worse than single residue substrates. The P_1 amide group of single residue substrates form a hydrogen bond only with the carbonyl of amino acid residue 214, but the amino terminal blocking group may interact with the determinants of the S_2 subsite of the Trp-215 side chain. The S214Q mutation causes consistently higher K_m values than the S214E mutation and apparently so high in the case of the MeOSuc-AFK-AMC and Suc-AFK-AMC substrates as to be undeterminable. For 214Q trypsin, the K_m for Z-R-AMC is inexplicably higher than the K_m for Z-GPR-AMC. The S214E mutation causes a 12x increase in the K_m of the single residue Z-R-AMC substrate but a 20-30x increase in the K_m of tripeptide amide substrates (Table 6). The S214E mutation interferes with the binding of tripeptide substrates more than with the single residue Z-R-AMC. This suggests that the S214E mutation causes a disruption of the S_2 or S_3 subsites on the protein presumably due to the altered position of the Trp-215 side chain. The most direct comparison of the functional effects of the altered position of the Trp-215 side chain by the S214E and S214K mutations results from an electrostatically neutral polypeptide substrate. The 30x increase in K_m observed for the hydrolysis of Suc-AFK-AMC and MeOSuc-AFK-AMC by S214E trypsin is also observed for the hydrolysis of the electrostatically neutral Suc-AFK-AMC substrate by S214K trypsin. These results lend further support to the argument that functionally the structural perturbations caused by the S214E and S214K mutations are similar. Moreover, the observation of higher K_m values caused by the S214Q mutation suggests that the structural perturbations are more severe than the perturbations caused by the S214E and S214K mutations. Yet, as noted below, these perturbations do not appreciably affect the turnover rate of the enzyme.

Ionization State of Side Chain 214

Analytical isoelectric focussing (Figure 26) clearly indicates that the S214R or S214K mutations introduce a positive charge into the molecule. Although we have no direct evidence that this charge resides on the Arg-214 or Lys-214 side chains, the basicity of the Arg-214 and Lys-214 side chains as well as detailed theoretical calculations based on the chemical environment of the Lys-214 Nζ inside the protein suggest that the charge resides on the side chains. The crystal structure[148] of the S214K mutant reveals that the ζ-amino group replaces the buried water molecule which forms a hydrogen bond to Ser-214 in the wild type enzyme. This Lys-214 ζ-amino group is stabilized by interactions with the side chain of Thr-229 (3.2Å), the backbone carbonyl of Leu-99 (2.9Å), and a buried water molecule (2.5Å). Although the ζ-amino group of Lys-214 is 4.9Å from the Asp-102 Oδ2, ~7 kcal/mol of electrostatic stabilization due to the opposite charges of these two side chains has been estimated.[148]

The S214A, S214Q, and S214E trypsins all focus at the same isoelectric point. The exact cause of the apparently subtle but consistent perturbation of the electrostatic charge of these proteins from wild type is uncertain. However, Ser-214 Oγ normally forms a hydrogen bond with the catalytic, buried Asp-102 carboxylate, and it has been proposed that Ser-214 plays a significant role in stabilizing this charge.[233] It may be possible that the negative charge on the buried carboxylate of Asp-102 is less well stabilized due to the mutation of Ser-214. It is clear, however, that the S214E mutation does not introduce a negative charge into trypsin since the isoelectric point of S214E trypsin is the same as the isoelectric point which results due to the electrostatically neutral S214A mutation as well as the isosteric and electrostatically neutral S214Q mutation. Furthermore, we observe no perturbation of the acidic pK_a of the pH-k_{cat}/K_m profiles of S214E trypsin which is normally ascribed to the ionization of His-57.

The Glu-214 carboxylate could be easily neutralized by protonation. The stereoelectronic interaction between the Glu-214 and Asp-102 carboxylates suggests that Glu-214, not Asp-102, is protonated according to Gandour values Rule.[73,106,169,248] Furthermore, the serine proteases have evolved an electrostatic potential field to stabilize a negative charge of the catalytic aspartic acid but not at postion 214 (Lei Jin and Robert J. Fletterick, *personal communication*). Evidently, the carboxylate of Glu-214 is protonated under the influence of the powerful, stabilized electrostatic field of Asp-102. This proton would be sitting in a medium with a dielectric constant[78] ~4 positioned 5Å in the "down" conformation, or 2Å, in the "up" conformation, from the Asp-102 Oδ2. Such a carboxylate would be expected to have a dramatically shifted pK_a. The absence of a basic limb in the pH-k_{cat} profile of wild type rat anionic trypsin suggests that the basic limb in the pH-k_{cat} profile of the S214E mutant is due to the Glu-214 side chain. Although there are no spectroscopic data, the author proposes that the basic pK_a of 9.2 in the pH-k_{cat} profile is due to the deprotonation of the Glu-214 side chain. Further, the inactivation of the enzyme at high pH is due to the unfolding of the protein which would be necessary to relieve the proximity of two negatively charged carboxylates. A pK of 9.2 would represent an approximately 4.5 unit shift in the pK_a of a carboxylic acid which would be equivalent to a $\Delta\Delta G \approx 6$ kcal/mol at 25°C according to (30).

$$\Delta\Delta G = (\Delta pK_a)\ (1.36)\ \tfrac{kcal}{mol})^a \qquad (30)$$

The Lys-214 Nζ and the carboxylate in the Glu-214 *down* conformation occupy approximately the same position in the protein. The electrostatic interaction energy due to the negative charge on the carboxylate of Asp-102 has been calculated by the FDPB method to be ~−7 kcal/mol (stabilizing) at Lys-214 Nζ,[148] ~+6 kcal/mol (destabilizing) at the

---

[a] $\Delta G = -RT\ \ln K \Rightarrow \Delta\Delta G = \Delta G_1 - \Delta G_2 = -RT\ \ln K_1 + RT\ \ln K_2 = -RT\ (\ln K_1 - \ln K_2) = -RT\ (2.303 \log K_1 - 2.303 \log K_2) = 2.303RT\ (-\log K_1 - -\log K_2) = 2.303RT\ (pK_1 - pK_2) = (2.303)\ (1.987 \tfrac{cal}{mol})\ (298°)\ \Delta pK = (1.36 \tfrac{kcal}{mol})\ \Delta pK$

Glu-214 *down* Cγ, and ~+7.5 kcal/mol at the Glu-214 up Cγ (J. Vásquez, M. McGrath, J. Guenot, and R. Fletterick, *personal communication*) according to finite difference Poisson-Boltzmann calculations which take into account the different charge distributions on Glu-, Gln-, and Lys-214. The ostensible 4.5 unit shift in the pK_a of the Glu-214 carboxylate observed in the pH-k_{cat} profile agrees well with the electrostatic theory. The ostensible 6–7 unit shift in the pK_a for the Gln-214 side chain observed in the pH-K_m profile is within the range of energies calculated for the Glu-214 *up* conformation. Because of the opposite charges of the ionized Glu-214 and Gln-214 side chains, to a first approximation, the interaction energies could be expected to be of equal value but opposite sign. The different charge distribution on an amide versus a carboxylate would also contribute to a difference in interaction energy. Nevertheless, the S214E and S214K pH-K_m profiles serve as negative and positive controls for the pH-K_m profile of S214Q trypsin. The Gln-214 side chain can be assigned to the ionization which is evident in the pH-K_m profile with reasonable certainty because the Glu-214 side chain does not cause a similar effect on the pH-K_m profile while the Lys-214 side chain *does* cause such an effect. Alternatively, it would have to be explained why burying a neutral amide side chain in a protein would induce some other side chain to protonate at low pH. Calculations of the electrostatic energies for interactions this close become susceptible to large errors.[82] The calculation of the electrostatic interaction energy for the Gln-214 up conformation is actively being pursued.

Table 8. *The effect of the negative charge of Asp-102 on electrostatic potentials for glutamine, glutamate, and lysine side chains at position 214. (see the text for references)*

Site of Electrostatic Potential	Δφ Due to Asp-102 (kcal/mol)	Predicted ΔpK_a	Putative ΔpK_a
Lys-214 Nζ	7	+5	?
Glu-214 *down* Cγ	6	+4.5	4.5
Glu-214 *up* Cγ	7.5	+5.5	4.5
Glu-214 *down* Oε2	8.5	+6	6–7
Glu-214 *up* Oε2	14	+10	6–7

ELECTROSTATIC EFFECTS OF THE MUTATIONS

Electrostatic Effects of the Mutations on Substrate Binding

Electrostatic repulsion between the Lys-214 ζ-amino group and positively charged substrates is evident in three cases. First, the observed increase of the k_{cat} for the Z-GPR-S-Bzl ester substrate (Table 6) may reflect electrostatic repulsion of the product. In particular, the more severe steric perturbations caused by the S214Q mutation do not result in a similar rate acceleration. Second, consistent with this hypothesis, the positively charged Lys-214 side chain, but not the Glu-214 side chain, apparently induces pH dependent binding of the positively charged, tripeptide amide substrate Z-GPR-AMC. While the increase in K_m at low pH might have been ascribed to unfolding of the protein, we note that at pH 6.0 the S214K protein binds specifically to both BPTI and benzamidine affinity resins and that the S214K crystal structure was determined at pH 8.0.[148] Finally, we note that although crystallographic data reveal no gross unfolding of S214K trypsin at pH 8.0, and although crystallographic and kinetic data suggest that the steric perturbations of the S214E, S214Q, and S214K mutations are similar the K_m values for the hydrolysis of positively charged amide substrates by S214K trypsin are 76-210x increased above the K_m values for rat anionic trypsin. No substrate on S214E trypsin has a K_m more than 35x greater than that of wild type while S214Q trypsin as amide K_m values 16x and 44x wild type. Two observations in particular highlight the electrostatic nature of the increased K_m values. The single residue amide substrate Z-R-AMC apparently has a K_m so high as to be unmeasurable for S214K trypsin although hydrolysis of the substrate is catalyzed at an easily measurable rate and this K_m can be measured for S214E trypsin. Apparently, electrostatic repulsion confounds a kinetic determination of the binding constants of Z-R-AMC for the S214K mutant. Secondly, the S214K mutant shows a relative preference for binding the electrostatically neutral Suc-AFK-AMC substrate over the positively charged MeOSuc-AFK-AMC substrate (Table 6) although the two substrates

differ structurally only by a methyl group which does not sterically affect binding of the substrate to any of the other enzymes.

The electrostatic effects of the S214K mutation are considerable. What the bovine cationic trypsin cannot accomplish with a net surface charge 12.5 units more positive than rat anionic trypsin,[201] the S214K mutation accomplishes with a single buried point charge: repulsion of a positively charged substrate. Still, the primary determinant of substrate specificity in trypsin, Asp-189, productively binds the substrate near the catalytic machinery of the enzyme and allows us to explore the effects of the mutation on the catalytic function of the enzyme.

Model **(28)** gives excellent agreement with the observed pH-K_m plot of S214Q trypsin and was derived by adding the pH dependence to equation **(5)** (p. 16). Therefore,

$$k_{lim} = \frac{k_3}{k_2 + k_3} \qquad (31)$$

and the observed perturbations of k_{lim} represent the shift in the relative rates of acylation (k_2) and deacylation (k_3). For amides, because $k_2 \approx 0.1k_3$, $k_{lim} \approx 0.9$, does not change, and therefore, does not show up in the pH-K_m plot. At acidic pH, the Gln-214 side chain evidently is protonating, resulting in electrostatic repulsion of the positively charged substrate, exactly as was seen for the Lys-214 side chain. Detailed FDPB calculations (Table 8) indicate that this ionization state is energetically accessible. At low pH, the acylation rate is slowed and the deacyation rate increased as is seen for the S214K trypsin. At high pH, the deacylation rate apparently slows relative to the acylation rate.

The Effects of Buried Charges on Catalysis

The catalytic efficiency parameter k_{cat}/K_m is inappropriate for the comparisons made in this paper because the mutations affect both catalytic function which is primarily reflected in k_{cat} and binding affinity which is primarily reflected in K_m. Therefore, we

examine k_{cat} as the kinetic parameter which best describes the catalytic effect of these mutations.

The S214K mutation has a significant effect on k_{cat} at pH 8.0 whenever k_{cat} is a measure of the rate of a chemical reaction (and therefore not in the case of the Z-GPR-X substrates); the surface charges of the bovine cationic isozyme, the S190G mutation, and the S214(A,D,Q,E) mutations do not have a consistent significant effect on k_{cat}. The S214(A,D,Q,E) mutations cause a reduction in k_{cat} to approximately 40% of the catalytic function of trypsin while the S214K mutation causes a reduction to approximately 1% of trypsin activity. Moreover, despite the more severe structural perturbations caused by the S214Q mutation, values of k_{cat} are 9 – 60% of wild type. The same effects on the rate are seen for acylation by electrostatically positive (MeOSuc-AFK-AMC) and neutral (Suc-AFK-AMC) substrates, deacylation of a specific, positively charged substrate (Z-R-S-Bzl), and deacylation of a non-specific, negatively charged substrate (Suc-AAPF-S-Bzl). The variety in charge and chemistry of these substrates indicates that the decreases in k_{cat} are caused by effects of the mutations on the function of the enzyme and are not due to the differences among the substrates.

PERTURBATION OF THE ELECTROSTATIC FIELD IN THE ACTIVE SITE OF TRYPSIN – CONCLUSIONS

Nature has evolved the structure and composition of the catalytic triad[a] – Asp-102, His-57, and Ser-195 – for the hydrolysis of peptide bonds by the trypsin and subtilisin families of the serine proteases and for the hydrolysis of phospholipids by lipase.[34,242] The workings of the serine protease catalytic triad have long been studied as a quintessential problem of enzymology and the kinetic and chemical mechanisms of the serine proteases are well described.[68] The residues in the catalytic triad and the amide protons which define the oxyanion hole have been implicated as catalytic determinants

[a] The catalytic triad numbered here is for the trypsin family only.

based on chemical modification experiments,[165] structure analysis,[28] and, more recently, mutagenesis experiments.[37,41,49,203,238]

Like any catalyst, trypsin must stabilize the transition state of a reaction relative to the ground state thus decreasing the free energy of activation, $\Delta G^{\ddagger}$, for each microscopic step along the reaction coordinate.[138,187,243] The chemistry of the catalytic residues suggests possible mechanisms for transition state stabilization, yet the exact nature of the forces at work in the transition state are in dispute. Proton inventory studies have been interpretted to indicate that at least two protons move in the transition states of certain serine protease-catalyzed reactions, and it has been proposed that these protons move between Asp-102 and His-57 as well as between His-57 and Ser-195 in a charge relay mechanism[189,205] similar to that proposed by Blow.[27] The charge relay and electrostatic mechanisms involve mutually exclusive arrangements of the proton between Asp-102 and His-57 in the transition states of the reactions (Figure 7).

The electrostatic effects of surface point mutations on the function of subtilisin have been extensively examined.[177,178,220,239] The effects of these mutations on the function of the catalytic triad were apparently minimal due to screeening by bulk solvent.[69,80,81] The effectiveness of bulk water at screening surface charges is apparent in the observed Michaelis-Menten parameters for cow cationic trypsin and rat anionic trypsin. Although these isozymes differ by 12.5 units (histidine is counted as 0.5 units) of surface charge,[201] both trypsins hydrolyze the tripeptide lysine substrates with nearly equal efficiency (Table 6).

Therefore, in an attempt to perturb the electrostatic potential at the active site of trypsin, we have explored the bulk and charge tolerance of the rat anionic trypsin molecule by directed mutagenesis of the solvent inaccessible residue, Ser-214, to Ala, Asp, Gln, Glu, Lys, and Arg. Rat anionic trypsin apparently accepts these mutations and folds correctly. Each of these mutant trypsins is present in the periplasm at comparable levels,

specifically binds to immobilized BPTI during purification, and catalyzes the hydrolysis of trypsin substrates. A mutant protein which fails any of these tests might be misfolded. The broad tolerance of trypsin to mutagenesis of Ser-214 seems to result from the presence in the structure of buried water molecules as well as the powerful electrostatic field of Asp-102.[201]

That similar or more severe structural perturbations arise in the S214E and S214Q mutants without a significant effect on the chemical activity of the enzyme indicates that the loss of activity caused by the S214K mutation is not due to structural effects. Stabilization of the negative charge on Asp-102 by the "aspartate hole"has been proposed as an important determinant of catalytic function.[233] For the *down* position of the Glu-214 side chain, the *down* position of the Gln-214 side chain, and the Lys-214 side chain, this hydrogen bond is lost. Without a crystal structure, it is not possible to know if the S214A mutation deletes this particular hydrogen bond or if a water molecule occupies the site. However, it appears that this particular hydrogen bond is not an important component of the aspartate hole. We note that these are the stereochemically less active *anti* lone pair electrons of Asp-102.[73,106,169,248] Finally, the loss of stabilization of the Asp-102 charge by this hydrogen bond is compensated by the positive charge of Lys-214.[148] Therefore, we conclude that the major effect of the S214K mutation is electrostatic in nature.

These results are inconsistent with a charge relay mechanism insofar as charge relay requires the deprotonation of His-57. The finite difference Poisson-Boltzmann method[79] for calculating electrostatic potentials in rat anionic trypsin are corroborated by the experimental results in this thesis.[148] Those calculations indicate that protonation of Asp-102 results in an electrostatic potential field within trypsin which is defined by the surface charges of the protein.[201] The two isozymes of trypsin, anionic and cationic, set up negative or neutral electrostatic fields in the active site when Asp-102 is neutralized as it would be upon protonation, yet both hydrolyze substrates at comparable rates. Since a

transition state which includes a protonated Asp-102 carboxylate and deprotonated His-57 imidazole would lack a positively charged species, the activation energy for the transition states in the two isozymes would be different. An electrostatically neutral transition state, which would be unaffected by the net surface charge, might be set up by protonating the oxyanion of the tetrahedral intermediate, but then electrostatic mutations should have no effect on the activation energy of the reaction. However, electrostatic mutations *do* have an effect on the activation energy of the reaction indicating the presence of a positive charge in the transistion state. Therefore, Asp-102 cannot be protonated in the transition-state of trypsin catalyzed reactions.

In order to characterize the effects of the Lys-214 mutation on the function of the catalytic triad, the hydrolysis rates of several different substrates were determined and the hydrolysis of these substrates by different trypsin mutants were examined to better define the relationships between chemistry, structure, and function in this system. The same effect on k_{cat} is seen in two different reactions: acylation and deacylation of the same enzyme. Common to both of these reactions is the function of the catalytic His-57 which activates Ser-195 during acylation and a water molecule during deacylation (Figure 7). The destabilization of positive charge is the primary electrostatic effect observed when a positive charge is fixed at one end of the catalytic triad 4.9Å from the carboxylate of Asp-102. The data presented are consistent with the interpretation that a positive charge is developed on His-57 in the transition states for acylation and deacylation. We propose this is direct experimental evidence that one function of Asp-102 is to stabilize positive charge which develops on His-57 during catalysis.[201,233] Given the option of carrying out a proton transfer reaction or maintaining a charge couple, the catalytic triad of trypsin evidently maintains a charge couple. To test this hypothesis further, the orientation of the carboxylate in the catalytic triad was shifted. These experiments are described in Chapter 4.

The magnitude of the effects of charges in the interior of a protein are significant. The charge couple in the transistion states of trypsin-catalyzed reactions is evidently stabilized by the exclusion of bulk solvent. The magnitude of this effect evidently manifests itself in (1) the stabilization of Lys-214, (2) the protonation of Glu-214, and (3) the ionization of Gln-214. Thus, although much more work remains to be done, some indication of the powerful electrostatic interactions of positive, negative, and neutral side chains with a negative charge in the interior of a protein has been obtained. Theoreticians have appreciated the possibility of the importance of such interactions[201,233] but have been unable to demonstrate that it exists. It may now be justifiable to describe the catalytic mechanisms of other enzymes which catalyze ionic reactions in terms of the large energies available due to electrostatic interactions of charges in the absence of bulk solvent.

CHAPTER 4

RECONSTRUCTION OF THE CATALYTIC MECHANISM OF MUTANT TRYPSINS WHICH LACK A CATALYTIC ASPARTIC ACID

INTRODUCTION

As the field of protein engineering advances through the *de novo* design of polypeptides which mimic the catalytic function of known enzymes,[89] information on the subtleties of the function of the known enzymes becomes more important. It is not enough to mimic the known spatial arrangements of catalytic residues in solvent if a less critical arrangement removed from bulk solvent is the key. Extensive theoretical calculations,[122,201,224,233] and previous results (Chapter 3)[37,41,49,148,203,229,238] suggest that the electrostatic charge couple of Asp-102 and His-57 is the primary determinant of the chemical activity of the nucleophilic Ser-195. As an experimental test of this theory, we have attempted to regenerate the catalytic function of mutant trypsins which have no catalytic aspartic acid. It is assumed in this chapter that wild type or near wild type catalytic activity is observed only when the established catalytic mechanism entailing activation of the Ser-195 hydroxyl group by the His-57 imidazole is at work.

MATERIALS AND METHODS

DOUBLE PRIMER MUTAGENESIS OF M13

The D102A, D102S, and D102N[49] mutants were prepared before the advent of plasmid-based Kunkel mutagenesis using double primer mutagenesis of M13 phage.[46,152,249] This technique seems baroque compared to Kunkel mutagenesis but is subject to different problems and so the techniques are complementary. The ability to identify single base-pair differences at a specific site could still prove useful for difficult constructions or mutations. This or a similar approach could even be coupled with the Kunkel method to identify events which are extremely rare for whatever reason. In addition, the screening technique used in this approach forms one basis for cloning new genes using oligonucleotides.

M13 is a single-stranded DNA phage which infects male *E. coli*; *i.e.*, those which have an F pilus. The life cycle of the M13 phage has been described in almost chemical detail[123] and complete systems (which are not discussed here) have been engineered for cloning, sequencing, and mutagenesis.[152] It suffices here to explain that the phage is propagated by growing on F^+ or F' strains of *E. coli* which have been selected for good growth and maintenance of M13.[152] M13 does not lyse the cells it infects so growth in liquid medium gives a good yield of phage, yet M13 does slow the growth of *E. coli* enough to allow the discrimination of plaques (small clear spots of infected cells on a turbid lawn of growing *E. coli*) on solid media.

When nitrocellulose paper is laid on the surface of a plate which has plaques growing on it, some phage adhere to the paper. The nitrocellulose is removed from the plate and baked to denature the proteins of the phage and bond the DNA. The DNA adheres to the paper even through several steps of heating in aqueous baths of various buffers. A radiolabeled oligonucleotide with the proper sequence will bind ("hybridize") to the DNA on the nitrocellulose filter depending on how well the sequence of the oligonucleotide matches the sequence of the DNA. The nitrocellulose filter with oligonucle-

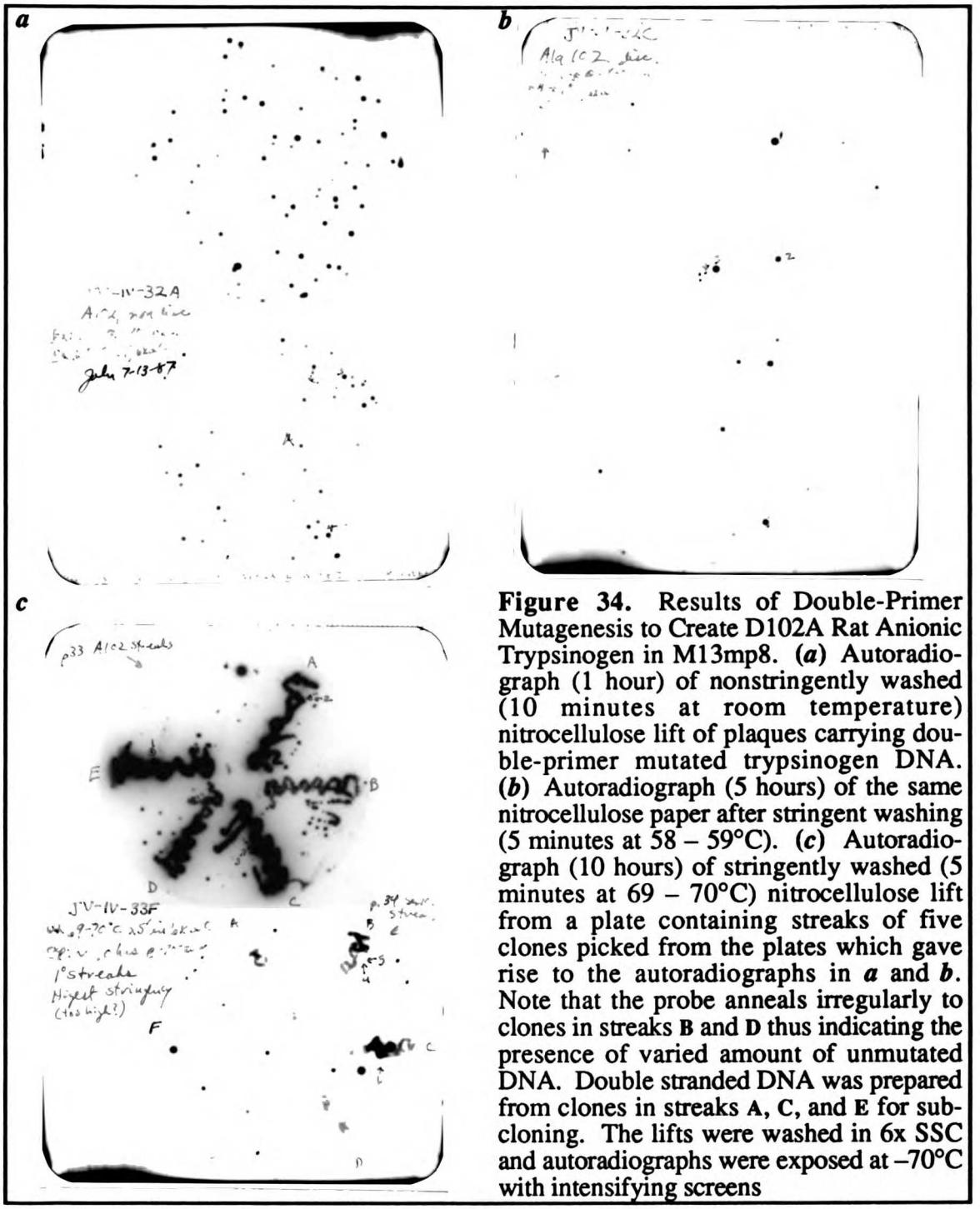


Figure 34. Results of Double-Primer Mutagenesis to Create D102A Rat Anionic Trypsinogen in M13mp8. (***a***) Autoradiograph (1 hour) of nonstringently washed (10 minutes at room temperature) nitrocellulose lift of plaques carrying double-primer mutated trypsinogen DNA. (***b***) Autoradiograph (5 hours) of the same nitrocellulose paper after stringent washing (5 minutes at 58 – 59°C). (***c***) Autoradiograph (10 hours) of stringently washed (5 minutes at 69 – 70°C) nitrocellulose lift from a plate containing streaks of five clones picked from the plates which gave rise to the autoradiographs in ***a*** and ***b***. Note that the probe anneals irregularly to clones in streaks **B** and **D** thus indicating the presence of varied amount of unmutated DNA. Double stranded DNA was prepared from clones in streaks **A**, **C**, and **E** for subcloning. The lifts were washed in 6x SSC and autoradiographs were exposed at –70°C with intensifying screens

otide annealed to the bound DNA is then washed[249] in 6x SSC[a] at incrementally increasing temperatures so that an oligonucleotide with even a single mismatched base will melt from

[a] 20x SSC (3M NaCl, 0.3M sodium citrate, 0.01M EDTA, pH'd to 7.2 with HCl) is prepared and diluted.

the bound DNA,[199] but a perfectly matched oligonucleotide will not. Radiation emanating from the probe will produce an image on X-ray film (autoradiography) indicating the location of the radioactive label on the nitrocellulose. This image is a replica of the original plate and pinpoints the location of a plaque which contains the sequence of the probe. The phage in that plaque are then genetically purifed and tested again to verify the presence of the sequence before the DNA sequence is determined (Figure 34). Detailed protocols for these techniques have been published.[141,152,249]

Double Primer Mutagenesis Protocol

Double primer mutagenesis was successfully carried out according to this protocol by Dr. Charles Craik using the M13 sequencing primer and a mismatched oligonucleotide to introduce mutations into single-stranded M13mp8-trypsinogen DNA.

Solutions

10x Kinase buffer: 0.5M Tris HCl, pH 7.8
0.1M $MgCl_2$
0.1M dithiothreitol (DTT)
0.01M spermidine

10mM ATP

10x RTase buffer: 0.5M Tris HCl, pH 8.0
0.2M KCl
70mM $MgCl_2$
1mM EDTA
0.1M β-mercaptoethanol

10x dNTP values: 1.5mM each dGTP, dTTP, dCTP
0.5mM each dATP, ATP

Kinasing

1. In a 0.5ml eppendorf tube combine
   - 100pmol of mutagenic oligonucleotide
   - 1.5µl 10 kinase buffer
   - 0.5µl 10mM ATP
   - water to 19µl
   - 1.0µl T_4 polynucleotide kinase (10U / New England Biolabs)
2. Incubate at 37°C for 45 – 60 minutes.
3. Heat inactivate the kinase at 65°C for 10 minutes.

Annealing

1. In a 0.5 ml eppendorf tube, combine the following in a total volume of 6µl or dry down and resuspend in the next step
   – 1pmol of single-stranded template DNA
   – 20pmol of M13 sequencing primer
2. Add
   – 1µl 10x RTase buffer
   – 3µl of phosphorylated mutagenic oligonucleotide from above
   – water to make 10µl total volume
3. a. Heat at 65°C for 5 minutes.
   b. Leave at room temperature for 5 minutes.

Synthesis

1. Add to the annealing reaction
   – 2µl 10x dNTP values
   – 1µl 10x RTase buffer
   – 3µl water
   – 3µl T_4 DNA Ligase
   – 1µl Klenow fragment (9U)
2. Incubate at 15°C for 6 – 8 hours.
3. a. Add 2µl of synthesis reaction to 100µl of Ca^{++} competent *E. coli* (F^+)
   b. Incubate on ice 1 hour.
4. Heat shock at 37°C for 5 minutes (or 42° for 1 minute).
5. a. Add 900µl of LB media to cells. This is the 1x stock.
   b. Serially dilute to make 10^{-1} and 10^{-2} stocks.
6. Mix the following in a sterile tube and pour quickly on a warmed LB plate:
   – 3 mls of melted top agar (0.8% agarose in LB)
   – 75 µl of a saturated overnight of F^+ *E. coli*
   – 100 µl of one of the 1x, 10^{-1}, or 10^{-2} dilutions from Step 5.
7. Incubate the cooled plates at 37°C overnight. Plaques should be evident.

Screening

1. Chill the plates at 4°C for approximately 30 minutes.
2. Lay the sterile (?) nitrocellulose filter paper on the surface of the plate. Mark the orientation of the filter on the plate.
3. After 2 minutes, peel the nitrocellulose from the plate and air dry 10 minutes.
4. Bake the filters under vacuum at 80°C for 2 hours.
5. Prehybridize the filters in 0.2% SDS, 10x Denhardt values Solution, 6x SSC.
6. Radiolabel the mutagenic oligonucleotide to prepare the probe.
   a. In a 1.5ml eppendorf tube combine
   – 2µl 10x Kinase buffer
   – 10pmol of mutagenic oligonucleotide

– 1µl T_4 polynucleotide kinase (10U)
– 200µCi γ-^{32}P-ATP
– water to 20µl

b. Incubate at 37°C for 45 minutes.
c. Heat inactivate kinase 2 minutes at 95°C.
d. Add 5µg of poly-dA in 500µl water.

7. Sep Pak purify the radiolabled probe to remove unincorporated ^{32}P.
   a. Activate Sep Pak with 5mls MeOH.
   b. Wash Sep Pak with 10mls water.
   c. Apply kinased sample to Sep Pak by gravity feed in 100µl lots.
   d. Rinse probe kinasing tube contents into Sep Pak.
   e. Wash out unincorporated ^{32}P with 20mls of water.
   f. Elute with 3ml 1:1 MeOH:water into 10 eppendorf tubes.
   g. Speed Vac samples to dryness (~3.5 hours).
   h. Resuspend probe in 500µl water.
8. Hybridize the probe to the nitrocellulose filters in hybridization solution (6x Denhardt values solution, 6x SSC) at room temperature overnight with rocking.
9. a. Wash the nitrocellulose filters at room temperature 10 minutes in ~500mls of 6x SSC (non-stringent).
   b. Autoradiograph at –70°C with intensifying screens.
10. a. Using a Geiger counter at point blank range, monitor the radioactivity on the filters and reduce the signal to approximately 1% of the non-stringent activity or to <0.5mR/hr by repeated washing as in b.
    b. Wash the nitrocellulose filters in 6x SSC for 5 minutes at temperatures incrementally increasing by 2 – 5°C beginning at 10° below theoretical T_m.
    $T_m = 4°$ per G or C + 2° per A or T – 5°C per mismatch
    c. Autoradiograph at –70°C with intensifying screens.

Genetic Purification and Identification

1. a. Streak potentially positive plaques onto a lawn of *F+ E. coli.*
   b. Lift streaks of plaques, wash and autoradiograph filters to identify pure positives.
2. To prepare single-stranded DNA, in a sterile culture tube
   a. Add
   – 2mls of 2xYT
   – 100µl of a saturated overnight of *F+ E. coli.*
   – one plaque cored from a plate with a sterile pasteur pipet.
   b. Grow ~8 hours
   c. Prepare single stranded DNA as described under "Preparation of Uracil-laden Single-stranded DNA." Save an aliquot of the phage as a stock.
3. To prepare double-stranded viral DNA (RF), do a plasmid miniprep using the pellets from Step 2.

CONSTRUCTION OF DOUBLE MUTANT PLASMIDS

Expression plasmids encoding D102(D,A,S,N):S214(S,D,E,Q) trypsins were constructed in a p^mT_3 plasmid by mutating an Eco RI site in the promoter region to a Bgl II site to generate p^mT_c (Figure 12). The D102X mutants were prepared first by cloning the Eco RI – Sty I fragments containing the D102X mutation into p^mT_c. The D102(D,A,S,N):S214(S,D,E,Q) series was prepared by cloning the Xho I – Sty I fragment containing the S214X mutation into each of the p^mT_cD102(D,A,S,N) plasmids.

Restriction Digests

Plasmid DNA from the miniprep or midiprep protocols in Chapter 3 or the RF DNA above can be used to construct plasmids. Double-stranded fragments can be cut using restriction enzymes[a] according to the manufacturer values instructions. The New England Biolabs Catalog should always be consulted for notes on the use of each new enzyme. In general, analytical digests were performed in 20µl incubations while preparative digests were done in 50µl or 100µl incubations.

DNA Fragment Isolation and Purification Protocol

DNA fragments were routinely isolated and purified according to the following protocol.

Set Up:

For each fragment, label for **LOW** salt buffer:
- 5ml syringe.
- 17x100mm snap-capped culture tube.
- small weigh boat.

Fragment Isolation by Gel Purification:

1. Pour a 1% Agarose in 1x TAE gel with a single wide well.[b]
2. Prepare sample
   a. Make the sample 1x in sample buffer *i.e.*, add 20 µl of 5x dye to a 100 µl sample)

---

[a] Preferably from New England Biolabs, Inc., 32 Tozer Road, Beverley, MA 01915-9931 (617) 927-5054

[b] The standard thin well will hold ~120 µl while the thicker well will hold ~250 µl of sample.

b. Heat the sample 10 minutes at 60°C. Do not spin down; use the evaporation to help concentrate the sample.

3. Electrophoresis
   a. Load the sample in the large lane.
   b. Load 10 µl of DNA Markers in the small lane.
   c. Electrophorese the sample 30 minutes at 200V.
4. Fragment Isolation (Keep the gel shielded from light as much as possible!)
   a. Stain gel 5 minutes in a 500ml solution of 1.4µg/ml ethidium bromide in 1x TAE.
   b. Destain the gel in a 500ml solution of 1x TAE for 10seconds..
   c. Photograph the gel using the Foto/Phoresis unit.
   d. Cut the band from the gel using a new razor blade.
   e. Put the agarose sliver in a labeled 5ml syringe.
5. Preparation of fragment for purification.
   a. Extrude the fragment through a 22 gauge needle into a labelled 17x100mm snap-capped culture tube.
   b. Deliver 3mls of Elutip-d **LOW**[b] salt buffer into the small weigh boat.
   c. Rinse the needle and syringe with 3x1ml lots of **LOW** salt buffer into the 17x100mm snap-capped culture tube.
   d. Seal the cluture tube, vortex the sample to homogeneity, and cover the sealed tube with foil.
   e. Shake the sample tube overnight at 37°C x 300rpm.
6. Purify fragment according to Elutip-d standard protocol.

Fragment Purification by Elutip-d[c] Chromatography:

1. Activate Elutip-d Column with 2mls of **HIGH.**[d]
2. Prime Elutip-d Column with 5mls of **LOW.**
3. Load the DNA in **LOW** through a 0.45 µ filter onto the Elutip-d Column.
4. Rinse the filter and the Elutip-d Column with 3mls of **LOW.**
5. Elute the DNA from the Elutip-d Column with 400 µl of **HIGH.**
6. Ethanol precipitate the DNA.
   a. Add 1.1ml of -20°C 100% ethanol.
   b. Mix and incubate the preparation on ice 10 minutes.
7. a. Microfuge the preparation 30 minutes.
   b. Aspirate and discard the supernatant.
   c. Rinse the pellet with ~100 µl of -20°C, 70% Ethanol.
8. Dry the DNA in a speed vac.[e]
9. Resuspend the resulting pellet in 20 µl TE.

---

[b] LOW Salt Buffer is 0.2M NaCl, 1.0mM EDTA, 20mM Tris•HCl, pH 7.3 – 7.5

[c] These columns (cat #NA010/2) are available from Schleicher & Schuell, Inc., Keene, NH 03431, (800) 245-4024.

[d] HIGH Salt Buffer is 1.0M NaCl, 1.0mM EDTA, 20mM Tris•HCl, pH 7.3 – 7.5

[e] 5 minutes under a reasonable vacuum (1-2 Torr) is usually sufficient.

10. Analyze a 2 µl sample on a 1% agarose gel (TAE).

DNA Ligation Protocol

Fragments were ligated according to the following protocol.

Set Up:

1. Prepare **Buffer** Mix:
   - -2 µl 0.7 mM ATP per ligation
   - -2 µl 10x Ligase Buffer[a] per ligation
   - -2 µl 10 mM Spermidine per ligation
2. Prepare T_4 DNA Ligase **Enzyme** mix:
   - -0.5 µl 1 M dithiothreitol per ligation
   - -1 µl T_4 DNA Ligase per ligation

Procedure:

1. Add water to each ligation mix to adjust the final volume to 20 µl.
2. Add 6 µl of **Buffer** mix.
3. Add VECTOR and INSERT DNA.
4. Add 1.5 µl of **Enzyme** Mix.
5. Incubate at room temperature 1-2 hours.[b]
6. Transform 3µl into frozen Ca^{++} cells.

PROTEIN PREPARATION

p^mT_c plasmids were transformed into *E. coli* strain X90, genetically purified, and sequenced. The relative expression and activity of the sixteen proteins was tested using a western blot and the trypsin activity gel. The trypsins listed in Table 9 were then purified from high density 10l fermentor cultures. Each of these techniques is described in Chapter 3.

KINETICS

Either 2.5 nM or 50 nM enzyme was incubated with 1 mM to 500 nM substrate in 100 mM NaCl, 20 mM $CaCl_2$, 50 mM Tris, pH 8.0 with 1% DMF for amides or 2% DMF and 250 µM 4,4'-dithiodipyridine for esters. The progress of the hydrolysis

---

[a] 10x Ligase buffer is 0.5M Tris-HCl (pH 8.0), 0.1M $MgCl_2$

[b] Sticky end ligations run 1-2 hours. Blunt end ligations should be run overnight at room temperature.

reactions was monitored by visible spectroscopy for thioesters[49] or fluoresence spectroscopy for amides[84] as described. Background solvolysis rates were subtracted for esters substrates. The kinetic constants were determined from the initial velocities for five concentrations of substrate using Eadie-Hofstee plots.[68] These constants did not vary significantly from the statistically weighted equations of Cornish-Bowden and Wharton[45] which were used to calculate k_{cat}/K_m from velocities determined at five substrate concentrations when k_{cat} and K_m could not be measured.

RECONSTRUCTION OF THE CATALYTIC MECHANISM OF MUTANT TRYPSINS WHICH LACK A CATALYTIC ASPARTIC ACID - RESULTS

SCREENING IN CRUDE EXTRACTS

Periplasmic fractions were prepared from overnight cultures of *E. coli* strain X90 transformed with $p^mT_cD102(D,A,S,N):S214(S,D,E,Q)$ plasmids and grown in the presence of 1mM IPTG to induce production of trypsin. After fractionation by denaturing polyacrylamide gel electrophoresis,[135] and blotting to nitrocellulose, the presence of trypsin proteins was determined by immunoaffinty of trypsin antibodies (Figure 35).

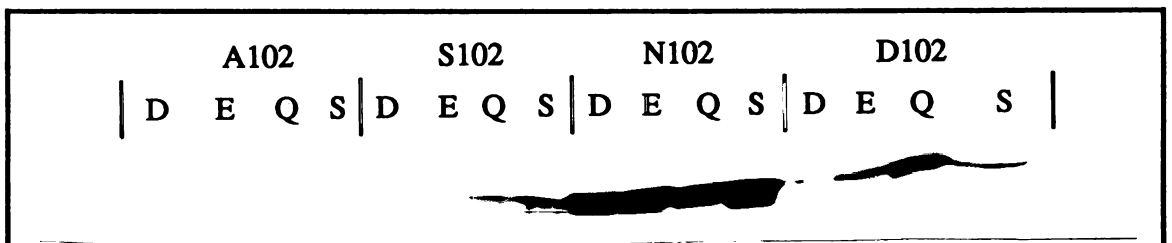


Figure 35. Rabbit Anti-Rat Anionic Trypsin Western Blot of Denaturing PAGE-fractionated Periplasmic Extracts of *E. coli* Strain X90 Transformed with One of the Plasmids $p^mT_cD102(D,A,S,N):S214(S,D,E,Q)$ and Induced with 1mM IPTG.

To test for trypsin activity, the periplasms were also fractionated by non-denaturing polyacrylamide gel electrophoresis. The SDS detergent was soaked out of the gel and Ca^{++} soaked into the gel before the gel was sealed on top of a polyacrylamide gel impregnated with phenol red pH indicator and tosyl-arginyl methyl ester (TAME) at basic pH.

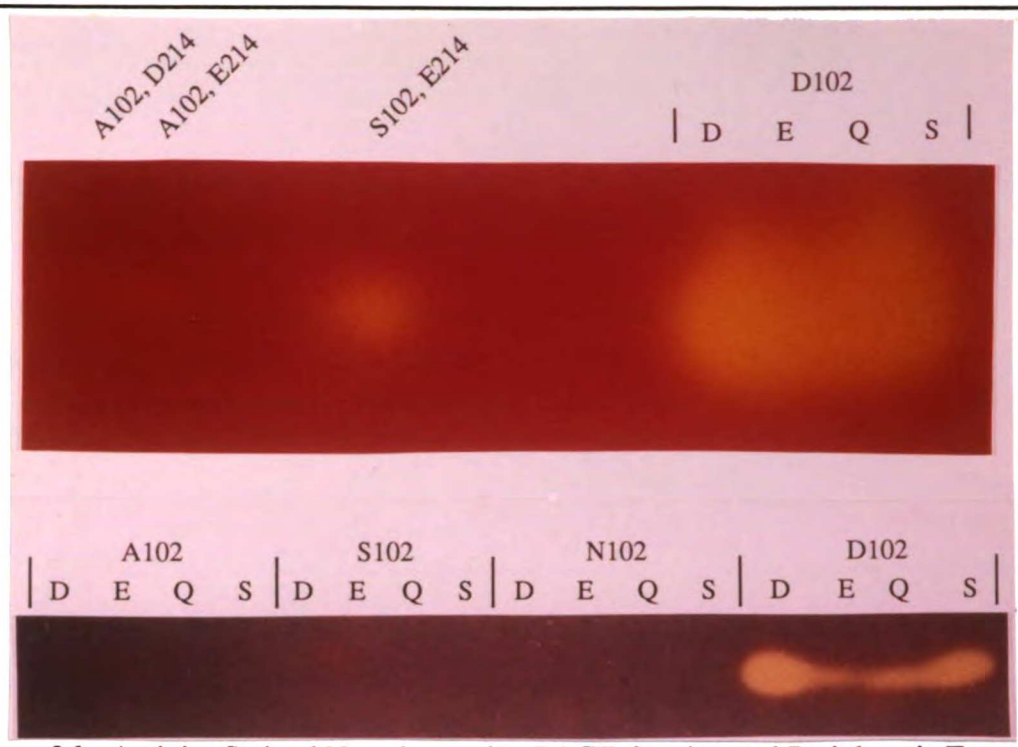


Figure 36. Activity Stained Non-denaturing PAGE-fractionated Periplasmic Extracts of *E. coli* Strain X90 Transformed with $p^mT_cD102(D,A,S,N):S214(S,D,E,Q)$ and Induced with 1mM IPTG. ***Lower*** 2 hours at 37°C. ***Upper*** Plus ~12 hours at room temperature

ENZYMOLOGICAL CHARACTERIZATION OF PURIFIED PROTEIN

The Michaelis-Menten paramters k_{cat}, K_m, and k_{cat}/K_m are listed in Table 9. Normalized values are listed in Table 10. The D102A mutant was tested using only the Z-GPR-S-Bzl and Z-GPR-AMC substrates due to the limited availability of protein. Previous results suggest that product dissociation is the rate limiting step in the hydrolysis of Z-GPR-S-Bzl by rat anionic trypsin.[229] Therefore, comparisons of the rates determined for Z-GPR-S-Bzl between mutants and trypsin are generally not valid, and comparisons between mutants must be made with care. The K_m values for Suc-AFK-AMC and MeOSuc-AFK-AMC were too high to be determined for the D102(A,S,N):S214E mutants. The K_m for Suc-AFK-AMC was too high to be measured for the D102S:S214D trypsin. The MeOSuc-AFK-AMC substrate was turned over at a rate too low for determination of k_{cat}/K_m for D102N:S214E trypsin. Overall, each of the mutants caused an increase of the K_m for the lysine amide substrates. Control reactions using a new

preparation of rat anionic trypsin gave results which agree with those previously obtained with these substrates.

Kinetic Parameters for Single Mutants D102(A,S,N)

As has been reported,[41,49] the D102(A,S,N) mutations in general caused an approximately 10^{-4} decrease in k_{cat} with the D102N mutation being slightly less disruptive than the D102(A,S) mutations. The D102(A,S,N) mutations caused an increase in K_m which was less than 10x above trypsin with a

Table 9. *Michaelis-Menten Steady State Parameters for the Hydrolysis of Various Substrates by D102X:S214Y Mutants.*

Enzyme	Substrate	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($min^{-1}\mu M^{-1}$)
WT	Suc-AFK-AMC	490 ± 20	190 ± 10	2.6 ± 0.06
	MeOSuc-AFK-AMC	440 ± 0.04	58 ± 0.8	7.7 ± 0.1
	Z-GPR-S-Bzl	5,000 ± 400	12 ± 0.6	430 ± 20
	Z-GPR-AMC	5,300 ± 100	20 ± 0.6	270 ± 10
	Z-R-S-Bzl	5,300 ± 100	3.6 ± 0.2	1,500 ± 60
	Z-R-AMC	39 ± 0.6	62 ± 10	0.63 ± 0.1
D102A	Z-GPR-S-Bzl	13 ± 0.4	3.5 ± 0.3	3.7 ± 0.1
	Z-GPR-AMC	2.3 ± 0.2	130 ± 20	0.017 ± 0.001
D102S	Suc-AFK-AMC	0.22 ± 0.03	2,500 ± 200	0.000087± 0.000005
	MeOSuc-AFK-AMC	0.026 ± 0.007	180 ± 100	0.00016 ± 0.00005
	Z-GPR-S-Bzl	5.4 ± 0.5	8.1 ± 2	0.68 ± 0.1
	Z-GPR-AMC	2.0 ± 0.2	41 ± 1	0.049 ± 0.003
	Z-R-S-Bzl	14 ± 4	30 ± 20	0.55 ± .2
	Z-R-AMC	0.033 ± 0.005	370 ± 100	0.000092 ± 0.00001
D102N	Suc-AFK-AMC	2.1 ± 0.4	2,800 ± 500	0.00075 ± 0.00001
	MeOSuc-AFK-AMC	0.62 ± 0.06	470 ± 50	0.0013 ± 0.00001
	Z-GPR-S-Bzl	2.7 ± 0.05	1.5 ± 0.1	1.9 ± 0.1
	Z-GPR-AMC	1.7 ± 0.01	5.4 ± 0.02	0.32 ± 0.0007
	Z-R-S-Bzl	7.5 ± 0.5	8.7 ± 0.3	0.86 ± 0.03
	Z-R-AMC	0.38 ± 0.01	530 ± 10	0.00073 ± 0.00001
D102S S214D	Suc-AFK-AMC	-	-	0.0019 ± 0.00005
	MeOSuc-AFK-AMC	2.6 ± 0.7	780 ± 300	0.0034 ± 0.0004
	Z-GPR-S-Bzl	30 ± 0.4	3.7 ± 0.6	8.2 ± 1
	Z-GPR-AMC	5.2 ± 0.4	18 ± 2	0.28 ± 0.006
	Z-R-S-Bzl	35 ± 0.3	4.1 ± 0.2	8.4 ± 0.3
	Z-R-AMC	0.26 ± 0.003	180 ± 1	0.0015 ± 0.00003
D102A S214E	Suc-AFK-AMC	-	-	0.0017
	MeOSuc-AFK-AMC	-	-	0.0019
	Z-GPR-S-Bzl	1,100 ± 80	59 ± 2	18 ± 2
	Z-GPR-AMC	190 ± 4	1,000 ± 30	0.19 ± 0.0007
	Z-R-S-Bzl	510 ± 40	20 ± 7	26 ± 7
	Z-R-AMC	1.1 ± 0.3	620 ± 300	0.0018 ± 0.0003
D102S S214E	Suc-AFK-AMC	-	-	0.0018
	MeOSuc-AFK-AMC	-	-	0.0022
	Z-GPR-S-Bzl	1,800 ± 90	73 ± 4	24 ± 0.2
	Z-GPR-AMC	1,100 ± 90	2,800 ± 200	0.39 ± 0.004
	Z-R-S-Bzl	740 ± 200	24 ± 8	31 ± 3
	Z-R-AMC	0.98 ± 0.2	1,000 ± 200	0.00099 ± 0.00006
D102N S214E	Suc-AFK-AMC	-	-	0.00036
	MeOSuc-AFK-AMC	-	-	-
	Z-GPR-S-Bzl	410 ± 20	81 ± 6	5.0 ± 0.2
	Z-GPR-AMC	110 ± 8	1,900 ± 200	0.057 ± 0.001
	Z-R-S-Bzl	370 ± 60	54 ± 20	7.1 ± 1
	Z-R-AMC	0.18 ± 0.04	1,800 ± 300	0.000098 ± 0.000004

slightly greater increase in the K_m values for the single residue substrate over the tripeptide substrates. Overall the catalytic efficiency (k_{cat}/K_m) of the D102S trypsin is reduced to 10^{-4} to 10^{-5} of trypsin while the catalytic efficiency of D102N trypsin is reduced to 10^{-3} to 10^{-4} of trypsin (Table 10) as has been reported.[49]

Table 10. *Normalized Michaelis-Menten Steady State Parameters for the Hydrolysis of Various Substrates by D102X:S214Y Mutants.*

Enzyme	Substrate	k_{cat}	K_m	k_{cat}/K_m
WT	Suc-AFK-AMC	1.0	1.0	1.0
	MeOSuc-AFK-AMC	1.0	1.0	1.0
	Z-GPR-S-Bzl	1.0	1.0	1.0
	Z-GPR-AMC	1.0	1.0	1.0
	Z-R-S-Bzl	1.0	1.0	1.0
	Z-R-AMC	1.0	1.0	1.0
A102	Z-GPR-S-Bzl	2.6×10^{-3}	0.29	8.6×10^{-3}
	Z-GPR-AMC	4.3×10^{-4}	6.5	6.3×10^{-5}
S102	Suc-AFK-AMC	4.5×10^{-4}	13	3.3×10^{-5}
	MeOSuc-AFK-AMC	5.9×10^{-5}	3.1	2.1×10^{-5}
	Z-GPR-S-Bzl	1.1×10^{-3}	0.68	1.6×10^{-3}
	Z-GPR-AMC	3.8×10^{-4}	2.1	1.8×10^{-4}
	Z-R-S-Bzl	2.6×10^{-3}	8.3	3.7×10^{-4}
	Z-R-AMC	8.5×10^{-4}	6.0	1.5×10^{-4}
N102	Suc-AFK-AMC	4.3×10^{-3}	15	2.9×10^{-4}
	MeOSuc-AFK-AMC	1.4×10^{-3}	8.1	1.7×10^{-4}
	Z-GPR-S-Bzl	5.4×10^{-4}	0.13	4.4×10^{-3}
	Z-GPR-AMC	3.2×10^{-4}	0.27	1.2×10^{-3}
	Z-R-S-Bzl	1.4×10^{-3}	2.4	5.7×10^{-4}
	Z-R-AMC	9.7×10^{-3}	8.5	1.2×10^{-3}
S102D214	Suc-AFK-AMC	-	-	7.3×10^{-4}
	MeOSuc-AFK-AMC	5.9×10^{-3}	13	4.4×10^{-4}
	Z-GPR-S-Bzl	6.0×10^{-3}	0.31	1.9×10^{-2}
	Z-GPR-AMC	9.8×10^{-4}	0.90	1.0×10^{-3}
	Z-R-S-Bzl	6.6×10^{-3}	1.1	5.6×10^{-3}
	Z-R-AMC	6.7×10^{-3}	2.9	2.4×10^{-3}
A102E214	Suc-AFK-AMC	-	-	6.5×10^{-4}
	MeOSuc-AFK-AMC	-	-	2.5×10^{-4}
	Suc-AAPF-S-Bzl	-	-	8.6×10^{-4}
	Z-GPR-S-Bzl	2.2×10^{-1}	4.9	4.2×10^{-2}
	Z-GPR-AMC	3.6×10^{-2}	50	7.0×10^{-4}
	Z-R-S-Bzl	9.6×10^{-2}	5.6	1.7×10^{-2}
	Z-R-AMC	2.8×10^{-2}	10	2.9×10^{-3}
S102E214	Suc-AFK-AMC	-	-	6.9×10^{-4}
	MeOSuc-AFK-AMC	-	-	2.9×10^{-4}
	Z-GPR-S-Bzl	3.6×10^{-1}	6.1	5.6×10^{-2}
	Z-GPR-AMC	2.1×10^{-1}	140	1.4×10^{-3}
	Z-R-S-Bzl	1.4×10^{-1}	6.7	2.1×10^{-2}
	Z-R-AMC	2.5×10^{-2}	16	1.6×10^{-3}
N102E214	Suc-AFK-AMC	-	-	1.4×10^{-4}
	MeOSuc-AFK-AMC	-	-	-
	Z-GPR-S-Bzl	8.2×10^{-2}	6.8	1.2×10^{-2}
	Z-GPR-AMC	2.1×10^{-2}	95	2.1×10^{-4}
	Z-R-S-Bzl	7.0×10^{-2}	15	4.7×10^{-3}
	Z-R-AMC	4.6×10^{-3}	29	1.6×10^{-4}

Kinetic Parameters for Double Mutants

D102S:S214D

As noted in Chapter 3, the S214D mutation causes approximately a 40% reduction of k_{cat} relative to trypsin and an increase of K_m up to 16x above wild type. The resulting k_{cat}/K_m values are 10^{-1} to 10^{-2} of trypsin (Table 6). The D102S:S214D mutations cause an almost uniform 7×10^{-3} decrease in k_{cat} regardless of the chemistry of the substrate. The K_m

values for these substrates vary from 0.3x to 7.1x relative to wild type. The resulting catalytic efficiencies are approximately 10^{-3} of trypsin (Table 10).

D102(A,S,N):S214E Mutants

The D102(A,S):S214E double mutations cause a decrease in k_{cat} to 10^{-1} to 10^{-2} of trypsin. The D102N:S214E double mutation causes a decrease in k_{cat} to 10^{-2} to 10^{-3} of trypsin. The D102(A,S,N):S214E mutations change K_m from 0.7x to 160x relative to wild type depending on the chemistry and particularly on the size of the substrates.

RECONSTRUCTION OF THE CATALYTIC MECHANISM OF MUTANT TRYPSINS WHICH LACK A CATALYTIC ASPARTIC ACID – DISCUSSION

BIOLOGICAL STABILITY OF THE MUTANTS

Figure 35 shows the amount of trypsin and mutant trypsin protein which can be recovered from the periplasm of *E. coli* at steady state. Proteins are recovered at a relative yield determined by the amino acid at position 102 where Asn > Asp > Ser > Ala. These results have been confirmed in a different strain of *E. coli* using the *phoA*^{227} expression system (data not shown) indicating this phenomenon is a function of the protein. The production and stability of RNA has not been ascertained, but it seems unlikely that a one or two base-pair change could determine the stability of the RNA molecules in this case. The stability of the proteins in the bacterial cells has not been determined directly by pulse-labeling experiments. The data strongly indicate such an experiment would reveal that the protein stability would parallel the operational recoveries. Figure 27 shows that the Asp-102 carboxylate forms hydrogen bonds with the backbone amide bonds of residues Ala-56 and His-57. Evidently, disrupting these hydrogen bonds makes the structure more flexible and thereby more labile to endogenous proteases.

CHEMICAL MECHANISM OF D102X TRYPSINS

Replacing the catalytic Asp-102 side chain with alanine, serine, or asparagine causes similar effects on the function of the resulting mutant trypsin. Indeed, the observed

k_{cat} values are almost identical for the substrates tested. The D102N mutation causes a smaller increase in the K_m values of amide substrates than the D102S mutation thus suggesting less disruption of substrate binding by the D102N mutation. Coupled with the disparate biological stabilities of the mutants in this series, these results suggest that the D102N mutant structure is more "rigid" than the D102A or D102S mutants.

Particularly sensitive to this putative rigidity are the kinetic parameters for the hydrolysis of the single residue arginine and tripeptide lysine amide substrates. Although the K_m values for these substrates are only slightly higher for D102N trypsin, the k_{cat} values for these substrates are an order of magnitude higher for D102N trypsin than for D102S trypsin. Evidently the structural effects of the D102S mutation allows slightly better substrate binding, but perturbs the ability of the enzyme to stabilize the transtion states in the reactions of these substrates.

The same rates are observed for the hydrolysis of ester substrates and the tripeptide Z-GPR-AMC amide substrate by D102S and D102N trypsin. The ester substrates require less chemical activation than the amide substrates, and the Z-GPR-AMC substrate apparently forms a Michaelis complex with rat anionic trypsin that activates the amide bond particularly well in an, as yet, undefined manner.[101,229] The perturbation caused by the D102S mutation may be compensated by, or irrelevant to, the function of rat anionic trypsin on Z-GPR-AMC.

The S214D single mutation clearly does not cause the same steric perturbations which result from the S214E mutation (Table 6). While some increase in K_m is detectable due to the S214D mutation, some K_m values are unaffected and others are not affected to the same extent as was detected due to the S214E mutation. The sensitivity of the lysine amide substrates to perturbation of the trypsin structure is again clear however. Although no pH-profile is presented, the Asp-214 side chain is presumably protonated due to the

stabilized electrostatic field of Asp-102. Deprotonation of Asp-214 would probably result in inactivation of the protein in a mechanism similar to that proposed for the S214E mutant.

Recovery of Structural Function in D102S:S214D Trypsin

The swap of the position 102 and 214 side chains in the D102S:S214D double-mutant trypsin apparently results in a remarkably unperturbed structure as can be discerned from the function reported by the observed K_m values. With the exception of the apparently finical lysine tripeptide amide substrates, the observed K_m values are lower than those measured for the D102S mutant. However, the k_{cat} values for all substrates are approximately 7×10^{-3} lower than for trypsin; almost a 10x improvement over the catalytic power of the D102(A,S,N) mutants and a 100x improvement for the MeOSuc-AFK-AMC substrate.

Recovery of Catalytic Function in D102(A,S,N):S214E Trypsins

The elevated K_m values observed for the D102(A,S,N):S214E mutants are reminiscent of the K_m values observed for the S214E single mutant (Table 6) which were ascribed to the displacement of the Trp-215 side chain based on X-ray crystallography results (Figure 32). However, this perturbation of the structure has not prevented the S214E mutation from restoring a significant amount of catalytic power to each of these mutants. The S214E second site mutation increases the k_{cat} of D102N:S214E trypsin by approximately 100x relative to D102N trypsin, increases the k_{cat} of D102A:S214E trypsin by 10x or 100x relative to D102A trypsin, and increases the k_{cat} of D102S:S214E trypsin by up to 500x relative to D102S trypsin. The observed increase in rate apparently depends on the inherent chemical activity of the substrate, as in the case of the esters, or on the ability of the Michaelis complex to activate the amide bond, as in the case of the Z-GPR-AMC substrate. The decreased k_{cat} for Z-R-AMC in D102N:S214E trypsin relative to D102N trypsin probably results from the disruption of the basic D102N structure by the S214E mutation. The disrupted structure of the D102N:S214E trypsin might also

interfere with functions critical to the activation of the amide bond in Z-GPR-AMC. However, the 50x increase of k_{cat} suggests that the perturbations of the D102N structure by the S214E mutation may not be relevant to the chemical activation of Z-GPR-AMC by the D102N:S214E enzyme.

RECONSTRUCTION OF THE CATALYTIC MECHANISM OF MUTANT TRYPSINS WHICH LACK A CATALYTIC ASPARTIC ACID – CONCLUSIONS

It is with some surprise that we report that the k_{cat} values for the D102A and D102S mutants are the same or lower than the k_{cat} values for the D102N mutant. The chemically reasonable orientation of the Asn-102 side chain suggests that the imidazole of His-57 has been forced into a tautomer with a pyrrolic nitrogen pointed towards the Ser-195 hydroxyl. Such an imidazole could not act as an effective general base during catalysis. We conclude that the functional effect of the D102A or D102S mutations in trypsin is to also deactivate the His-57 imidazole so that it does not function as a general base catalyst. The disruption of the trypsin structure by the D102A and D102S mutations evidently negates the proximity of His-57 by shifting the position of the imidazole relative to the Ser-195 hydroxyl so that it cannot act as a general base. The crystal structure of D102N trypsin reveals that the His-57 side chain can swing out of the active site to a low-energy conformation where it could not act as a general base.[203] The catalytic His-64 – Ser-221 couple in subtilisin retains ~10x activity over Ser-221 alone.[41] If His-64 is acting as a general base in the D32A subtilisin, it apparently is worth no more than 10x in k_{cat} in the absence of the catlytic aspartic acid.

Mutagenesis experiments on subtilisin have shown that the absence of the nucleophilic serine results in a decrease in rate which is more dramatic than the decrease detected in these experiments. Those results suggest a role for the Ser-195 hydroxyl in serine proteases in which the catalytic aspartic acid as been removed.[41] The observed difference in catalytic rates is consistent with the difference in nucleophilicity between water

and hydroxide ion[56,174] where water is the model nucleophile in the absence of a general base catalyst and hydroxide is the model nucleophile in the presence of a general base catalyst. The loss of a functional general base suggests the mechanism presented in Figure 37 for the D102(A,S,N) mutants in which the Ser-195 hydroxyl acts as the nucleophile without general base catalysis. This mechanism could be accelerated by deprotonation of the Ser-195 hydroxyl prior to substrate binding or by specific base catalysis with the resulting water molecule perhaps acting as a general acid in the decomposition of the tetrahedral intermediate. Either mechanism would be consistent with the pH-profile reported for D102N trypsin. It is not clear whether the proton transfers occur in concert with or precede the collapse of the tetrahedral intermediate, but it seems chemically unreasonable for them to follow. Comparison to the mechanism of trypsin suggests the importance of the proposed moving histidine for shuttling protons[7,125] in a general-acid catalyzed decomposition of the tetrahedral intermediate.

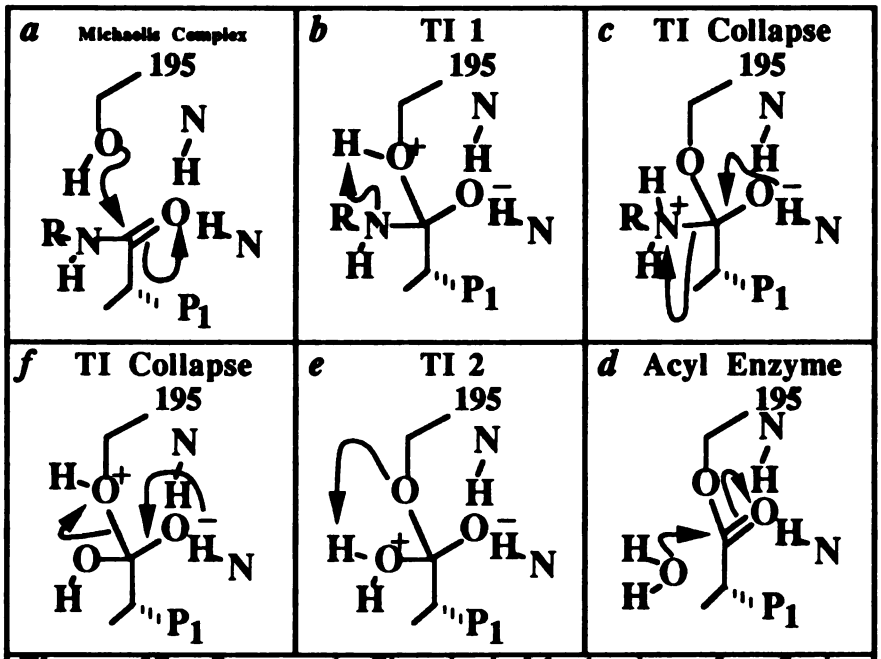


Figure 37. Proposed Chemical Mechanism for Serine Proteases Which Lack a Catalytic Aspartic Acid. Proton transfers on the reaction pathway to product are invoked from ***b*** to ***c*** and ***e*** to ***f***. Hydroxide may also attack the acyl enzyme in ***d***. TI: Tetrahedral Intermediate.

The steric perturbations caused by the D102S mutation can apparently be relieved to some degree by the S214D mutation. It is unclear what portion of the activity recovered

because of the S214D mutation is due to (1) structural corrections as seems to occur for the D102N single mutant over the D102S mutant, (2) the possible reactivation of the imidazole of His-57 as a general base, or (3) electrostatic transition state stabilization. The improvement in the structure of the D102S mutants pays off well for the hydrolysis of MeOSuc-AFK-AMC. But the low k_{cat} values overall suggest that the His-57 – Ser-195 couple has not been activated by the S214D mutation.

Significantly more of the function of the catalytic triad is restored to the D102(A,S,N) mutants by the S214E second site mutation than by the S214D second site mutation. This trend is presumably due to the ability of the Glu-214 carboxylate to approximate more closely than the Asp-214 carboxylate the position of the Asp-102 carboxylate found in trypsin. The amide of Gln-214 is not able to activate His-57 (Figure 36) indicating the importance of the ionizability of the carboxylate. The recovery of activity is most pronounced for Z-GPR-AMC. Presumably, the D102(A,S):S214E double mutants are able to activate the imidazole of His-57 while the D102S:S214D double mutant cannot. The Asn-102 side chain may deactivate the imidazole in the D102N:S214E trypsin or may be sterically hindering the activation of the His-57 imidazole by the S214E mutation. The enhancement of rate for the D102N mutant by the S214E mutation is not consistent with the mechanism presented in Figure 37. The mechanism without general base catalysis proposed above does involve the development of positive charge in the transition state and the presence of a positive charge in the tetrahedral intermediate which might be stabilized by electrostatic interaction with the Glu-214 side chain in the D102N:S214E double mutant which might explain the rate acceleration. However, in a similar experiment in a similar system, a rate enhancement is not seen in the H64A subtilisin versus the H64A:D32A double mutant in which steric perturbation or solvent screening of the negative charge might account for this result.[41]

The combination of the Ser-102 and Glu-214 side chains is apparently more effective at activating the His-57 imidazole than the combination of the Ala-102 and Glu-214 side chains. Presumably, the Glu-214 carboxylate is finding the position formerly occupied by the Asp-102 carboxylate. In this position, the Ser-102 hydroxyl stabilizes the Glu-214 carboxylate better than the Ala-102 methyl group. The Glu-214 carboxylate might also be located in the *down* orientation found in the electron density maps of the S214E trypsin. Or there could be a partitioning in which the *up* orientation is better stabilized by a hydrogen bond from the Ser-102 hydroxyl group.

The inability of the Asp-214 carboxylate to activate the His-57 imidazole suggests that the His-57 imidazole may not be in the catalytically competent position. This would suggest that the position of the carboxylate is important for orienting or even holding the His-57 imidazole in the Michaelis complex.[203] Alternatively, the His-57 side chain may be close to the catalytic position, but unactivated due to the inability of the electrostatic field of the protein to stabilize the negative charge of a carboxylate at the 214 position.

Modeling indicates that the Glu-214 carboxylate can only interact with His-57 using the *anti* lone pair electrons (Robert J. Fletterick, *personal communication*). The putative chemical activation of the Ser-195 hydroxyl group by the Glu-214:His-57 couple would indicate that it is the position of the negative charge which is important, not the orientation of carboxylate. A less-well characterized Asp – His – Ser catalytic triad has been reported in phospholipase with this orientation of the carboxylate.[242] This position of the Glu-214 carboxylate might help orient the His-57 imidazole in the ground state, but also seems capable of activating the His-57:Ser-195 couple in the transistion state. According to experiments supporting Gandour values Rule, the chemical activity of the carboxylate has been reduced 10^4 to 10^5 yet the net loss of chemical activity of the Ser-195 hydroxyl is no more than 10x. The experiments in Chapter 3 indicate that the interactions of buried charges are more energetic in the absence of dipoles which can reorient to counter the

electric field of the charges. The activation of the His-57 imidazole by a "misoriented" carboxylate would further suggest that an important feature of the catalytic triad is not the orientation of the residues so much as the fact that they are buried charges.

REFERENCES

1. Ako, H., Foster, R.J. & Ryan, C.A. "Mechanism of Action of Naturally Occuring Proteinase Inhibitors. Studies with Anhydrotrypsin and Anhydrochymotrypsin Purified by Affinity Chromatography" *Biochemistry* **13**, 132-139 (1974).
2. Ako, H., Ryan, C.A. & Foster, R.J. "The Purification by Affinity Chromatography of a Proteinase Inhibitor Binding Species of Anhydro-chymotrypsin" *Biochem. Biophys. Res. Comm.* **46**, 1639-1645 (1972).
3. Alber, T., Petsko, G.A. & Tsernglou, D. "Crystal structure of elastase-substrate complex at -55°C" *Nature* **263**, 297-300 (1976).
4. Albery, A.J. "Solvent isotope effects" in *Proton-Transfer Reactions* Caldin, E. & Gold, V., eds., (Chapman and Hall, London) pp. 263 (1975).
5. Allison, R.D. & Purich, D.L. "Practical Considerations in the Design of Initial Velocity Enzyme Rate Assays" *Meth. in Enzymol.* **63**, 3-22 (1979).
6. Amman, E., Brosius, J. & Ptashne, M. "Vectors bearing a hybrid *trp-lac* prometer useful for regulated expression of cloned genes in *Escherichia coli*" *Gene* **25**, 167-170 (1983).
7. Bachovchin, W.W. "^{15}N NMR Spectroscopy of Hydrogen Bonding Interactions in the Active Site of Serine Proteases: Evidence for a Moving Histidine Mechanism" *Biochemistry* **25**, 7751-7759 (1986).
8. Bachovchin, W.W., Wong, W.Y.L., Farr-Jones, S., Shenvi, A.B. & Ketner, C.A. "Nitrogen-15 NMR Spectroscopy of the Catalytic-Triad Histidine of a Serine Protease in Peptide Boronic Acid Inhibitor Complexes" *Biochemistry* **27**, 7689-7697 (1988).
9. Barbas, C.F., III, Matos, J.R., West, J.B. & Wong, C.-H. "A Search for Peptide Ligase: Cosolvent-Mediates Conversion of Proteases to Esterases for Irreversible Synthesis of Peptides" *J. Am. Chem. Soc.* **110**, 5162-5166 (1988).
10. Bazan, J.F. & Fletterick, R.J. "Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: Structural and functional implications" *Proc. Natl. Acad. Sci. USA* **85**, 7872-7876 (1988).
11. Beattie, R.E., Elmore, D.T., Williams, C.H. & Gutherie, D.J.S. "The behaviour of leucine aminopeptidase towards thionopeptides" *Biochem. J.* **245**, 285-288 (1987).
12. Bell, G.I., Quinto, C., Quiroga, M., Valenzuela, P., Craik, C.S. & Rutter, W.J. "Isolation and Sequence of a Rat Chymotrypsin B Gene" *J. Biol. Chem.* **259**, 14270-14265 (1984).
13. Bender, M. & Kaiser, E.T. "The Mechanism of Trypsin-catalyzed Hydrolyses. The Cinnamoyl-trypsin Intermediate" *J. Am. Chem. Soc.* **84**, 2556-2561 (1962).
14. Bender, M. & Zerner, B. "The Formation of an Acyl-enzyme Intermediate in the α-Chymotrypsin-catalyzed Hydrolyses of Non-labile *trans*-Cinnamic Acid Esters" *J. Am. Chem. Soc.* **84**, 2550-2555 (1962).
15. Bender, M.L. "Oxygen Exchange as Evidence for the Existence of an Intermediate in Ester Hydrolysis" *J. Am. Chem. Soc.* **73**, 1626-1629 (1951).
16. Bender, M.L., Clement, G.E., Gunter, C.R. & Kézdy, F.J. "The Kinetics of α-Chymotrypsin Reactions in the Presence of Added Nucleophiles" *J. Am. Chem. Soc.* **86**, 3697-3703 (1964).
17. Bender, M.L. & Ginger, R.D. "Intermediates in the Reactions of Carboxylic Acid Derivatives. IV. The Hydrolysis of Benzamide" *J. Am. Chem. Soc.* **77**, 348-351 (1957).
18. Bender, M.L. & Nakamura, K. "The Effect of Structure on the Rates of Some α-Chymotrypsin-catalyzed Reactions" *J. Am. Chem. Soc.* **84**, 2577-2582 (1962).
19. Bender, M.L., Schonbaum, G.R. & Zerner, B. "Spectrophotometric Investigations of the Mechanism of α-Chymotrypsin-catalyzed Hydrolyses. Detection of the Acyl-enzyme Intermediate" *J. Am. Chem. Soc.* **84**, 2540-2550 (1962).
20. Berg, C.M. & Berg, D.E. "Uses of Transposable Elements and Maps of Known Insertions" in *Eschericia coli and Salmonella typhimurium Cellular and Molecular Biology* Neidhardt, F.C., *et al.*, eds., (American Society for Microbiology, Washington, D. C.) pp. 1071-1109 (1987).

21. Bergmann, M. & Fruton, J.S. "On Proteolytic Enzymes XIII. Synthetic Substrates for Chymotrypsin" *J. Biol. Chem.* **118**, 405-415 (1937).

22. Bergmann, M., Fruton, J.S. & Pollok, H. "The Specificity of Trypsin" *J. Biol. Chem.* **127**, 643-648 (1939).

23. Bergmann, M. & Ross, W.F. "Synthetic Substrates for Protein-digesting Enzymes" *J. Am. Chem. Soc* **58**, 1503 (1936).

24. Bernhard, S.A. & Gutfreund, H. "The Optical Detection of Transients in Trypsin- and Chymotrypsin-catalyzed Reactions" *Proc. Natl. Acad. Sci. USA* **53**, 1238-1243 (1965).

25. Bernhard, S.A., Lau, S.J. & Noller, H. "Spectrophotometric Identification of Acyl Enzyme Intermediates" *Biochemistry* **4**, 1108-1118 (1965).

26. Birnboim, H.C. & Doly, J. "A rapid alkaline extraction procedure for screening recombinant plasmid DNA" *Nucleic Acids Res.* **7**, 1513 (1979).

27. Blow, D.M., Birktoft, J.J. & Hartley, B.S. "Role of a Buried Acid Group in the Mechanism of Action of Chymotrypsin" *Nature* **221**, 337-340 (1969).

28. Bode, W. & Huber, R. "Crystal structures of pancreatic serine endopeptidases" in *Molecular and Cellular Basis of Digestion* Desnuelle, P., Sjoestroem, H. & Noren, O., eds., (Elsevier, New York) pp. 213-234 (1986).

29. Bode, W. & Schwager, P. "The Refined Crystal Structure of Bovine β-Trypsin at 1.8Å Resolution" *J. Mol. Biol.* **98**, 693-717 (1975).

30. Bode, W., Walter, J., Huber, R., Wenzel, H.R. & Tschesche, H. "The refined 2.2-Å (0.22nm) X-ray crystal structure of the ternary complex formed by bovine trypsinogen, valine-valine and the Arg^{15} analogue of bovine pancreatic trypsin inhibitor" *Eur. J. Biochem.* **144**, 185-190 (1984).

31. Bone, R., Shenvi, A.B., Kettner, C.A. & Agard, D.A. "Serine Protease Mechanism: Structure of an Inhibitory Complex of α-Lytic Protease and a Tightly Bound Peptide Boronic Acid" *Biochemistry* **26**, 7609-7614 (1987).

32. Bone, R., Silen, J.L. & Agard, D.A. "Structural plasticity broadens the specificity of an engineered protease" *Nature* **339**, 191-195 (1989).

33. Brady, K., Wei, A., Ringe, D. & Abeles, R.H. "Structure of Chymotrypsin Trifluoromethyl Ketone Inhibitor Complexes: Comparison of Slowly and Rapidly Equilibrating Inhibitors" *Biochemistry* **29**, 7600-7607 (1990).

34. Brady, L., Brzozowski, A.M., Derewenda, Z.S., Dodson, E., Dodson, G., Tolley, S., Turkenburg, J.P., Christiansen, L., Huge-Jensen, B., Norskov, L., Thim, L. & Menge, U. "A serine protease triad forms the catalytic centre of a triacylglycerol lipase" *Nature* **343**, 767-770 (1990).

35. Brayer, G.D., Delbaere, L.T.J. & James, M.N.G. "Molecular Structure of Crystalline *Streptomyces griseus* Protease A at 2.8Å Resolution II. Molecular Conformation, Comparison with α-Chymotrypsin and Active-site Geometry" *J. Mol. Biol.* **124**, 261-283 (1978).

36. Briggs, G.E. & Haldane, J.B.S. "A Note on the Kinetics of Enzyme Action" *Biochem. J.* **19**, 338-339 (1925).

37. Bryan, P., Pantoliano, M.W., Quill, S.G., Hsiao, H.-Y. & Poulos, T. "Site-directed mutagenesis and the role of the oxyanion hole in subtilisin" *Proc. Natl. Acad. Sci. U. S. A.* **83**, 3743 (1986).

38. Calos, M.P. "DNA sequence for a low-level promoter of the *lac* repressor gene and an 'up' promoter mutation" *Nature* **274**, 762-765 (1978).

39. Caplow, M. & Jencks, W.P. "The Effect of Substituents on the Deacylation of Benzoyl-Chymotrypsins" *Biochemistry* **1**, 883-893 (1962).

40. Carter, P. & Wells, J.A. "Engineering Enzyme Specificity by 'Substrate-Assisted Catalysis'" *Science* **237**, 394-399 (1987).

41. Carter, P. & Wells, J.A. "Dissecting the catalytic triad of a serine protease" *Nature* **332**, 564-568 (1988).

42. Carter, P. & Wells, J.A. "Functional Interaction Among Catalytic Residues in Subtilisin BPN'" *Proteins* **7**, 335-342 (1990).

43. Castillo, M.J., Nakajima, K., Zimmerman, M. & Powers, J.C. "Sensitive Substrates for Human Leukocyte and Porcine Pancreatic Elastase: A Study of the Merits of Various Chromophoric and Fluorogenic Leaving Groups in Assays for Serine Proteases" *Anal. Biochem.* **99**, 53-64 (1979).

44. Chung, C.H., Ives, H.E., Almeda, S. & Goldberg, A.L. "Purification from *Escherichia coli* of a Periplamic Protein That Is a Potent Inhibitor of Pancreatic Proteases" *J. Biol. Chem.* **258**, 11032-11038 (1983).

45. Cornish-Bowden, A. & Wharton, C.W. *Enzyme Kinetics* (IRL Press, Washington, D.C.) pages (1988).

46. Craik, C.S. "Use of Oligonucleotides for Site-Specific Mutagenesis" *BioTechniques* **Jan/Feb**, 12-19 (1985).

47. Craik, C.S., Choo, Q.-L., Swift, G.H., Quinto, C., MacDonald, R.J. & Rutter, W.J. "Structure of Two Related Pancreatic Trypsin Genes" *J. Biol. Chem.* **259**, 14255-14264 (1984).

48. Craik, C.S., Largman, C., Fletcher, T., Roczniak, S., Barr, P.J., Fletterick, R. & Rutter, W.J. "Redesigning Trypsin: Alteration of Substrate Specificity" *Science* **228**, 291-297 (1985).

49. Craik, C.S., Roczniak, S., Largman, C. & Rutter, W.J. "The Catalytic Role of the Active Site Aspartic Acid in Serine Proteases" *Science* **237**, 909-913 (1987).

50. Debye, P. *Polar Molecules* (Dover Publishing Co., New York) 172 pages (1929).

51. Delbaere, L.T.J., Hutcheon, W.L.B., James, M.N.G. & Thiessen, W.E. "Tertiary structural differences between microbial serine proteases and pancreatic serine enzymes" *Nature* **257**, 758-763 (1975).

52. Dente, L., Cesareni, G. & Cortese, R. "pEMBL: A New Family of Single-Stranded Plasmids" *Nuc. Acids Res.* **11**, 1645-1655 (1983).

53. Dixon, G.H., Go, S. & Neurath, H. "Peptides combined with ^{14}C-di*iso*propyl phosphoryl following following degradation of ^{14}C-DIP-trypsin with α-chymotrypsin" *Biochim. Biophys. Acta* **19**, 193-195 (1956).

54. Dorovska-Taran, V., Momtcheva, R., Gulubova, N. & Martinek, K. "The Specificity in the Elementary Steps of α-Chymotrypsin Catalysis" *Biochim. Biophys. Acta* **702**, 37-53 (1982).

55. Dotto, G.P., Enea, V. & Zinder, N.D. "Functional Analysis of the Bacteriophage f1 Intergeneic Region" *Virology* **114**, 463-473 (1981).

56. Edwards, J.O. & Pearson, R.G. "The Factors Determining Nucleophilic Reactivities" *J. Am. Chem. Soc.* **84**, 16-24 (1962).

57. Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water* (Oxford University Press, New York) 296 pages (1969).

58. Elrod, J.P., Hogg, J.L., Quinn, D.M. & Schowen, R.L. "Protonic reorganization and substrate structure in catalysis by serine proteases" *J. Am. Chem. Soc.* **102**, 3917 (1980).

59. Epand, R.M. & Wilson, I.B. "Evidence for the Formation of Hippuryl Chymotrypsin during the Hydrolysis of Hippuric Acid Esters" *J. Biol. Chem.* **238**, 1718-1723 (1963).

60. Epand, R.M. & Wilson, I.B. "Comparison of the Relative Rates of the Chymotrypsin-catalyzed and Nonenzymic Reactions of Esters with Hydroxylamine and Water" *J. Biol. Chem.* **240**, 1104-1107 (1965).

61. Evnin, L.B. *Investigation of the Substrate Specificity of Trypsin* (University of California, San Francisco, 1990).

62. Evnin, L.B. & Craik, C.S. "Development of an Efficient Method for Generating and Screening Active Trypsin and Trypsin Variants" *Annals of the New York Academy of Sciences* **542**, 61-74 (1988).

63. Evnin, L.B., Vásquez, J.R. & Craik, C.S. "Substrate specificity of trypsin investigated by using a genetic selection" *Proc. Natl. Acad. Sci. USA* **87**, 6659-6663 (1990).

64. Fahrney, D.E. & Gold, A.M. "Sulfonyl Fluorides as Inhibitors of Esterases. I. Rates of Reaction with Acetylcholinesterase, α-Chymotrypsin, and Trypsin" *J. Am. Chem. Soc.* **85**, 997-1000 (1963).

65. Farmer, D.A. & Hageman, J.H. "Use of *N*-Benzoyl-L-tyrosine Thiobenzyl Ester as a Protease Substrate: Hydrolysis by α-Chymotrypsin and Subtilisin BPN'" *J. Biol. Chem.* **250**, 7366-7371 (1975).

66. Farr-Jones, S., Smith, S.O., Kettner, C.A., Griffin, R.G. & Bachovchin, W.W. "Crystal versus solution structure of enzymes: NMR spectroscopy of a peptide boronic acid-serine protease complex in the crystalline state" *Proc. Natl. Acad. Sci. USA* **86**, 6922-6924 (1989).

67. Fastrez, J. & Fersht, A.R. "Demonstration of the Acyl-Enzyme Mechanism for the Hydrolysis of Peptides and Anilides by Chymotrypsin" *Biochemistry* **12**, 2025-2034 (1973).

68. Fersht, A.R. *Enzyme Structure and Mechanism, 2nd Edition* (W. H. Freeman and Company, San Francisco) 475 pages (1985).

69. Fersht, A.R. & Sternberg, M.J.E. "Can a simple function for the dielectric response model electrostatic effects in globular proteins?" *Prot. Eng.* **2**, 527-530 (1989).

70. Fletcher, T.S., Alhadeff, M., Craik, C.S. & Largman, C. "Isolation and Characterization of a cDNA Encoding Rat Cationic Trypsinogen" *Biochemistry* **26**, 3081-3086 (1987).

71. Fruton, J.S. "Proteinase-catalyzed Synthesis of Peptide Bonds" *Adv. Enzymol.* **53**, 239-306 (1982).

72. Fujinaga, M., Delbaere, L.T.J., Brayer, G.D. & James, M.N.G. "Refined Structure of α-Lytic Protease at 1.7Å Resolution: Analysis of Hydrogen Bonding and Solvent Structure" *J. Mol. Biol.* **183**, 479-502 (1985).

73. Gandour, R.D. "On the Importance of Orientation in General Base Catalysis by Carboxylate" *Bioorganic Chemistry* **10**, 169-176 (1981).

74. Garg, S.K. & Smyth, C.P. "Microwave Absorption and Molecular Structure in Liquids. LXVI. The Dielectric Relaxation of the Water-Dioxane System and the Structure of Water" *J. Chem. Phys.* **43**, 2959-2965 (1965).

75. Gelb, M.H. & Abeles, R.H. "Substituted Isatoic Anhydrides: Selective Inactivators of Trypsin-like Serine Proteases" *J. Med. Chem.* **29**, 585-589 (1986).

76. Gibson, Q.H. "Apparatus for the study of rapid reactions" *J. Physiol.* **117**, 49P-50P (1952).

77. Gibson, Q.H. & Roughton, F.J.W. "The kinetics of dissociation of the first oxygen molecule from fully saturated oxyhaemoglobin in sheep blood solutions" *Proc. Roy. Soc. (London) Series B* **143**, 310-334 (1955).

78. Gilson, M.K. & Honig, B. "The Dielectric Constant of a Folded Protein" *Biopolymers* **25**, 2097-2119 (1986).

79. Gilson, M.K. & Honig, B. "Calculation of the Total Electrostatic Energy of a Macromolecular System: Solvation Energies, Binding Energies, and Conformational Analysis" *Proteins* **4**, 7-18 (1988).

80. Gilson, M.K. & Honig, B.H. "Calculation of electrostatic potentials in an enzyme active site" *Nature* **330**, 84-86 (1987).

81. Gilson, M.K. & Honig, B.H. "Energetics of Charge-Charge Interactions in Proteins" *Proteins* **3**, 32-52 (1988).

82. Gilson, M.K., Sharp, K.A. & Honig, B.H. "Calculating the Electrostatic Potential of Molecules in Solution: Method and Error Assessment" *J. Comp. Chem.* **9**, 327-335 (1987).

83. Glusker, J.P. & Trueblood, K.N. *Crystal Structure Analysis: A Primer, Second Edition* (Oxford University Press, New York) 269 pages (1985).

84. Graf, L., Craik, C.S., Patthy, A., Roczniak, S., Fletterick, R.J. & Rutter, W.J. "Selective Alteration of Substrate Specificity by Relacement of Aspartic Acid-189 with Lysine in the Binding Pocket of Trypsin" *Biochemistry* **26**, 2616-2623 (1987).

85. Gutfreund, H. & Hammond, B.R. "Steps in the Reactions of Chymotrypsin with Tyrosine Derivatives" *Biochem. J.* **73**, 526-530 (1959).

86. Gutfreund, H. & Sturtevant, J.M. "The Mechanism of Chymotrypsin-catalyzed Reactions" *Proc. Nat. Acad. Sci. USA* **42**, 719-728 (1956).

87. Gutfreund, H. & Sturtevant, J.M. "The Mechanism of the Reaction of Chymotrypsin with *p*-Nitrophenyl Acetate" *Biochem. J.* **63**, 656-661 (1956).

88. Hadorn, B., Tarlow, M.J., Lloyd, J.K. & Wolff, O.H. "Intestinal Enterokinase Deficiency" *Lancet* **1**, 812-813 (1969).

89. Hahn, K.W., Klis, W.A. & Stewart, J.M. "Design and Synthesis of a Peptide Having Chymotrypsin-Like Esterase Activity" *Science* **248**, 1544-1547 (1990).

90. Hanahan, D. "Studies on the Transformation of *Escherichia coli* with Plasmids" *J. Mol. Biol.* **166**, 557-580 (1983).

91. Hanahan, D. "Techniques for Transformation of *E. coli*" in *DNA Cloning, a practical approach* Glover, D.M., eds., pp. (1985).

92. Harper, J.W., Ramirez, G. & Powers, J.C. "Reaction of Peptide Thiobenzyl Esters with Mammalian Chymotrypsinlike Enzymes: A Sensitive Assay Method" *Anal. Biochem.* **118**, 382-387 (1981).

93. Hartley, B.S., Brown, J.R., Kauffman, D.L. & Smillie, L.B. "Evolutionary Similarities Between Pancreatic Proteolytic Enzymes" *Nature* **207**, 1157-1159 (1965).

94. Hartley, B.S. & Kilby, B.A. "The Reaction of *p*-Nitrophenyl Esters with Chymotrypsin and Insulin" *Biochem. J.* **56**, 288-297 (1954).

95. Hartley, B.S. & Massey, V. "The Active Centre of Chymotrypsin I. Labelling with a Fluorescent Dye" *Biochim. et Biophys. Acta* **21**, 58-70 (1956).

96. Harvey, S.C. "Treatment of Electrostatic Effects in Macromolecular Modeling" *Proteins* **5**, 78-92 (1989).

97. Hasted, J.B. "Liquid Water: Dielectric Properties" in *The Physics and Physical Chemistry of Water* Franks, F., eds., (Plenum Press, New York) pp. 255-309 (1972).

98. Hedstrom, L., Moorman, A.R., Dobbs, J. & Abeles, R.H. "Suicide Inactivators of Chymotrypsin by Benzoxazinones" *Biochemistry* **23**, 1753-1759 (1984).

99. Henderson, R. "Structure of Crystalline α-Chymotrypsin IV. The Structure of Indoleacryloyl-α-Chymotrypsin and its Relevance to the Hydrolytic Mechanism of the Enzyme" *J. Mol. Biol.* **54**, 341-354 (1970).

100. Hendrickson, W.A. & Konnert, J.H. "Incorporation of Stereochemical Information into Crystallographic Refinement" eds., (Indian Academy of Sciences, Bangalore) pp. 1-13.01 (1980).

101. Higaki, J.N., Evnin, L.B. & Craik, C.S. "Introduction of a Cysteine Protease Active Site into Trypsin" *Biochemistry* **28**, 9256-9263 (1989).

102. Higaki, J.N., Haymore, B.L., Chen, S., Fletterick, R.J. & Craik, C.S. "Regulation of Serine Protease Activity by an Engineered Metal Switch" *Biochemistry* **29**, 8582-8586 (1990).

103. Hofmann, K. & Bergmann, M. "The Specificity of Trypsin. II" *J. Biol. Chem.* **130**, 81-86 (1939).

104. Huber, R., Bode, W., Kukla, D., Kohl, U. & Ryan, C.A. "The Structure of the Complex Formed by Bovine Trypsin and Bovine Pancreatic Trypsin Inhibitor III. Structure of the Anhydro-Trypsin-Inhibitor Complex" *Biophys. Struct. Mechanism* **1**, 189-201 (1975).

105. Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J. & Steigemann, W. "Structure of the Complex formed by Bovine Trypsin and Bovine Pancreatic Trypsin Inhibitor II. Crystallographic Refinement at 1.9Å Resolution" *J. Mol. Biol.* **89**, 73-101 (1974).

106. Huff, J.B., Askew, B., Duff, R.J. & Julius Rebek, J. "Stereoelectronic Effects and the Active Site of the Serine Proteases" *J. Am. Chem. Soc.* **110**, 5908-5909 (1988).

107. James, M.N.G., Delbaere, L.T.J. & Brayer, G.D. "Amino acid sequence alignment of bacterial and mammalian pancreatic serine proteases based on topological equivalences" *Can. J. Biochem.* **56**, 396-402 (1978).

108. James, M.N.G., Sielecki, A.R., Brayer, A.R., Delbaere, L.T.J. & Bauer, C.-A. "Structures of Product and Inhibitor Complexes of *Streptomyces griseus* Protease A at 1.8Å Resolution: A Model for Serine Protease Catalysis" *J. Mol. Biol.* **144**, 43-88 (1980).

109. Jameson, G.W., Roberts, D.V., Adams, R.W., Kyle, W.S.A. & Elmore, D.T. "Determination of the Operational Molarity of Solutions of Bovine α-Chymotrypsin, Trypsin, Thrombin, and Factor Xa by Spectrofluorimetric Titration" *Biochem. J.* **131**, 107-117 (1973).

110. Jansen, E.F., Nutting, M.-D. & Balls, A.K. "Mode of Inhibition of Chymotrypsin by Diisopropyl Fluorphosphate I. Introduction of Phosphorus" *J. Biol. Chem.* **179**, 201-204 (1949).

111. Jencks, W.P. *Catalysis in Chemistry and Enzymology* (McGraw-Hill Book Company, San Francisco) 644 pages (1969).

112. Jencks, W.P. "Binding Energy, Specificity, and Enzymic Catalysis: The Circe Effect" in *Advances in Enzymology and Other Related Areas of Molecular Biology* Meister, A., eds., (Wiley, New York) pp. 219-410 (1975).

113. Jencks, W.P. "Economics of Enzyme Catalysis" *Cold Spring Harbor Symposia on Quantitative Biology* **52**, 65-73 (1987).

114. Joyce, C.M. & Grindley, N.D.F. "Method for Determining Whether a Gene of *Escherichia coli* Is Essential: Application to the *polA* Gene" *J. Bacteriol.* **158**, 636-643 (1984).

115. Kahne, D. & Still, W.C. "Hydrolysis of a Peptide Bond in Neutral Water" *J. Am. Chem. Soc.* **110**, 7529-7534 (1988).

116. Kallos, J. & Rizok, D. "The Study of the Active Center of Chymotrypsin. II. *p*-Toluenesylfonyl-Chymotrypsin" *J. Mol. Biol.* **7**, 599-601 (1963).

117. Kallos, J. & Rizok, D. "Heavy Atom Labelling of the Serine of the Active Centre of Chymotrypsin: Pipsyl-chymotrypsin" *J. Mol. Biol.* **9**, 255-259 (1964).

118. Kaufman, S., Neurath, H. & Schwert, G.W. "The Specific Peptidase and Esterase Activities of Chymotrypsin" *J. Biol. Chem.* **177**, 793-814 (1949).

119. Kézdy, F.J., Clement, G.E. & Bender, M.L. "The Observation of Acyl-Enzyme Intermediates in the α-Chymotrypsin-Catalyzed Reactions of N-Acetyl-L-tryptophan Derivatives at Low pH" *J. Am. Chem. Soc.* **86**, 3690-3696 (1964).

120. Klapper, I., Hagstrom, R., Fine, R., Sharp, K. & Honig, B. "Focusing of Electric Fields in the Active Site of Cu-Zn Superoxide Dismutase: Effects of Ionic Strength and Amino-Acid Modification" *Proteins* **1**, 47-59 (1986).

121. Knowles, J. "Intermolecular Forces and Enzyme Specificity" in *Peptides 1968: Proceedings of the Ninth European Peptide Symposium in Orsay France* Bricas, E., eds., (North-Holland Publishing Corp., Amsterdam) pp. 310-323 (1968).

122. Kollman, P.A. & Hayes, D.M. "Theoretical Calculations on Proton-Transfer Energetics: Studies of Methanol, Imidazole, Formic Acid, and Methanethiol as Models for the Serine and Cysteine Proteases" *J. Am. Chem. Soc.* **103**, 2955-2961 (1981).

123. Kornberg, A. *DNA replication* (W. H. Freeman, San Francisco) 724 pages (1980).

124. Koshland, D.E., Strumeyer, D.H. & Ray, W.J., Jr. "Amino Acids Involved in the Action of Chymotrypsin" *Brookhaven Symp. Biol.* **15**, 101-133 (1962).

125. Kossiakoff, A.A. "Catalytic Properties of Trypsin" in *Active Sites of Enzymes* Jurnak, F.A. & McPherson, A., eds., (John Wiley & Sons, New York) pp. 367-412 (1987).

126. Kossiakoff, A.A. & Spencer, S.A. "Neutron diffraction identifies His 57 as the catalytic base in trypsin" *Nature* **288**, 414-416 (1980).

127. Kraut, J. "How Do Enzymes Work?" *Science* **242**, 533-540 (1988).

128. Kresge, A.J. "Solvent isotope effects and the mechanism of trypsin action" *J. Am. Chem. Soc.* **95**, 3065 (1973).

129. Krieger, M., Kay, L.M. & Stroud, R.M. "Structure and Specific Binding of Trypsin: Comparison of Inhibited Derivatives and a Model for Substrate Binding" *J. Mol. Biol.* **83**, 209-230 (1974).

130. Kunitz, M. "Formation of Trypsin from Crystalline Trypsinogen by Means of Enterokinase" *J. Gen. Physiol.* **22**, 429-446 (1939).

131. Kunitz, M. & Northrop, J.H. "Crystalline Chymo-trypsin and Chymo-trypsinogen I. Isolation, Crystallization, and General Properties of a New Proteolytic Enzyme and Its Precursor" *J. Gen. Physiol.* **18**, 433-458 (1935).

132. Kunitz, M. & Northrop, J.H. "Isolation from Beef Pancreas of Crystalline Trypsinogen, Trypsin, A Trypsin Inhibitor, and an Inhibitor-Trypsin Compound" *J. Gen. Physiol.* **19**, 991-1007 (1935).

133. Kunkel, T.A. "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Proc. Natl. Acad. Sci. USA* **82**, 488-492 (1985).

134. Kunkel, T.A., Roberts, J.D. & Zakour, R.A. "Rapid and Efficient Site-Specific Mutagenesis without Phenotypic Selection" *Meth. Enzymol.* **154**, 367-382 (1987).

135. Laemmli, U.K. "Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4" *Nature* **227**, 680-685 (1970).

136. Laidler, K.J. & King, M.C. "The Development of Transition-State Theory" *J. Phys. Chem.* **87**, 2657-2664 (1983).

137. Liao, D.I. & Remington, S.J. "Structure of Wheat Serine Carboxypeptidase II at 3.5-Å Resolution: A New Class of Serine Proteinase" *J. Biol. Chem.* **265**, 6528-6531 (1990).

138. Lienhard, G.E. "Enzymatic Catalysis and Transition-State Theory" *Science* **180**, 149-154 (1973).

139. Mackenzie, N.E., Malthouse, J.P.G. & Scott, A.I. "Cryoenzymology of trypsin: ^{13}C N.M.R. detection of an acyl-trypsin intermediate in the trypsin-catalyzed hydrolysis of a highly specific substrate at subzero temperature" *Biochem J.* **219**, 437-444 (1984).

140. Mangel, W.F., Singer, P.T., Cyr, D.M., Umland, T.C., Toledo, D.L., Stroud, R.M., Pflugrath, J.W. & Sweet, R.M. "Structure of an Acyl-Enzyme Intermediate during Catalysis: (Guanisinobenzoyl)trypsin" *Biochemistry* **29**, 8351-8357 (1990).

141. Maniatis, T., Fritsch, E.F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) 545 pages (1982).

142. Maroux, S., Baratti, J. & Desnuelle, P. "Purification and Specificity of Porcine Enterokinase" *J. Biol. Chem.* **246**, 5031-5039 (1971).

143. Marquart, M., Walter, J., Deisenhofer, J., Bode, W. & Huber, R. "The Geometry of the Reactive Site and of the Peptide Groups in Trypsin, Trypsinogen, and Its Complexes with Inhibitors" *Acta Cryst.* **B39**, 480-490 (1983).

144. Massey, V., Harrington, W.F. & Hartley, B.S. "Certain Physical Properties of Chymotrypsin and Chymotrypsinogen using the Depolarization of Fluorescence Technique" *Discuss. Faraday Soc.* **20**, 24-32 (1955).

145. Matthews, B.W., Sigler, P.B., Henderson, R. & Blow, D.M. "Three-dimensional Structure of Tosyl-α-chymotrypsin" *Nature* **214**, 652-656 (1967).

146. McGrath, M.E., Fletterick, R.J. & Craik, C.S. "Rapid Preparation of Proteins for Crystallization Trials" *Biotechniques* **7**, 246-247 (1989).

147. McGrath, M.E., Hines, W.M., Sakanari, J.A., Fletterick, R.J. & Craik, C.S. "The Sequence and Reactive Site of Ecotin: A General Inhibitor of Pancreatic Serine Proteases from *E. coli*" *J. Biol. Chem. (submitted)* (1990).

148. McGrath, M.E., Vásquez, J.R., Craik, C.S., Yang, A.-S., Honig, B. & Fletterick, R.J. "Electrostatics in Serine Protease Catalysis" *Submitted to Science* (1990).

149. McGrath, M.E., Wilke, M.E., Higaki, J.N., Craik, C.S. & Fletterick, R.J. "Crystal Structure of Two Engineered Thiol Trypsins" *Biochemistry* **28**, 9264-9270 (1989).

150. McPherson, A. *Preparation and Analysis of Protein Crystals* (John Wiley & Sons, New York) 371 pages (1982).

151. McRae, B.J., Kurachi, K., Heimark, R.L., Fujikawa, K., Davie, E.W. & Powers, J.C. "Mapping the Active Sites of Bovine Thrombin, Factor IXa, Factor Xa, Factor XIa, Factor XIIa, Plasma Kallikrein, and Trypsin with Amino Acid and Peptide Thioesters: Development of New Sensitive Substrates" *Biochemistry* **20**, 7196-7206 (1981).

152. Messing, J. "New M13 Vectors for Cloning" *Meth. Enzymol.* **101**, 20-78 (1983).

153. Meyer, E., Cole, G., Radhakrishnan, R. & Epp, O. "Structure of Native Porcine Pancreatic Elastase" *Acta Cryst.* **B44**, 26-38 (1988).

154. Michaelis, L. & Menten, M.L. "Die Kinetik der Invertinwirkung" *Biochem. Z.* **49**, 333-369 (1913).

155. Morgan, P.H., Robinson, N.C., Walsh, K.A. & Neurath, H. "Inactivation of Bovine Trypsinogen and Chymotrypsinogen by Diisopropylphosphorofluoridate" *Proc. Nat. Acad. Sci. USA* **69**, 3312-3316 (1972).

156. Müller-Hill, B., Crapo, L. & Gilbert, W. "Mutants that make more lac repressor" *Proc. Natl. Acad. Sci. USA* **59**, 1259-1264 (1968).

157. Najarian, R. *Snipper* (Chiron, Corp., Emeryville, CA).

158. Nakagawa, Y. & Bender, M.L. "Methylation of Histidine-57 in α-Chymotrypsin by Methyl *p*-Nitrobenzenesulfonate. A New Approach to Enzyme Modification" *Biochemistry* **9**, 259-267 (1970).

159. Neuhard, J. & Nygaard, P. "Biosynthesis and Conversion of Nucleotides: Purines and Pyrimidines" in *Eschericia coli and Salmonella typhimurium Cellular and Molecular Biology* Neidhardt, F.C., *et al.*, eds., (American Society for Microbiology, Washington, D. C.) pp. 445-473 (1987).

160. Olafson, R.W., Jurásek, L., Carpenter, M.R. & Smillie, L.B. "Amino Acid Sequence of *Streptomyces griseus* Trypsin. Cyanogen Bromide Fragments and Complete Sequence" *Biochemistry* **14**, 1168-1177 (1975).

161. Ong, E.B., Shaw, E. & Schoellmann, G. "An Active Center Histidine peptide of α-Chymotrypsin" *J. Am. Chem. Soc.* **86**, 1271-1272 (1964).

162. Page, M.I. "Theories of Enzyme Catalysis" in *Enzyme Mechanisms* Page, M.I. & Williams, A., eds., (Royal Society of Chemistry, London) pp. 1-13 (1987).

163. Perkin-Elmer Ltd. *LS-5B Luminescence Spectrometer Operator's Manual* (Perkin-Elmer Limited, Beaconsfield, England) pages (1986).

164. Pitt, G.S. & Devreotes, P.N. "A simple BASIC program allows the rapid entry of DNA nucleotide sequences into personal computers" *Biotechniques* **6**, 122-123 (1988).

165. Polgar, L. *Mechanisms of Protease Action* (CRC Press, Inc., Boca Raton, Florida) 223 pages (1989).

166. Pratt, L.R. & Chandler, D. "Theory of the Hydrophobic Effect" *J. Chem. Phys.* **67**, 3683-3704 (1977).

167. Rampolla, R.W., Miller, R.C. & Smyth, C.P. "Microwave Absorption and Molecular Structure in Liquids. XXV. Measurements of Dielectric Constant and Loss at 3.1-mm Wavelength by an Interferometric Method" *J. Chem. Phys.* **30**, 566-573 (1959).

168. Read, R.J. & James, M.N.G. "Refined Crystal Structure of *Streptomyces griseus* Trypsin at 1.7Å Resolution" *J. Mol. Biol.* **200**, 523-551 (1988).

169. Rebek, J., Duff, R.J., Gordon, W.E. & Parris, K. "Convergent Functional Groups Provide a Measure of Stereoelectronic Effects at Carboxyl Oxygen" *J. Am. Chem. Soc.* **108**, 6068-6069 (1986).

170. Reitz, J.R., Milford, F.J. & Christy, R.W. *Foundations of Electromagnetic Theory* (Addison-Wesley Publishing Company, Menlo Park, California) 534 pages (1979).

171. Richardson, J.S. "The Anatomy and Taxonomy of Protein Structure" *Adv. Prot. Chem.* **34**, 167-339 (1981).

172. Ringe, D., Mottonen, J.M., Gelb, M.H. & Abeles, R.H. "X-ray Diffraction Analysis of the Inactivation of Chymotrypsin by 3-Benzyl-6-chloro-2-pyrone" *Biochemistry* **25**, 5633-5638 (1986).

173. Ringe, D., Seaton, B.A., Gelb, M.H. & Abeles, R.H. "Inactivation of Chymotrypsin by 5-Benzyl-6-chloro-2-pyrone: ^{13}C NMR and X-ray Diffraction Analyses of the Inactivator-Enzyme Complex" *Biochemistry* **24**, 64-68 (1985).

174. Ritchie, C.D. "Cation-Anion Combination Reactions. XIII. Correlation of the Reactions of Nucleophiles with Esters" *J. Am. Chem. Soc.* **97**, 1170-1179 (1975).

175. Robertus, J.D., Kraut, J., Alden, R.A. & Birktoft, J.J. "Subtilisin; a Stereochemical Mechanism Involving Transition-State Stabilization" *Biochemistry* **11**, 4293-4303 (1972).

176. Rühlmann, A., Kukla, D., Schwager, P., Bartels, K. & Huber, R. "Structure of the Complex formed by Bovine Trypsin and Bovine Pancreatic Trypsin Inhibitor: Crystal Structure Determination and Stereochemistry of the Contact Region" *J. Mol. Biol.* **77**, 417-436 (1973).

177. Russell, A.J. & Fersht, A.R. "Rational modification of enzyme catalysis by engineering surface charge" *Nature* **328**, 496-500 (1987).

178. Russell, A.J., Thomas, P.G. & Fersht, A.R. "Electrostatic Effects on Modification of Charged Groups in the Active Site Cleft of Subtilisin by Protein Engineering" *J. Mol. Biol.* **193**, 803-813 (1987).

179. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) pages (1989).

180. Sanger, F., Nicklen, S. & Coulson, A.R. "DNA sequencing with chain-terminating inhibitors" *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467 (1977).

181. Sauer, R.T., Smith, D.L. & Johnson, A.D. "Flexibility of the yeast α2 repressor enables it to occupy the ends of its operator, leaving the center free" *Genes Dev.* **2**, 807-816 (1988).

182. Schaffer, N.K., May, S.C. & Summerson, W.H. "Serine Phosphoric Acid from Diisopropylphosphoryl Chymotrypsin" *J. Biol. Chem.* **202**, 67-76 (1953).

183. Schecter, I. & Berger, A. "On the Size of the Active Site in Proteases. I. Papain" *Biochem. Biophys. Res. Commun.* **27**, 157-162 (1967).

184. Schimmel, P. "Hazards of Deducing Enzyme Structure-Activity Relationships on the Basis of Chemical Applications of Molecular Biology" *Acc. Chem. Res.* **22**, 232-233 (1989).

185. Schoellman, G. & Shaw, E. "Direct Evidence for the Presence of Histidine in the Active Center of Chymotrypsin" *Biochemistry* **2**, 252-255 (1963).

186. Scholten, J.D., Hogg, J.L. & Raushel, F.M. "Methyl Chymotrypsin Catalyzed Hydrolyses of Specific Substrate Esters Indicate Multiple Proton Catalysis is Possible with a Modified Charge Relay Triad" *J. Am. Chem. Soc.* **110**, 8246-8247 (1988).

187. Schowen, R.L. "Catalytic Power and Transition-State Stabilization" in *Transition States of Biochemical Processes* Gandour, R.D. & Schowen, R.L., eds., (Plenum, New York) pp. 77-114 (1978).

188. Schowen, R.L. "Structural and Energetic Aspects of Protolytic Catalysis by Enzymes. Charge Relay Catalysis in the Function of Serine Proteases" in *Mechanistic Principles of Enzyme Activity* Liebman, J.F. & Greenberg, A., eds., (VCH Publishers, Inc., Deerfield Beach, FL) pp. 400 (1988).

189. Schowen, R.L. "Structural and Energetic Aspects of Protolytic Catalysis by Enzymes. Charge-Relay Catalysis in the Function of Serine Proteases" in *Principles of Enzyme Activity* Liebman, J.F. & Greenberg, A., eds., (VCH Publishers, Inc., Deerfield Beach, FL) pp. (1988).

190. Schramm, H.J. "Modifizierung eines Histidin-Restes in Chymotrypsin durch 2-Phenyl-1,4-dibrom-acetoin" *Biochem. Z.* **342**, 139-142 (1965).

191. Schwert, G.W. & Eisenberg, M.A. "The Kinetics of the Amidase and Esterase Activities of Trypsin" *J. Biol. Chem.* **179**, 665-672 (1949).

192. Schwert, G.W., Neurath, H., Kaufman, S. & Snoke, J.E. "The Specific Esterase Activity of Trypsin" *J. Biol. Chem.* **172**, 221-239 (1948).

193. Sharp, K. & Honig, B. "Electrostatic Interactions in Macromolecules: Theory and Applications" *Ann. Rev. Biophys. & Biophys. Chem.* **19**, 301-332 (1990).

194. Shaw, E., Mares-Guia, M. & Cohen, W. "Evidence for an Active-Center Histidine in Trypsin through Use of a Specific Reagent, 1-Chloro-3-tosylamido-7-amino-2-heptone, the Chloromethyl Ketone derived from N^{α}-Tosyl-L-lysine" *Biochemistry* **4**, 2219-2224 (1965).

195. Shlomai, J. & Kornberg, A. "Deoxyuridine Triphosphatase of *Eschericia coli* Purification, properties, and use as a reagent to reduce uracil incorporation into DNA" *J. Biol. Chem.* **253**, 3305-3312 (1978).

196. Shotton, D.M. & Hartley, B.S. "Amino-acid Sequence of Porcine Pancreatic Elastase and its Homologies with other Serine Proteinases" *Nature* **225**, 802-806 (1970).

197. Shotton, D.M. & Watson, H.C. "Three-dimensional Structure of Tosyl-elastase" *Nature* **225**, 811-816 (1970).

198. Smillie, L.B. & Hartley, B.S. "Histidine Sequences in the Active Centres of Some 'Serine' Enzymes" *J. Mol. Biol.* **10**, 183-185 (1964).

199. Smith, M. "Synthetic Oligodeoxyribonucleotides as Probes for Nucleic Acids and as Primers in Sequence Determination" in *Methods of DNA and RNA Sequencing* Weissman, S.M., eds., (Praeger Publishers, New York) pp. 23-68 (1983).

200. Smith, S.O., Farr-Jones, S., Griffin, R.G. & Bachovchin, W.W. "Crystal Versus Solution Structures of Enzymes: NMR Spectroscopy of a Crystalline Serine Protease" *Science* **244**, 961-964 (1989).

201. Soman, K., Yang, A.-S., Honig, B. & Fletterick, R. "Electrical Potentials in Trypsin Isozymes" *Biochemistry* **28**, 9918-9926 (1989).

202. Spomer, W.E. & Wootton, J.F. "The Hydrolysis of α-*N*-Benzoyl-L-Argininamide Catalyzed by Trypsin and Acetyltrypsin: Dependence on pH" *Biochim. Biophys. Acta* **235**, 164-171 (1971).

203. Sprang, S., Standing, T., Fletterick, R.J., Stroud, R.M., Finer-Moore, J., Xuong, N.-H., Hamlin, R., Rutter, W.J. & Craik, C.S. "The Three-Dimensional Structure of Asn^{102} Mutant of Trypsin: Role of Asp^{102} in Serine Protease Catalysis" *Science* **237**, 905-909 (1987).

204. Stein, R.L. "Catalysis by Human Leukocyte Elastase. 4. Role of Secondary-Subsite Interactions" *J. Am. Chem. Soc.* **107**, 5767-5775 (1985).

205. Stein, R.L. & Strimpler, A.M. "P_1 Residue Determines the Operation of the Catalytic Triad of Serine Proteases during Hydrolyses of Acyl-Enzymes" *J. Am. Chem. Soc.* **109**, 4387-4390 (1987).

206. Stein, R.L., Strimpler, A.M., Hori, H. & Powers, J.C. "Catalysis by Human Leukocyte Elastase: Proton Inventory as a Mechanistic Probe" *Biochemistry* **26**, 1305-1314 (1987).

207. Steitz, T.A., Henderson, R. & Blow, D.M. "Structure of Crystalline α-Chymotrypsin III. Crystallographic Studies of Substrates and Inhibitors bound to the Active Site of α-Chymotrypsin" *J. Mol. Biol.* **46**, 337-348 (1969).

208. Sternberg, M.J.E., Hayes, F.R.F., Russell, A.J. & Fersht, A.R. "Prediction of electrostatic effects of engineering of protein charges" *Nature* **330**, 86-88 (1987).

209. Stevenson, K.J. & Smillie, L.B. "The Reaction of Phenoxymethyl Chloromethyl Ketone with Nitrogen 3 of Histidine-57 of Chymotrypsin" *J. Mol. Biol.* **12**, 937-941 (1965).

210. Stoddard, B.L., Bruhnke, J., Koenigs, P., Porter, N., Ringe, D. & Petsko, G.A. "Photolysis and Deacylation of Inhibited Chymotrypsin" *Biochemistry* **29**, 8042-8051 (1990).

211. Stoddard, B.L., Bruhnke, J., Porter, N., Ringe, D. & Petsko, G.A. "Structure and Activity of Two Photoreversible Cinnamates Bound to Chymotrypsin" *Biochemistry* **29**, 4871-4879 (1990).

212. Stroud, R.M. "A Family of Protein-Cutting Proteins" *Sci. Amer.* **231**, 74-88 (1974).

213. Stroud, R.M., Kay, L.M. & Dickerson, R.E. "The Crystal and Molecular Structure of DIP-inhibited Bovine Trypsin at 2.7Å Resolution" *Cold Spring Harbor Symposia in Quantitative Biology* **36**, 125-140 (1971).

214. Stroud, R.M., Kay, L.M. & Dickerson, R.E. "The Structure of Bovine Trypsin: Electron Density Maps of the Inhibited Enzyme at 5Å and at 2.7Å Resolution" *J. Mol. Biol.* **83**, 185-208 (1974).

215. Strumeyer, D.H., White, W.N. & Koshland, D.E., Jr. "Role of Serine in Chymotrypsin Action. Conversion of the Active Serine to Dehydroalanine" *Proc. Natl. Acad. Sci. USA* **50**, 931-935 (1963).

216. Swift, G.H., Craik, C.S., Stary, S.J., Quinto, C., Lahaie, R.G., Rutter, W.J. & MacDonald, R.J. "Structure of the Two Related Elastase Genes Expressed in the Rat Pancreas" *J. Biol. Chem.* **259**, 14721-14278 (1984).

217. Tabor, S. & Richardson, C.C. "DNA sequence analysis with a modified bacteriophage T7 DNA polymerase" *Proc. Natl. Acad. Sci. USA* **84**, 4767-4771 (1987).

218. Takagi, H., Morinaga, Y., Ikemura, H. & Inouye, M. "Mutant Subtilisin E with Enhanced Protease Activity Obtained by Site-directed Mutagenesis" *J. Biol. Chem.* **263**, 19592-19596 (1988).

219. Tanford, C. & Kirkwood, J.G. "Theory of Protein Titration Curves. I. General Equations for Impenetrable Spheres" *J. Am. Chem. Soc.* **79**, 5333-5339 (1957).

220. Thomas, P.G., Russell, A.J. & Fersht, A.R. "Tailoring the pH dependence of enzyme catalysis using protein engineering" *Nature* **318**, 375-376 (1985).

221. Trowbridge, C.G., Krehbiel, A. & Laskowski, M., Jr. "Substrate Activation of Trypsin" *Biochemistry* **2**, 843-850 (1963).

222. Tye, B.-K., Chen, J., Lehman, I.R., Duncan, B.K. & Warner, H.R. "Uracil incorporation: A source of pulse-labeled DNA fragments in the replication of the *Escherichia coli* chromosome" *Proc. Natl. Acad. Sci. USA* **75**, 233-237 (1978).

223. Tye, B.-K., Chien, J., Lehman, I.R., Duncan, B.K. & Warner, H.R. "Uracil incorporation: A source of pulse-labeled DNA fragments in the replication of the *Eschericia coli* chromosome" *Proc. Natl. Acad. Sci. USA* **75**, 233-237 (1978).

224. Umeyama, H., Nakagawa, S. & Kudo, T. "Role of Asp102 in the Enzymatic Reaction of Bovine β-Trypsin: A Molecular Orbital Study" *J. Mol. Biol.* **150**, 409-421 (1981).

225. UPI "Scientists Redesign an Enzyme to Alter Biological Function" (New York Times, page 1) April 13, 1985

226. USB Research. *SEQUENASE™: Step-by-Step Protocols For DNA Sequencing With Sequenase* (United States Biochemical Corp., Cleavland, OH)

227. Vásquez, J.R., Evnin, L.B., Higaki, J.N. & Craik, C.S. "An Expression System for Trypsin" *J. Cell. Biochem.* **39**, 265-276 (1989).

228. Vásquez, J.R., McGrath, M.E., Guenot, J.M., ?, I., Kollman, P.A., Fletterick, R.J. & Craik, C.S. "The Report of an Electrostatic Probe from the Intetrior of Trypsin" *in preparation* (1990).

229. Vásquez, J.R., McGrath, M.E., Higaki, J.N., Evnin, L.B., Fletterick, R.J. & Craik, C.S. "Experimental Evidence for Catalysis by Electrostatic Transition State Stabilization by Trypsin" *in preparation* (1990).

230. Venkatasubban, K.S. & Schowen, R.L. "The Proton Inventory Technique" *CRC Crit. Rev. Biochem.* **17**, 1-44 (1985).

231. Walsh, C. *Enzymatic Reaction Mechanisms* (W. H. Freeman and Company, San Francisco) 978 pages (1979).

232. Warshel, A. & Levitt, M. "Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme" *J. Mol. Biol.* **103**, 227-249 (1976).

233. Warshel, A., Naray-Szabo, G., Sussman, F. & Hwang, J.-K. "How Do Serine Proteases Really Work?" *Biochemistry* **28**, 3629-3637 (1989).

234. Warshel, A. & Russell, S.T. "Calculations of electrostatic interactions in biological systems and in solutions" *Q. Rev. Biophys.* **17**, 283-422 (1984).

235. Watson, H.C., Shotton, D.M., Cox, J.M. & Muirhead, H. "Three-dimensional Fourier Synthesis of Tosyl-elastase at 3.5Å Resolution" *Nature* **225**, 806-811 (1970).

236. Weidmann, B. & Abeles, R.H. "Mechanism of Inactivation of Chymotrypsin by 5-Butyl-3*H*-1,3-oxazine-2,6-dione" *Biochemistry* **23**, 2373-2376 (1984).

237. Weil, L., James, S. & Buchert, A.R. "Photooxidation of Crystalline Chymotrypsin in the Presence of Methylene Blue" *Arch. Biochem. Biophys.* **46**, 266-278 (1953).

238. Wells, J.A., Cunningham, B.C., Graycar, T.P. & Estell, D.A. "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" *Phil. Trans. R. Soc. London* **A317**, 415-423 (1986).

239. Wells, J.A., Powers, D.B., Bott, R.R., Graycar, T.P. & Estell, D.A. "Designing substrate specificity by protein engineering of electrostatic interactions" *Proc. Natl. Acad. Sci. USA* **84**, 1219-1223 (1987).

240. West, J.B., Scholten, J., Stolowich, N.J., Hogg, J.L., Scott, A.I. & Wong, C.-H. "Modification of Proteases to Esterases for Peptide Synthesis: Methylchymotrypsin" *J. Am. Chem. Soc.* **110**, 3709-3710 (1988).

241. Westkaemper, R.B. & Abeles, R.H. "Novel Inactivators of Serine Proteases Based on 6-Chloro-2-pyrone" *Biochemistry* **22**, 3256-3264 (1983).

242. Winkler, F.K., D'arcy, A. & Hunziker, W. "Structure of human pancreatic lipase" *Nature* **343**, 771-774 (1990).

243. Wolfenden, R. & Frick, L. "Transition State Affinity and the Design of Enzyme Inhibitors" in *Enzyme Mechanisms* Page, M.I. & Williams, A., eds., (Royal Society of Chemistry, London) pp. 97-122 (1987).

244. Woolfson, M.M. *An Introduction to X-Ray Crystallography* (Cambridge University Press, Cambridge) 380 pages (1970).

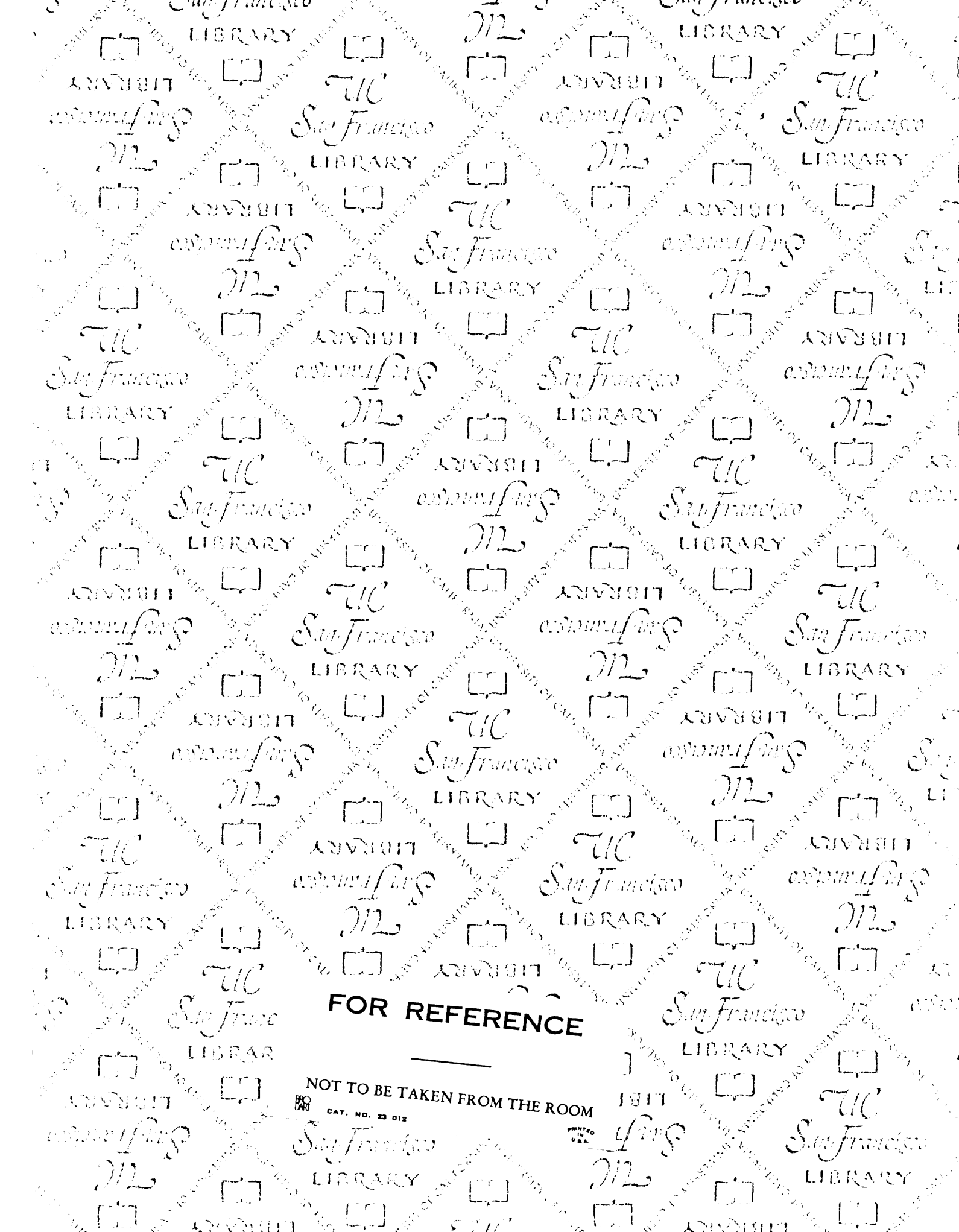
245. Wright, C.S., Alden, R.A. & Kraut, J. "Structure of Subtilisin BPN' at 2.5Å Resolution" *Nature* **221**, 235-242 (1969).

246. Zerner, B. & Bender, M. "The Kinetic Consequences of the Acyl-Enzyme Mechanism for the Reactions of Specific Substrates with Chymotrypsin" *J. Am. Chem. Soc.* **86**, 3669-3674 (1964).

247. Zerner, B., Bond, R.P.M. & Bender, M.L. "Kinetic Evidence for the Formation of Acyl-Enzyme Intermediates in the α-Chymotrypsin-Catalyzed Hydrolyses of Specific Substrates" *J. Am. Chem. Soc.* **86**, 3674-3679 (1964).

248. Zimmerman, S.C. & Cramer, K.D. "Stereoelectronic Effects at Carboxylate: A Syn Oriented Model for the Histidine-Aspartate Couple in Enzymes" *J. Am. Chem. Soc.* **110**, 5906-5908 (1988).

249. Zoller, M.J. & Smith, M. "Oligonucleotide-Directed Mutagenesis of DNA Fragments Cloned into M13 Vectors" *Meth. Enzymol.* **101**, 468-500 (1983).



FOR REFERENCE

NOT TO BE TAKEN FROM THE ROOM