Serine Protease Mechanism and Specificity

Lizbeth Hedstrom

Department of Biochemistry, MS 009, Brandeis University, Waltham, Massachusetts 02454

Received May 14, 2002

Contents

* E-mail: hedstrom@brandeis.edu; phone: 781-736-2333; FAX: 781-736-2349.

I. Introduction

Almost one-third of all proteases can be classified as serine proteases, named for the nucleophilic Ser residue at the active site. This mechanistic class was originally distinguished by the presence of the Asp−His−Ser "charge relay" system or "catalytic triad".[1] The Asp−His−Ser triad can be found in at least four different structural contexts, indicating that this catalytic machinery has evolved on at least four separate occasions.[2] These four clans of serine proteases are typified by chymotrypsin, subtilisin, carboxypeptidase Y, and Clp protease (MEROPS nomenclature;[3] Table 1). More recently, serine proteases with novel catalytic triads and dyads have been discovered, including Ser−His−Glu, Ser−Lys/His, His−Ser−His, and N-terminal Ser.[2] Several of these novel serine proteases are subjects of accompanying articles in this issue.

This article will review recent work on the mechanism and specificity of chymotrypsin-like enzymes, with the occasional references to pertinent experiments with subtilisin. Chymotrypsin-like proteases are the most abundant in nature, with over 240 proteases recognized in the MEROPS database.[3]

10.1021/cr000033x CCC: $39.75
Published on Web 11/23/2002

Table 1. Serine Protease Clans[a]

clan	members	example	catalytic residues	distribution
		"Classic" Catalytic Triad Serine Proteases		
PA(S)	301	chymotrypsin	His−Asp−Ser	B, Ar, F, Pl, An, V
SB	91	subtilisin	Asp−His−Ser	B, Ar, Pr, F,Pl,An,V
SC	64	carboxypeptidase Y	Ser−Asp−His	B, Ar, Pr, F, Pl, An,
SK	14	Clp protease	Ser−His−Asp	B, Ar, Pr, F, Pl,An,V
		"Novel" Serine Proteases		
SE	16	D-Ala-D-Ala carboxypeptidase A	Ser−Lys	B, Ar, Pl, An
SF	24	signal peptidase I	Ser−Lys/His	B, Ar, F, Pl, An, V
SH	7	cytomegalovirus assemblin	His−Ser−His	V
SM	7	C-terminal processing protease-1	Ser−Lys	B, Ar, Pl
SN	4	dipeptidase E	Ser−His−Glu	B, An
PB(S)	4	penicillin amidohydrolase precursor	N-terminal Ser	B, Ar, Pr

[a] The MEROPS database classifies proteases based on structural similarity [Rawlings, N. D., and Barrett, A. J. (2000) MEROPS: The Peptidase Database. *Nucleic Acids Res. 28*, 323−325]. The following information is tabulated from MEROPS release 5.7 [http://merops.iapc.bbsrc.ac.uk/]. Catalytic residues are listed in the order that they appear in the primary sequence. B, bacteria; Ar, archaea; Az, archezoa; Pr, protozoa; F, fungi; Pl, plants; An, animals; V, viruses. In addition to the proteases listed below, 21 serine proteases have been identified that are not yet assigned to a clan.

These proteases can be found in eukaryotes, prokaryotes, archae, and viruses. Chymotrypsin-like proteases are involved in many critical physiological processes, including digestion, hemostasis, apoptosis, signal transduction, reproduction, and the immune response (Table 2).[4−8] "Cascades" of sequential activation of serine proteases drive blood coagulation, complement fixation, and fibrinolysis.[9−11] Similar protease cascades appear to be involved in development, matrix remodeling, differentiation, and wound healing.[12−14] These diverse physiological niches demand proteases with wildly varied specificities, ranging from digestive proteases that cleave after hydrophobic or positively charged residues, to proteases that recognize a five-residue cleavage site or even a single protein. Serine protease specificity can usually be rationalized by the topology of the substrate binding sites adjacent to the catalytic triad (the "active site cleft"). Several excellent reviews have recently reported on this topic.[15−17] However, while this perspective implies that specificity can be reduced to discrimination in substrate binding, specificity is primarily expressed in the rates of chemical transformation. This review will discuss the expression of specificity during catalysis and strategies for substrate discrimination, focusing on the chymotrypsin−trypsin−elastase paradigm and including examples from Factor D, alpha lytic protease, and thrombin.

Table 2. Representative Mammalian Chymotrypsin-Like Proteases

- digestive proteases
  - chymotrypsin
  - trypsin
  - pancreatic elastase
- immune response
  - tryptase
  - complement factor D
  - complement factor B
  - complement factor C
  - complement component 2
  - mast cell protease
  - cathepsin G
  - neutrophil elastase
- blood coagulation
  - coagulation factor VIIa
  - coagulation factor IXa
  - coagulation factor Xa
  - coagulation factor XIIa
  - thrombin
  - protein C
- fibrinolysis
  - urokinase
  - tissue plasminogen activator
  - plasmin
  - kallikrein
- reproduction
  - prostate specific antigen
  - acrosein
- bacterial homologues
  - alpha lytic protease
  - *Streptomyces griseus* protease A
  - *Streptomyces griseus* protease B

Dr. Lizbeth Hedstrom was born in Frederick, Maryland, in 1958. She obtained a Bachelor of Science in Chemistry at the University of Virginia and then moved to Brandeis University where she obtained a Ph.D. in Biochemistry under the direction of Robert Abeles in 1985. After postdoctoral work at the University of California San Francisco with C. C. Wang and William Rutter, Dr. Hedstrom returned to Brandeis University in 1992. She is currently the Markey Associate Professor of Biochemistry. Her research interests include structure/function relationships in enzyme catalysis, protein engineering, the design of novel enzyme inhibitors as mechanistic probes, and potential pharmaceutical agents. Her work currently focuses on serine proteases and enzymes involved nucleotide metabolism.

II. The Structure of Chymotrypsin-Like Serine Proteases

First we will consider the structure of chymotrypsin-like proteases (Figure 1). Chymotrypsin has

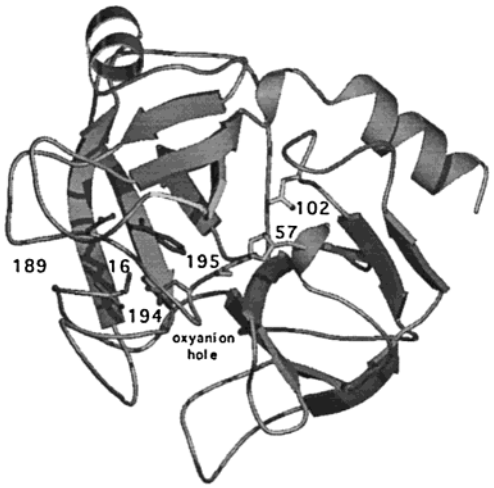


Figure 1. Structure of chymotrypsin. This figure was produced with Molscript v2.1 (ref 267) and POV-Ray by Annette Pasternak.

245 residues arranged in two six-stranded beta barrels.[18] The active site cleft is located between the two barrels. This structure is generally divided into catalytic, substrate recognition and zymogen activation domain components that are common to all chymotrypsin-like serine proteases. (The "loop 1" nomenclature used in some previous papers and reviews will be replaced with a system based on chymotrypsinogen numbering to facilitate comparison among different proteases.) This organization is derived from Kraut[19] and is convenient for discussion, but wrongly implies that catalysis, substrate recognition, and zymogen activation are separate processes. As will be seen, these three processes involve many of the same structural features and are intricately intertwined. In addition to the protease domain, many serine proteases are modular proteins containing auxiliary regulatory domains (e.g., kringle, EGF, etc.).[20] The functions of these auxiliary domains will not be considered here.

A. The Catalytic Components

1. The Catalytic Triad

The catalytic triad spans the active site cleft, with Ser195 on one side and Asp102 and His57 on the other (Figure 1). His57 assumes the less favorable Nδ1-H tautomer (Figure 2), indicating that the tautomer equilibrium is perturbed at least 100$\times$ from the solution position.[21] The catalytic triad is part of an extensive hydrogen bonding network (Figure 2A). Hydrogen bonds are generally observed between the Nδ1-H of His57 and Oδ1 of Asp102 and between the OH of Ser195 and the Nϵ2-H of His57, although the latter hydrogen bond is lost when His57 is protonated (Figure 2). Similar hydrogen bonds are observed in protein inhibitor (eglin C), acylenzyme, and transition state analogue (trifluoroketone) complexes, which suggests that these hydrogen bonds are present thoughout the catalytic cycle. Interestingly, the His57−Asp102 hydrogen bond has the syn orientation relative to the carboxylate, so that the hydrogen bond forms with the more basic electron pair.[22,23] The OH of Ser214 forms a hydrogen bond with Oδ1 of Asp102 in almost all chymotrypsin-like proteases. Ser214 was once considered the fourth member of the catalytic triad, although more recent evidence indicates that it is dispensable.[24,25] Hydrogen bonds are also observed between the Oδ2 of Asp102 and the main chain NHs of Ala56 and His57. These hydrogen

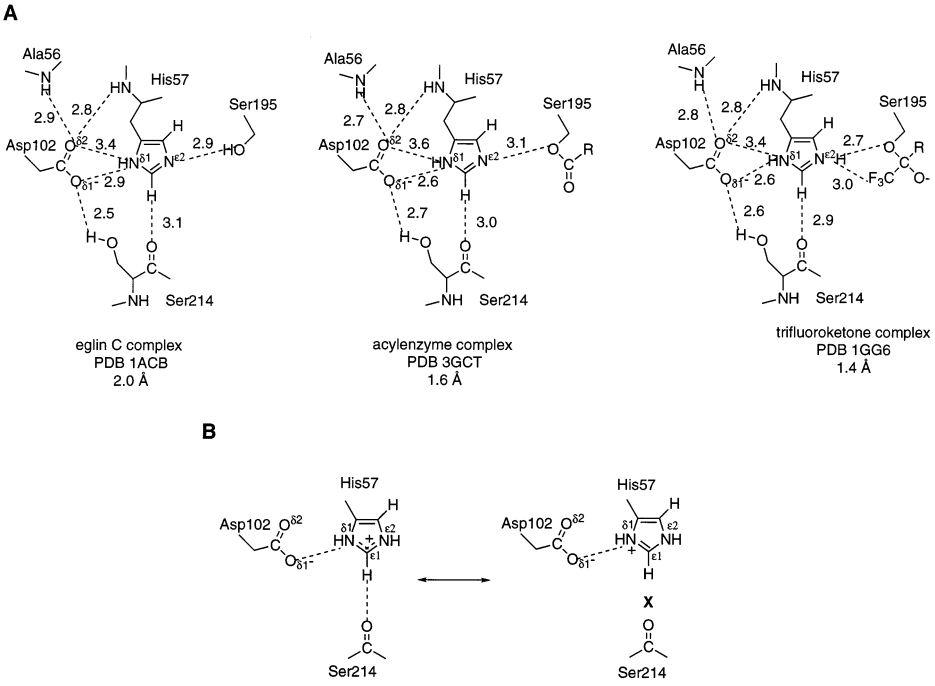


Figure 2. Hydrogen bonding in the catalytic triad. (A) The catalytic triad in chymotrypsin complexes. The hydrogen bonding network of the catalytic triad is depicted for the complex with a protein inhibitor eglin C, an acylenzyme intermediate and a trifluoroketone transition state analogue. The dotted lines represent potential hydrogen bonds and the numbers are the distances between heteroatoms in Å. (B) The effect of the Cϵ1-H hydrogen bond on charge distribution in His57. In the absence of the hydrogen bond, the positive charge will accumulate at Nδ1 where it can be stabilized by Asp102.

bonds are believed to orient Asp102 and His57. In addition, a novel hydrogen bond is observed between the Cϵ1-H of His57 and the main chain carbonyl of Ser214.[26,27] Similar hydrogen bonds are also observed in subtilisin and other classes of serine hydrolases, suggesting that this interaction may be important for catalytic triad function.[26] This hydrogen bond may prevent the localization of positive charge on Nδ1 of His57 (Figure 2B). Interestingly, the carbonyl oxygen of Ser214 is also part of the polypeptide binding site (see below), and may mediate communication between substrate and the catalytic triad.

2. The Oxyanion Hole

The oxyanion hole is formed by the backbone NHs of Gly193 and Ser195. These atoms form a pocket of positive charge that activates the carbonyl of the scissile peptide bond and stabilizes the negatively charged oxyanion of the tetrahedral intermediate (see below). The oxyanion hole is structurally linked to the catalytic triad and the Ile16−Asp194 salt bridge via Ser195. Subtilisin also contains an oxyanion hole, formed by the side chains of Asn155, Thr220, and the backbone NH of Ser221.[28]

B. The Substrate Recognition Sites

The substrate recognition sites include the polypeptide binding site and the binding pockets for the side chains of the peptide substrate. Investigations of protease specificity have generally focused on the P1/S1 interaction (based on the nomenclature of Schechter and Berger (ref 29), where P1−P1′ denotes peptide residues on the acyl and leaving group side of the scissile bond, respectively. The adjacent peptide residues are numbered outward, and S1, S1′, etc. denote the corresponding enzyme binding sites), followed by consideration of the P2−Pn/S2−Sn interactions. The S1′−Sn′ sites are generally overlooked due to the prevalence of assays using substrates with chromophoric leaving groups. We will consider the S1 site, the polypeptide binding site, the S2−Sn and S1′−Sn′ sites separately.

1. The S1 Site

Specificity of chymotrypsin-like serine proteases is usually categorized in terms of the P1−S1 interaction. The S1 site is a pocket adjacent to Ser195, formed by residues 189−192, 214−216, and 224−228. Specificity is usually determined by the residues at positions 189, 216, and 226 (reviewed in refs 15 and 17). For example, the specificity of chymotrypsin correlates with the hydrophobicity of the P1 residue, with P1−Phe preferred over Ala by 50000.[30,31] The combination of Ser189, Gly216, and Gly226 create a deep hydrophobic pocket in chymotrypsin that accounts for this specificity.[18] Asp189, Gly216, and Gly226 create a negatively charged S1 site that accounts for trypsin's specificity for substrates containing Arg or Lys at P1.[32,33] Elastase prefers substrates with small aliphatic residues at P1; the S1 site of elastase is smaller than the S1 sites of chymotrypsin and trypsin due to the presence of Val216 and Thr226.[34] The specificity of granzyme B for acidic P1 residues is attributed to Arg226.[35] These observations suggest that a small set of structural elements controls the specificity of these proteases. However, as detailed below, simple transposition of these structural elements does not switch S1 specificity.

2. The Polypeptide Binding Site

Protease−substrate interactions extend beyond the S1 site, minimally including the polypeptide binding site and often involving additional binding pockets. The polypeptide binding site refers to the main chain of residues 214−216 which form an antiparallel beta sheet with the backbone of the P1−P3 residues of a peptide substrate (Figure 3). In chymotrypsin, hydrogen bonds form between the carbonyl oxygen of Ser214 and the NH of the P1 residue, the NH of Trp215 and the carbonyl of P3 and the carbonyl of Gly216 and the NH of P3. These interactions are a general feature of chymotrypsin-like proteases and are critical for efficient substrate hydrolysis. Interestingly, Gly216 has different conformations in chymotrypsin, trypsin, and elastase,[16] which suggests that the strength of this hydrogen bond will vary. Note that residues 214−216 also form one wall of the S1 site, and that the carbonyl of Ser214 forms a hydrogen bond to His57. These structural interactions form a line of communication between the polypeptide binding site, the S1 site, and the catalytic triad.

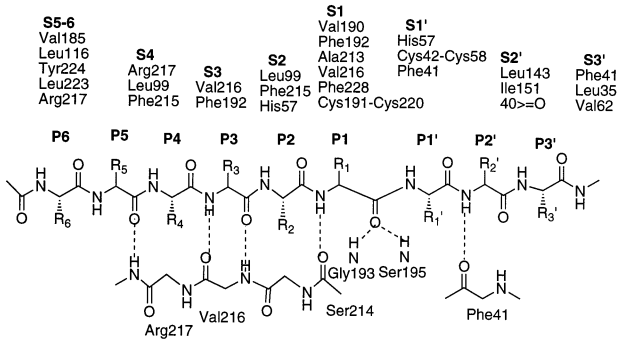


Figure 3. Subsite binding of porcine pancreatic and human leukocyte elastase. The hydrogen bonding interaction of the polypeptide binding site and S′ sites are shown by dotted lines. The residues that form each subsite are listed. Modified from ref 53.

Interestingly, the inhibitory loops of protein inhibitors of serine proteases are locked in the extended conformation required to form the antiparallel beta sheet with the polypeptide binding site. This extended conformation, termed the "canonical conformation", is observed in numerous protease−inhibitor complexes, including transition state analogues and peptidyl acylenzymes.[36−45] The canonical conformation includes the P3−P3′ residues and can extend to the amide NH of residue 218 and the carbonyl oxygen of P5 as observed in elastase (Figure 3). Importantly, the beta sheet structure causes the side chains the peptide substate to point in alternating directions.

The relevance of the antiparallel beta sheet interaction to peptide hydrolysis has recently been called into question because disruption of these hydrogen

bonding interactions had little effect on substrate hydrolysis.[46] However, the values of K_d for protein inhibitors and transition state analogues correlate with k_{cat}/K_m for hydrolysis of the analogous substrates, which argues strongly that substrates also assume the canonical conformation.[38,47−49]

3. The S2–Sn Sites

The S2−S3 sites of chymotrypsin display little substrate discrimination, in keeping with chymotrypsin's function as a digestive protease.[50] The S2 site is a shallow hydrophobic groove bounded by His57, Trp215, and Leu99; the specificity of the S2 site correlates with hydrophobicity of the P2 side chain.[49] The S3 site of chymotrypsin has little specificity, and can even accommodate D-amino acids; the side chain of the P3 residue protrudes out of the active site cleft.

Insertions in the loops surrounding the active site cleft create more defined pockets for the S2−Sn sites of other serine proteases. These sites can be the major determinant of specificity,[51] as exemplifed by elastase, where the P3/S3 interaction is dominant,[52,53] and enterokinase (also known as enteropeptidase), which recognizes the P5−P1 sequence Asp−Asp−Asp−Asp−Lys with restriction enzyme-like precision.[54]

4. The S1′–S3′ Sites

The interactions of the leaving group with the S1′−Sn′ sites are usually inferred from protein inhibitor structures that do not contain the optimum P1′−Pn′ residues. Therefore, the precise interactions of the P′ side chains have not been defined. A hydrogen bond between the mainchain NH of the P2′ residue and the mainchain carbonyl oxygen of Phe41 is a constant feature of the P′/S′ interactions. Chymotrypsin prefers substrates with P1′−Arg or Lys, which is attributed to electrostatic interactions with Asp35 and Asp64.[55] The P1′ and P3′ residues point in the same direction as a consequence of the beta sheet alignment of the substrate, so that the S1′ and S3′ sites overlap. The S1′/S3′ sites are bounded by His57, the 60's loop and the 40's loop. The P2′ residue points in the opposite direction and may interact with the 150's loop. As above, insertions in these loops can create more defined S1′−S3′ binding sites.[51]

C. The Zymogen Activation Domain

Chymotrypsin-like proteases are synthesized as inactive precursors ("zymogens") containing N-terminal extensions. Four segments are deformed in the zymogens of chymotrypsin and trypsin: the N-terminus to residue 19, residues 142−152, 184−193, and 216−223 (these regions are collectively termed the activation domain[36]). This deformed region includes the S1 site and oxyanion hole, which explains the low activity of the zymogen. Proteolytic processing activates the zymogen, releasing the N-terminal Ile16. The new N-terminus forms a buried salt bridge with Asp194, inducing a conformational change that orders the activation domain. The S1 site and oxyanion hole are formed, creating the active protease. Importantly, the loss of protease activity at high pH is attributed to the deprotonation of the N-terminus and disruption of Ile16−Asp194 salt bridge, shifting the conformational equilibrium toward an inactive zymogen-like conformation.[56,57] This mechanism of zymogen activation is conserved among mammalian chymotrypsin-like serine proteases. Bacterial homologues have a different mechanism of zymogen activation, but retain a salt bridge involving Asp194.[58]

III. The Mechanism of Serine Proteases

Obviously, a detailed understanding of the enzymatic reaction mechanism is required to understand how specificity is generated. Fortunately, chymotrypsin-like proteases occupy a distinguished place in the annals of biochemistry as perhaps the best studied enzyme system.

A. The Generally Accepted Chemical Mechanism of Serine Protease Catalysis

All proteases must overcome three obstacles to hydrolyze a peptide bond: (a) amide bonds are very stable due to electron donation from the amide nitrogen to the carbonyl. For comparison, a simple alkyl ester is ∼3000× more reactive than an amide bond, while a *p*-nitrophenyl ester is 300000× more reactive.[59] Proteases usually activate an amide bond via the interaction of the carbonyl oxygen with a general acid, and may also distort the peptide bond to disrupt resonance stabilization; (b) water is a poor nucleophile; proteases always activate water, usually via a general base; and (c) amines are poor leaving groups; proteases protonate the amine prior to expulsion. Serine proteases perform these tasks very efficiently: the rates of peptide hydrolysis by serine proteases are $\sim 10^{10}$-fold greater than the uncatalyzed reactions. Obviously, these mechanisms of catalysis are not confined to peptide hydrolysis; serine proteases also readily hydrolyze other acyl compounds, including amides, anilides, esters, and thioesters.

Figure 5 displays the generally accepted mechanism for chymotrypsin-like serine proteases. In the acylation half of the reaction, Ser195 attacks the carbonyl of the peptide substrate, assisted by His57 acting as a general base, to yield a tetrahedral intermediate. The resulting His57-H^+ is stabilized by the hydrogen bond to Asp102. The oxyanion of the tetrahedral intermediate is stabilized by interaction with the main chain NHs of the oxyanion hole. The tetrahedral intermediate collapses with expulsion of leaving group, assisted by His57-H^+ acting as a general acid, to yield the acylenzyme intermediate.

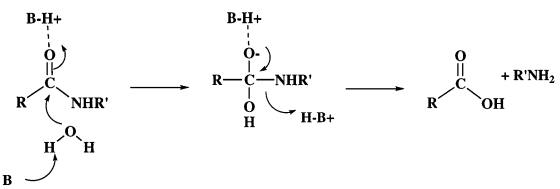


Figure 4. Catalysis of peptide hydrolysis. The catalysis of peptide hydrolysis involves activation of the carbonyl via a general acid, activation of water with a general base, and protonation of the amine leaving group.

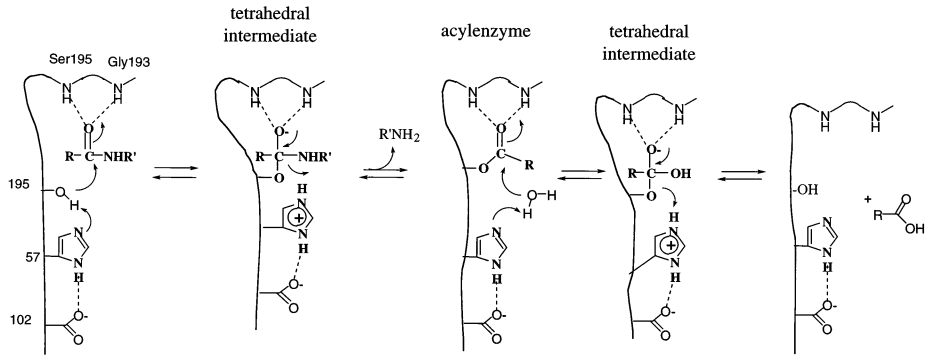


Figure 5. The generally accepted mechanism for serine proteases.

The deacylation half of the reaction essentially repeats the above sequence: water attacks the acylenzyme, assisted by His57, yielding a second tetrahedral intermediate. This intermediate collapses, expelling Ser195 and carboxylic acid product. The acylenzyme intermediate is well established as discussed below, but the tetrahedral intermediates are inferred from solution chemistry. The transition states of the acylation and deacylation reactions will resemble the high energy tetrahedral intermediates, and the terms "transition state" and "tetrahedral intermediate" are often used indiscriminately in the literature. It is worth noting that a web of hydrogen bonding interactions links the substrate binding sites to the catalytic triad. As the reaction proceeds, changes in bonding and charge at the scissile bond will propagate to more remote enzyme−substrate interactions, and vice versa. The evidence supporting this mechanism has been reviewed extensively, and will be only briefly described here.[19,56,60−62]

1. The Discovery of the Catalytic Triad and the Oxyanion Hole

Initial investigations of serine proteases focused on enzymes that could be readily obtained in large quantities: bovine chymotrypsin and trypsin and porcine elastase. Inactivators revealed the first clues into the catalytic machinery. Nerve gases such as diisopropylfluorophosphate (DFP) inactivate chymotrypsin by specifically modifying Ser195 (Figure 6)[63−65] and His57 is modified when chymotrypsin is inactivated by tosyl-L-phenylalanine chloromethyl ketone (TPCK).[66] The involvement of His57 was also suggested by the pH rate profiles which indicated that the deprotonation of a residue with p$K_a \sim 7$ was important for activity.[56] Subsequent NMR experiments showed that the pK_a of His57 is ~7.[21] The third member of the catalytic triad, Asp102, was discovered when Blow solved the crystal structure of chymotrypsin.[67] Blow proposed that the hydrogen bonding network between Asp102, His57, and Ser195—the charge relay system—activates Ser195 for nucleophilic attack. A charge relay system was rapidly identified in trypsin, elastase, and other related enzymes,[32,33,68] and a similar catalytic triad of Asp32, His 64, and Ser221 was discovered in subtilisin.[69,70] The oxyanion hole was first noted by Henderson, who proposed that the main chain NHs of Gly193 and Ser195 stabilize the negatively charged oxyanion of

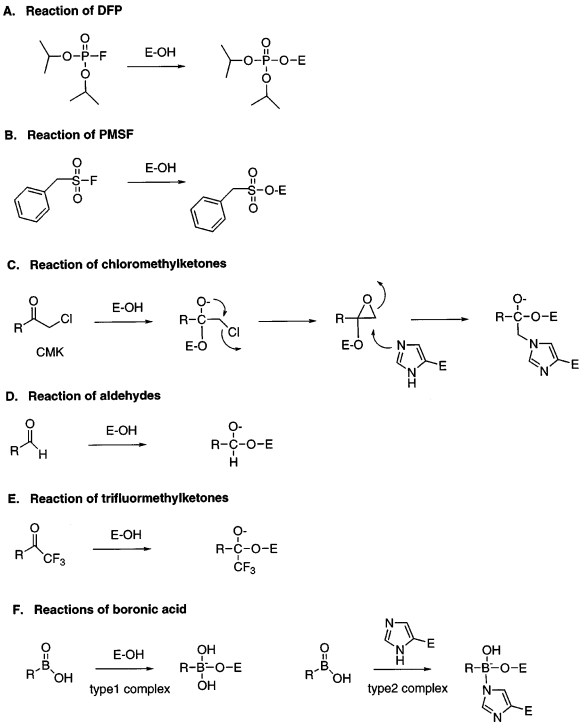


Figure 6. Mechanism inhibitors that form tetrahedral adducts.

the tetrahedral intermediate.[71] Crystal structures of several transition state analogue complexes confirm this interaction.[33,45,72−74]

2. Mutagenesis of the Catalytic Triad

Many experiments confirm the importance of the catalytic triad residues in serine protease catalysis. Chemical modification experiments were first used to disrupt the catalytic triad, either removing the hydroxyl of Ser195 to create anhydro proteases or methylating His57 at Nϵ2.[75,76] Either modification decreases protease activity by at least 10^4-fold.[77,78] Site-directed mutagenesis experiments provided a more incisive means to alter the catalytic triad. In trypsin, the substitution of either Ser195 or His57 with Ala decreases the value of k_{cat}/K_m by 10^5.[79] (Such values should most rigorously be considered upper limits for the activity of a given mutant enzyme. Water can substitute for a missing functional group, thus providing an alternative mechanism of catalysis and underestimating the effect of the mutation. In addition, factors such as translation misincorporation

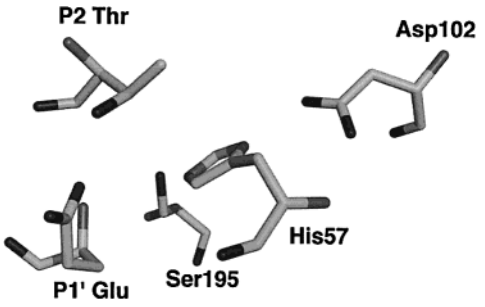


Figure 7. The P2 and P1′ residues bracket His57. The P2 and P1′ residues from the chymotrypsin-turkey ovomucoid third domain complex at 2.0 Å, PDB accession number 1CHO. This figure was produced with Molscript v2.1 (ref 267) and POV-Ray by Lu Gan.

and contaminating enzymes must also be considered when evaluating mutant enzymes with very low activities.) No further loss of activity is observed when the remaining catalytic triad residues are substituted. Thus, the substitution of either His57 or Ser195 is sufficient to completely disable the catalytic triad. The substitution of Asp102 to Asn in trypsin decreases k_{cat}/K_m by 10^4 at neutral pH.[80,81] The analogous mutations of the catalytic triad have similar effects in subtilisin.[82] However, it is important to recognize that the catalytic triad is not the sole source of protease catalytic power: even with the catalytic triad disabled, these mutant proteases hydrolyze substrates at $\sim 10^3\times$ the uncatalyzed rate.[79,82] This remaining activity reflects the contribution of the oxyanion hole, desolvation, and perhaps the canonical conformation to catalysis.

Interestingly, the activity of His57 mutations can be rescued by substrates containing His. This substrate assisted catalysis was first demonstrated in subtilisin and has recently been extended to trypsin.[83−86] In the absence of His57, the presence of P2−His or P1′-His increases peptide hydrolysis by as much as 300-fold. As noted above, His57 interacts with both the P2 and P1′ residues (Figure 7); it is believed that the P2- and P1′-His can replace His57 in the catalytic triad, forming similar interactions with Ser195 and Asp102.

A second instance of chemical rescue is found in mutations of Asp102. Hydroxide ion increases the activity of Asp102Asn trypsin to within 10% of wild type at pH 10.[80] Similar hydroxide activation is also observed in subtilisin.[82] The mechanism of this hydroxide rescue is not understood.

Unfortunately, since the oxyanion hole of chymotrypsin-like enzymes is composed of mainchain atoms, site directed mutagenesis cannot be used to confirm its role in catalysis. The oxyanion hole of subtilisin is composed of the side chains of Asn155 and Thr220 as well as the backbone NH of Ser221. Substitution of Asn155 decreases k_{cat}/K_m by $100-10^3$-fold, indicating that this residue provides 3−4 kcal/mol of transition state stabilization energy.[87−90] The substitution of Thr220 with Ala decreases k_{cat}/K_m by ∼20-fold (∼2 kcal/mol) and the removal of both Asn155 and Thr220 decreases activity by 10^4-fold, suggesting that the oxyanion hole can stabilize the transition state by as much as 6 kcal/mol. However, no further loss of activity is observed when oxyanion mutations are combined with mutations of the catalytic triad, indicating that the oxyanion and catalytic triad function cooperatively. Therefore, the 6 kcal/mol of transition state stabilization is not simply the measure of the interactions at the oxyanion hole.[90] The collaborative function of the catalytic machinery makes it impossible to precisely ascribe energies to individual interactions. Interestingly, Thr220 appears to be too far away to hydrogen bond directly to a tetrahedral intermediate; either Thr220 moves during catalysis, or stabilization results from long-range electrostatic interactions. Molecular modeling experiments support both mechanisms.[88,91]

3. Evidence for the Acylenzyme Intermediate

The famous burst experiment of Hartley and Kilbey was the first evidence for the acylenzyme intermediate.[92] The hydrolysis of *p*-nitrophenyl acetate by chymotrypsin is biphasic, with one enzyme-equivalent of *p*-nitrophenol produced in the rapid phase, indicating that an enzyme−acetate intermediate formed and decomposition of this intermediate was rate-limiting. Acylenzyme intermediates were subsequently isolated in several reactions of esters with chymotrypsin, and the chemical/kinetic competence of these intermediates was demonstrated.[56] However, similar experiments failed to detect acylenzyme intermediates in the hydrolysis of amide substrates. This failure could be readily explained if the formation of the acylenzyme was rate-limiting, so that the acylenzyme no longer accumulated to detectable levels during amide hydrolysis. The acylenzyme intermediate could be detected in acyltransfer experiments.[93] As shown in Figure 8, the acylenzyme

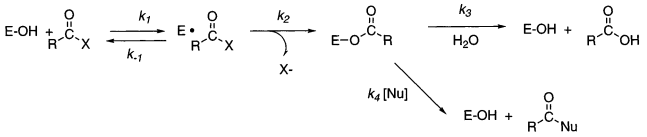


Figure 8. The mechanism of acyltransfer reactions of serine proteases. The acylenzyme normally reacts with water to form a carboxylic acid. However, other nucleophiles can also react with the acyl enzyme. If the added nucleophile is the N-terminus of a peptide, the reaction is the reverse of the hydrolysis reaction and displays the same specificity.

usually reacts with water, but can also react with other nucleophiles. The partitioning of acyltransfer between water and nucleophile will be constant for a given acylenzyme, regardless of how the acylenzyme was formed. Esters and amides have identical partition ratios in acyltransfer reaction, indicating that their reactions involves a common acylenzyme. Note that if the nucleophile is a peptide, the acyltransfer reaction is the reverse of peptide hydrolysis, and must exhibit the same specificity. These reactions have proven useful in mapping the specificity of the S′ sites.[94−99]

Several X-ray crystal structures of acylenzyme intermediates have been solved. The first structure of a peptidyl acylenzyme was a surprise revealed when high-resolution data uncovered a peptide in the active sites of γ-chymotrypsin.[41,100,101] The peptide is

probably Pro−Gly−Ala−Tyr, but may be a mixture derived from autoproteolysis. At low pH, continuous density is observed between Ser195 and the carbonyl of the peptide, indicating that the acylenzyme is formed. Interestingly, the carbonyl group is not completely in the oxyanion hole: the carbonyl oxygen forms a hydrogen bond with Gly193 (2.8 Å), but not with Ser195 (3.5 Å). At high pH, the covalent bond between Ser195 and the peptide is lost, indicating that the product complex has formed. The peptide can be readily replaced with inhibitors, further demonstrating that this acylenzyme is catalytically competent. A peptide was also discovered in the active site of *Streptomyces griseus* proteinase A (SGPA).[42] This peptide appears to be a mixture of autolysis products. As in γ-chymotrypsin, the peptide appears to be an acylenzyme at low pH and a product complex at high pH. Unlike γ-chymotrypsin, the carbonyl oxygen of the acylenzyme occupies the oxyanion hole. More recently, Wilmouth and colleagues screened a peptide library to isolate a stable peptidyl acylenzyme of elastase.[102] β-Casomorphin is a heptapeptide that forms a stable acylenzyme at low pH via the C-terminal carboxylate. The carbonyl oxygen is also found in the oxyanion hole in this acylenzyme. Although initial reports suggested that this carbonyl group was distorted toward a tetrahedral structure, more recent high-resolution data indicate that the ester is planar.[102] The stability of these acylenzymes can be attributed to the protonation of His57 at low pH. All three of these peptidyl acylenzymes display the canonical conformation.

4. Evidence for the Tetrahedral Intermediate

As yet, the tetrahedral intermediate has not been rigorously demonstrated in the serine protease reaction, but is included on the basis of solution chemistry.[103,104] Tetrahedral intermediates are unstable and have very short lifetimes.[78] Further, the lifetime of the intermediate decreases as the leaving group improves. In favorable circumstances, such leaving group effects can provide evidence for tetrahedral intermediates: as the pK_a of the leaving group decreases below that of the nucleophile, formation of the tetrahedral intermediate becomes rate-limiting and a characteristic break will be observed in the Bronsted plot.[105] In the case of the best leaving groups such as *p*-nitrophenol, the lifetime of the tetrahedral intermediate becomes less than a vibration, the reaction becomes concerted and the tetrahedral species must be considered a transition state.[106−108] As in solution, the hydrolysis of *p*-nitrophenylacetate by chymotrypsin is a concerted reaction,[109] and it is formally possible that enzyme-catalyzed peptide hydrolysis is also concerted. Of course, this would require an enormous activation of the amine leaving group. Unfortunately, the leaving group effects measured in Bronsted analysis are greatly reduced by general base catalysis, so that Bronsted analysis of chymotrypsin catalyzed reactions provides no information about the tetrahedral intermediate.[110,111] $^{14}N/^{15}N$ isotope effects have been similarly equivocal.[62]

Not surprisingly, the observation of the tetrahedral intermediate by X-ray crystallography has proven elusive.[112] The first report of a tetrahedral intermediate in the trypsin−BPTI complex could not be substantiated by NMR experiments; further refinement revealed that the scissile peptide bond is distorted toward tetrahedral, but no covalent bond is formed with Ser195.[113,114] Interestingly, this bond is also distorted in the absence of trypsin.[115] Another tetrahedral intermediate has recently been reported in crystals of elastase and the peptide β-casomorphin.[102] When crystals containing the peptidyl acylenzyme are transferred to pH 9.0 and flash-frozen, electron density is observed that suggests a tetrahedral adduct. However, this exciting result must be approached with great caution. X-ray crystallography is a time averaged method, and this tetrahedral adduct may be a mixture of complexes. The presence of sulfate in the cryoprotectant solution is especially troubling because sulfate binds in the active site of many serine proteases and can easily be mistaken for a tetrahedral interemediate (see refs 116 and 117 and references therein). The kinetics of intermediate formation are also not consistent with a pseudo first order process (the tetrahedral intermediate is present at 1 min but completely decomposed after 2 min), which further suggests that the "tetrahedral intermediate" is actually a mixture of complexes. A more promising tetrahedral intermediate appears when crystals the peptidyl acylenzyme of γ-chymotrypsin are soaked in hexane.[118] The electron density around the acyl carbonyl appears to be tetrahedral, and density is observed in the P1′ site. These observations suggest that a peptide from the crystallization mix has attacked the acylenzyme, forming a stable tetrahedral intermediate. The stability of this intermediate is not understood, but might be explained by a shift in equilibrium and/or hydrogen bonding energies in organic solvent. Unfortunately, tetrahedral intermediates have come and gone; the putative tetrahedral intermediates of both elastase and chymotrypsin should be substantiated by additional methods.

B. Mechanisms of Inactivation and Inhibition

A comprehensive discussion of inhibitor/inactivator mechanisms can be found in the Powers paper in this issue. The following will be a brief discussion of compounds that have played an important role in unraveling the mechanism of serine protease catalysis and substrate specificity.

1. Transition State Analogue Inhibitors

Several classes of inhibitors form stable tetrahedral adducts that mimic the tetrahedral intermediate/transition state of the serine protease reaction. The specificity of these reagents can be tailored to a given protease by combining the reactive groups with the appropriate peptidyl portions.[72,119] DFP and the less toxic phenylmethanesulfonylfluoride (PMSF) inactivate virtually all serine proteases and are diagnostic for this protease class. These compounds are very electrophilic, and react irreversibly with Ser195 to form stable tetrahedral adducts (Figure 6). Chloro-

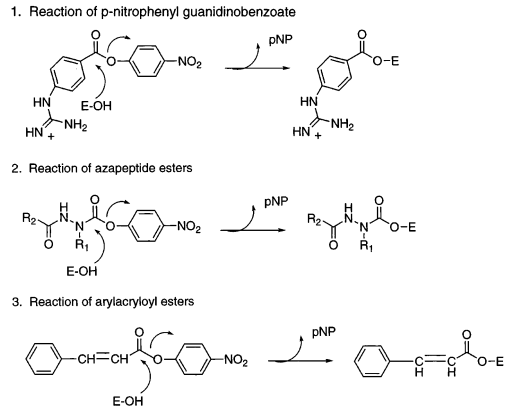


Figure 9. Mechanisms of inhibitors that form stable acylenzymes.

methyl ketones (CMK) are potent inactivators of both serine and cysteine proteases, and therefore are not diagnostic.[120] Chloromethyl ketones cross-link serine proteases by forming a hemiketal with Ser195 and alkylating His57;[121−123] the reaction proceeds via an epoxide intermediate as shown in Figure 6.[124−127] Interestingly, the formation of the epoxide intermediate requires that the oxyanion can escape the oxyanion hole.

Peptidyl aldehydes, trifluoromethyl ketones, boronic acids, and similar compounds form reversible tetrahedral adducts with Ser195. These complexes closely resemble the transition state/tetrahedral intermediate of the hydrolysis reaction: His57 is protonated and the hemiketal oxygen is negatively charged and bound in the oxyanion hole (Figure 6). The affinity of these inhibitors increases as they incorporate the features of good substrates; affinity correlates with the value of k_{cat}/K_m for the hydrolysis of the analogous substrate, as required for a transition state analogue.[47,49,128] Some aldehyde inhibitors also form an alternative adduct where the hemiacetal oxygen interacts with His57 rather than the oxyanion hole.[129−131] Because this adduct mimics the interactions of the leaving group, it can also be considered a transition state analogue. However, peptidyl boronic acids can also form an adduct at Nϵ2 of His57 that cannot be considered a transition state analog.[132] These Type 2 adducts are formed when inhibitors resemble poor substrates or in complexes of protease zymogens,[133] forming higher affinity complexes than would be expected based on the hydrolysis of analogous substrates.

2. Stable Acylenzymes

Many compounds inactivate serine proteases by forming stable acylenzyme intermediates. This stability derives from three general mechanisms:

(a) The intrinsic reactivity of the acyl group. The reactivity of esters can be decreased by increasing the electron density of the carbonyl. For example, the deacylation rates of benzoyl chymotrypsins decrease as the substituent becomes more electron-donating.[134] Similarly, alpha heteroatoms also deactivate esters, and acylchymotrypsins derived from azapeptides are stable.[135] Note that azapeptidyl acylenzymes are isosteric with peptidyl acylenzymes, and thus are expected to closely mimic the normal catalytic intermediates.[136] Of course, these effects also decrease the acylation rate constants, making it more difficult to form the acylenzyme. However, the effects on acylation can be offset with good leaving groups such as *p*-nitrophenol. Note that such reagents as *p*-nitrophenyl-guanidinobenzoate are active site titrants, reacting with trypsin to form a stable acylenzyme intermediate and one equivalent of *p*-nitrophenol (Figure 9).[137] Electron-withdrawing properties can also be disguised by more clever means and many classes of mechanism-based inactivators exploit this strategy (see Powers paper in this issue).

(b) The acylenzyme does not interact with the oxyanion hole. This mechanism is illustrated by arylacryloyl chymotrypsin, where the planarity of the arylacryloyl system does not allow the carbonyl to enter the oxyanion hole.[71,138−141] A similar mechanism may account for the stability of acetyl-chymotrypsin; this adduct lacks a large hydrophobic group that can bind in the S1 site and direct the carbonyl oxygen into the oxyanion hole.

(c) The catalytic triad is disrupted and cannot activate water. The simplest manifestation of this mechanism is the protonation of His57 so that it cannot serve as a general base catalyst as discussed above. A second example is the inactivation of chymotrypsin by 3-benzyl-6-chloro-2-pyrone.[142] This compound forms a stable acylenzyme where Ser195 attaches to C1 of (*Z*)-2-benzylpentenedioic acid. His57 forms a salt bridge with the free carboxylate of the inactivator.[143] This salt bridge prevents His57 from performing its role as the general base in the deacylation reaction. Nature has also used this strategy to inhibit serine proteases. Serpins form stable acylenzymes; this stability results from distortion of the active site upon complex formation[144] (see the Gettins paper in this issue).

C. Issues in the Function of the Catalytic Triad

1. Where Are the Protons? One-Proton versus Two-Proton-Transfer Mechanisms

Although the general features of the charge relay system are widely accepted, the details of its function have been surprisingly controversial.[21] The issue of whether the proton in the Asp102−His57 hydrogen bond resides on His57 or Asp102 has been particularly contentious. Blow's original conception of the charge relay system explained the reactivity of Ser195 in terms of resonance forms Asp-CO_2^-/H-His/Ser-OH and Asp-CO_2H/His-H/Ser-O^-.[1,67] This mechanism was later refined to involve actual proton transfers from Ser195 to His57 and His57 to Asp102, resulting in the tetrahedral intermediate and neutral Asp102 and His57.[145,146] This two-proton-transfer mechanism was controversial from the outset because it requires that the pK_a of His57 is *lower* than Asp102, i.e., the pK_a of His57 must decrease in the tetrahedral intermediate/transition state. Nevertheless, this mechanism was incorporated into many textbooks.

^{1}H NMR experiments strongly favor the one-proton mechanism. The Nδ1 proton of His57 has an unusu-

ally low field ^{1}H NMR signal (13.8−15 and 17−18 ppm in neutral and protonated His57, respectively) which can be used to determine the pK_a of His57 in various complexes.[147,148] The pK_a of His57 is ∼7 in the free enzyme, acylenzyme, and protein inhibitor complexes but *increases* to greater than 10 in transition state analogue complexes.[128,149−151] Neutron diffraction experiments also demonstrated that the proton is associated with His57 in monoisopropyl phosphate modified trypsin.[152] Although these experiments do not establish the position of the proton in the transition state of the catalytic reaction, they argue strongly for the one-proton mechanism, where protonated His57 is stabilized by an electrostatic interaction with a negatively charged Asp102.

2. Low Barrier Hydrogen Bonds

A more subtle version of the two-proton-transfer mechanism has emerged in the low barrier hydrogen bond (LBHB, also known as short, strong hydrogen bond) hypothesis of recent years.[153−159] This controversial theory developed from the observation that unusual hydrogen bonds form in the gas phase, characterized by short length (<2.6 Å) and extraordinary strength (30 kcal/mol).[160,161] The unusual properties of these hydrogen bonds is believed to derive from the short distance between donor and acceptor atoms with matched pK_a's, which eliminates the barrier for proton transfer so that the proton is shared in a broad single-potential well. The hydrogen bond between Asp102 and His57 appears to fulfill many of the criteria for a LBHB:[162−164] (a) The distance between Asp102 and His57 is <2.6 Å. (b) The ^{1}H NMR signal of the Nδ1 proton of His57 is unusually downfield.[21,147] (c) A deuterium isotope effect is observed on chemical shift of the His57 Nδ1 proton. (d) The D/H fractionation factor of the Nδ1 proton is low.[150,165] Unfortunately, these observations are not diagnostic for LBHBs, and the Asp−His hydrogen bond fails by other measures: (a) The pK_a's of Asp102 and His57 are not matched. (b) both ^{15}N−H spin couplings and ^{15}N chemical shifts indicate that the proton is ∼85% localized on His57.[158,166] (c) The His57−Nδ1 proton is not shielded from water and the environment is polar.[158] Moreover, both Asp102 and His57 are involved in multiple hydrogen bonding interactions; these interactions will compete with and weaken the Oδ1−Nδ1−H hydrogen bond. Further, disruption of the catalytic triad decreases the catalytic power of serine proteases by no more than ∼6 kcal/mol, which must be distributed among several hydrogen bonds in addition to the Asp102−His57 interaction. This energy loss is consistent with a simple electrostatic interaction.[167] Last, mutations at Asp102 can be rescued by replacing Ser214 with Asp.[168] Although this altered catalytic triad has not been thoroughly characterized, it would be surprising if a LBHB could be so easily reconstituted. Similarly, substitution of the catalytic triad Asp32 with Cys in subtilisin eliminates the putative LBHB, but decreases k_{cat}/K_m by no more than 50-fold.[169] These experiments suggest that if a LBHB exists, it contributes no more than 2 kcal/mol to the catalytic power of serine proteases.

The interactions of the oxyanion with the oxyanion hole have also been suggested to involve LBHBs.[153] However, the two hydrogen bonds of the oxyanion hole provide a total of ∼6.5 kcal/mol in transition state stabilization,[87,89,90,149,170] which is consistent with simple electrostatic interactions.[171]

3. Moving His Mechanisms

In the generally accepted mechanism for serine protease catalysis (Figure 5), His57 removes a proton from Ser195 and transfers it to the leaving group. This proton transfer has troubled some investigators, who argue that if His57 abstracts a proton from Ser195, then His57-H$^+$ will be positioned near Ser195 in the tetrahedral intermediate, which would favor re-protonation of Ser195 and regeneration of substrate.[172,173] For the reaction to proceed, His57-H$^+$ must be positioned to protonate the leaving group. The "His flip" mechanism of Figure 10 has been

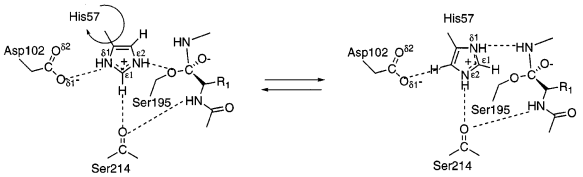


Figure 10. The "His flip" mechanism. After initial formation of the tetrahedral intermediate, the Nϵ1-H will be within hydrogen bonding distance of the γO of Ser195. Rotation of His57 supposedly places the Nϵ1-H near the leaving group NH. Although the hydrogen bonds between Nδ1-H and Asp102 and Cϵ1 and Ser214 would be disrupted, new hydrogen bonds will form to replace them.[21]

proposed to solve this problem:[27] after formation of the tetrahedral intermediate, His57-H$^+$ flips, which should place the Nδ1 proton near the leaving group. The flipped conformation of His57 has been observed in subtilisin in 50% dimethylformamide.[174] In addition, the presence of the Ser214−His57 hydrogen bond and the unusual dynamics of His57 have been taken as evidence for the "His flip" mechanism.[21]

Several considerations argue against the "His flip" mechanism. First, the disruption and reformation of so many hydrogen bonds in the short lifetime of the tetrahedral intermediate seems unlikely. The His flip mechanism appears to violate the principle of least motion, which suggests that enzymatic reactions occur with an economy of movement.[175] Moreover, His57 is bracketed by the P2 and P1′ residues of a peptide substrate, which creates a steric barrier to flipping (Figure 7). The unusual dynamics of His57 can be explained by substrate and solvent moving in and out of the active site, or even the chemical transformation—there is no need to invoke flipping. Most seriously, while the flipping mechanism looks reasonable in two dimensions, it fails in three dimensions. Analysis of peptide trifluoroketone or boronic acid complexes shows that the flip actually places proton donor farther away from the leaving group by 0.8 Å.

Of course, some movement of both protease and substrate is inevitable as the hydrolysis reaction proceeds. Ser195 must move at least 1 Å to form the tetrahedral intermediate and smaller movements of

the substrate will be required as the bonding changes from sp^2 to sp^3. Therefore, His57 may not be oriented toward Ser195 in the tetrahedral intermediate as assumed. Even if His57 is positioned so that Ser195 is protonated in the tetrahedral intermediate, the reaction is not doomed to collapse back to starting materials. The proton is less than 2.5 Å from the amine leaving group, which has the higher pK_a. Protonation of the amine is favored overall, and the proton needs only to transfer to the amine before the tetrahedral intermediate breaks down. In addition, several structures of transition state analogue complexes indicate that His57-H^+ can interact with the leaving group position while maintaining the Asp102−His57 hydrogen bond (ref 74 and references therein). Whatever reorientation that may be necessary for completion of the catalytic cycle can probably be accomplished with small adjustments in both protease and substrate and without breaking the Asp102−His57 hydrogen bond.

4. The Hydrolytic Water

The mechanism of serine protease catalysis places some rather rigorous constraints on the hydrolytic water: it should approach acylenzyme from the leaving group side of the active site, it must make a hydrogen bond to His57 and the angle of attack must approximate 109 degrees. Several crystal structures of serine proteases contain ordered water molecules that appear to fulfill one or more of these criteria. Henderson noted the first such water in a cinnamoyl enzyme complex of chymotrypsin, and similarly located waters have appeared in subsequent acylenzyme structures.[41,71] This water is 2.4 Å above the carbonyl of the acylenzyme and also interacts with His57 as would be expected of the hydrolytic water, but the angle of attack on the carbonyl is not favorable. In addition, a similarly placed water molecule is observed in product and transition state analogue complexes.[176] The latter complexes mimic the tetrahedral intermediate, i.e., the intermediate formed *after* the water has attacked; therefore, this water is more likely to play a structural role. Another candidate water molecule is observed in a benzoylenzyme complex of trypsin, poised 3.9 Å above the carbonyl and 3.4 Å from His57.[177,178] However, this position is normally occupied the P2−P3 residues of a peptide substrate, which suggests that this water is not involved in the hydrolysis of peptidyl acylenzyme. A better candidate for the hydrolytic water has been identified in peptidyl acylenzymes of SGPA and elastase.[39,42] This water molecule is found ∼3 Å from His57 and ∼3 Å from the carbonyl of the acylenzyme, at an optimal angle for nucleophilic attack from the leaving group side of the active site. However, as promising as this candidate is, it must be remembered that these complexes are unusually stable acylenzymes at a pH where catalysis does not occur readily; the actual hydrolytic water may not present.

5. Stereochemistry of the Tetrahedral Intermediate

The stereoelectronic requirements of formation and breakdown of the tetrahedral intermediate pose an additional conundrum.[62,173] When a nucleophile attacks an amide bond, the lone electron pairs of the oxyanion and the nitrogen of the leaving group must be antiperiplanar to the new bond (Figure 11). In a serine protease reaction, this stereochemistry leaves the lone electron pair of the amine leaving group pointing away from His57-H^+. The nitrogen must undergo an inversion to position the lone pair for protonation. Although it is unclear how the protease will influence the inversion process, such inversions occur readily in solution. Perhaps the enzyme can use interactions with the oxyanion hole and S′ sites to overcome the stereoelectronic imperative of solution chemistry.[62]

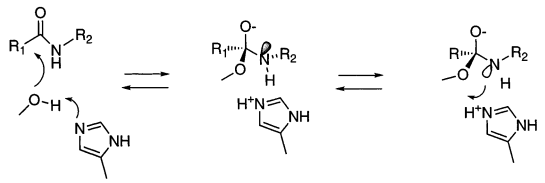


Figure 11. Stereoelectronic mechanism for serine protease catalysis.

6. Stabilization of the Tetrahedral Intermediate/Strain in the Acylenzyme Intermediate

Enzyme catalysis is most simply described in terms of transition state stabilization, but in practice interactions that stabilize the transition state will also strain the ground state. This idea can be demonstrated by considering the oxyanion hole, where interactions should both stabilize the tetrahedral intermediate and activate the substrate for nucleophilic attack.

Stabilization of tetrahedral intermediates is evident in the characterization of transition state analogue complexes (Figure 6). The pK_a of the hemiketal oxygen in chloromethyl ketone-inactivated chymotrypsin is ∼2 units lower than the pK_a in solution,[170] while the pK_a of the hemiketal oxygen of trifluoroketone adduct is >5 units lower.[149] These results suggest that the oxyanion hole interactions stabilize the tetrahedral intermediate by ∼6 kcal/mol.

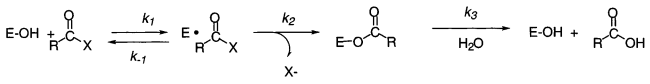
Unfortunately, distortion of the peptide substrate has been difficult to demonstrate. However, convincing evidence for such distortion/activation of the acylenzyme comes from resonance Raman spectroscopy. Tonge and Carey measured the carbonyl stretching frequencies of a series of arylacryloyl acylenzymes of chymotrypsin and subtilisin.[140,179−181] Recent work extends these observations to *p*-(dimethylamino)-benzoylenzymes.[182] Multiple features are observed in the carbonyl profile of these acylenzymes, suggesting that the carbonyl group has multiple conformations. Because hydrogen bonding decreases the stretching frequency of carbonyl bond, the lowest frequency band was assigned to the conformation of the acylenzyme with the carbonyl bound in the oxyanion hole. This assignment is substantiated by the following observations: (1) The intensity of this band increases with pH, displaying the same pH dependence as the deacylation rate constant.[140] (2) This band is absent in subtilisin variants where the oxyanion hole has been eliminated.[179] Importantly, the frequency of this

band ($\nu_{C=O}$) correlates with deacylation rate for a series of acylenzymes. This correlation spans 4 orders of magnitude in deacylation rate constants ($6.8 \times 10^{-6}-0.3\ s^{-1}$).[180,181] Further, $\nu_{C=O}$ can be related to carbonyl bond length, indicating that carbonyl bond length increases 0.025 Å as the deacylation rate constant increases 16300-fold. This increase is ~11% of the change incurred upon transformation of a carbonyl group to a C−O bond. If this correlation is extrapolated to the deacylation rates observed for good substrates (~100 s^{-1}), the carbonyl bond is distorted ~14% toward a C−O bond. Last, model systems indicate that the change in ΔH associated with the shifts in $\nu_{C=O}$ is 6.4 kcal/mol, similar to the changes in transition state stabilization of 5.7 kcal/mol over the range of compounds. There is sufficient energy in the hydrogen bonding interactions of the oxyanion hole to provide this transition state stabilization. Similarly, multiple ester bands are also observed in Fourier transform infrared spectroscopy of cinnamoyl-chymotrypsins, and have likewise been attributed to multiple conformations of the carbonyl group.[183] This work provides further evidence for distortion of the carbonyl bond in the acylenzyme. Interestingly, the ester carbonyl frequency correlates with dielectric constant in model compounds. This correlation suggests that the carbonyl of the acylenzyme experiences a dielectric constant of 70 in the oxyanion hole conformation and 40 in the nonproductive conformation.

It appears that the function of the catalytic triad and oxyanion hole can be explained by the electrostatic complementarity of the active site and transition state.[167] The preorganization of the enzyme active site provides a substantial advantage relative to the uncatalyzed reaction where solvent molecules must organize to stabilize the developing charges as the reaction proceeds.[262] Warshel has described this phenomenon as solvation substitution rather than desolvation.[167] This preorganization is paid for by the folding energy of the enzyme.[263]

IV. Kinetics of the Serine Protease Reaction

The serine protease reaction is generally considered to have three-step kinetic mechanism (shown in Figure 12): (a) formation of an enzyme−substrate (E·S) complex ($K_s = k_{-1}/k_1$); (b) acylation of the active site serine (k_2); and (c) hydrolysis of the acylenzyme



$$k_{cat} = k_2k_3/(k_2 + k_3)$$

$$K_m = k_3(k_{-1} + k_2)/k_1(k_2 + k_3)$$

$$k_{cat}/K_m = k_1k_2/(k_{-1} + k_2)$$

If $k_{-1} >> k_2$,

$$K_m = k_{-1}k_3/k_1(k_2 + k_3) = K_S k_3/(k_2 + k_3)$$

$$k_{cat}/K_m = k_1k_2/k_{-1} = k_2/K_S$$

If $k_3 >> k_2$,

$$K_m = (k_{-1} + k_2)/k_1$$

Figure 12. The kinetic mechanism of serine proteases.

intermediate (k_3). Although the three-step mechanism of Figure 12 is widely accepted, it is important to note that some of its key features have not been rigorously tested. First, both k_2 and k_3 are composite rate constants that include a chemical transformation and a product dissociation step. It is generally assumed that the product dissociation step is fast relative to the chemical transformation. The evidence usually cited to support this assumption is the presence of a large solvent kinetic isotope effect (~3) on k_{cat}, which is consistent with rate-limiting chemical transformation.[184] However, solvent isotope effects can be difficult to interpret because the solubility of substrates and products, the viscosity of the solution and even the structure of the protein can change in D_2O, so that the isotope effect may not be definitive proof of rate-limiting chemistry.[185] Second, the mechanism indicates that dissociation of the leaving group must occur prior to hydrolysis of the acylenzyme. Again, there is little evidence on this point. Acyltransfer experiments indicate that hydrolysis can occur from an acylenzyme−nucleophile complex, which appears to contradict this assumption.[97,186−188] However, these acylenzyme−nucleophile complexes may be nonproductive complexes that do not form during the normal catalytic cycle. For example, during the hydrolysis of a peptide substrate, the leaving group would be constrained to interact at the S1′ site. In contrast, an added nucleophile could form a nonproductive complex at another S′ site that would not impede hydrolysis of the acylenzyme. Thus, the order of hydrolysis and product release is not well established.

Several other assumptions are frequently made when evaluating the kinetics of serine protease reactions. The Michaelis−Menten parameters k_{cat}, K_m, and k_{cat}/K_m are composites of the rate constants as shown in Figure 12. In particular, the expression for K_m includes k_3 as well as k_1, k_{-1}, and k_2. Increasing values of k_2 will increase K_m until $k_2 \geq k_3$, but further increasing k_2 will decrease K_m (Figure 12), so that K_m can be exceedingly difficult to interpret with respect to substrate affinity. In addition, the expressions in Figure 12 are frequently simplified by assuming $k_{-1} \gg k_2$. This assumption is probably valid for poor substrates, but the hydrolysis of good substrates is often diffusion controlled, which indicates $k_2 \geq k_{-1}$.[189−191]

The assertion that acylation is rate-limiting for amide hydrolysis, i.e., $k_3 \gg k_2$, is by far the most indiscriminately applied assumption. Seminal work from Bender and co-workers demonstrated that k_2 is the rate-limiting step for the hydrolysis of *N*-acetyl-L-tryptophanamide, while k_3 is rate limiting for the hydrolysis of *N*-acetyl-L-tryptophan esters.[59] These observations have been generalized to statements such as "acylation is rate determining for amide substrates and deacylation is rate determining for ester substrates of serine proteases", as may be found in many textbooks. This statement has gained wide acceptance because amides are intrinsically much less reactive than esters. However, *N*-acetyl-L-tryptophanamide is a very poor substrate for chymotrypsin, with $k_{cat}/K_m = 3.6\ M^{-1}\ s^{-1}$ and $k_{cat} = 0.026\ s^{-1}$.

Enzymes often process poor substrates with different rate-determining steps than good substrates. Chymotrypsin and trypsin hydrolyze peptides with values of k_{cat}/K_m in the range of $\sim 10^7$ M^{-1} s^{-1} and $k_{cat} \approx 100$ s^{-1}.[192,193] Michaelis−Menten parameters of this order of magnitude are typical for the reactions of serine proteases; thrombin and tissue plasminogen activator hydrolyze their natural substrates with k_{cat}/K_m's $\approx 10^7$ M^{-1} s^{-1}.[194,195] Even Factor D, a protease with notoriously low activity toward oligopeptide amide and ester substrates, cleaves its natural protein substrate with $k_{cat}/K_m \geq 2 \times 10^6$ M^{-1} s^{-1}.[196,197] Given that the values of k_{cat}/K_m for peptide hydrolysis can be 10^6-fold greater than the amide substrates used in Bender's work, the acylation step should not be assumed to be generally rate-limiting.

Indeed, several experiments suggest that k_2 is not always rate-limiting for amide hydrolysis. Comparable values of k_{cat} have been reported for the hydrolysis of analogous oligopeptide amides and esters.[192,198−201] Such observations generally suggest that a common step is rate-limiting, i.e., hydrolysis of the acylenzyme intermediate. However, inferring individual rate constants from steady-state parameters is a precarious practice. The most convincing case for rate-limiting deacylation in amide hydrolysis is the observation of bursts of product formation under pre-steady state conditions. Such bursts are observed in the hydrolysis *p*-nitroanilidamides by human leukocyte elastase and in the hydrolysis of methylcoumarinamides by Kex2 protease and furin.[202−204] Similar bursts have been observed in the hydrolysis of peptide methylcoumarinamides by trypsin (X. Liu and L. Hedstrom, unpublished experiments). Most convincingly, a burst is observed in the hydrolysis of a peptide substrate by Kex2.[205] Thus, the generalization "k_2 is rate determining for the hydrolysis of amide substrates by serine proteases" is not warranted.

The above work is frequently criticized because the *p*-nitroanilides and methylcoumarinamides leaving groups are more reactive than a peptide bond. However, amide hydrolysis has little dependence on leaving group ($\beta_{lg} = 0.07$) and the plot of the values of k_{OH} for the hydrolysis of both amides and anilides fall on a single line when plotted against pK_a of the leaving group.[206] While *p*-nitroanilides are exceptionally reactive in solution,[207] this unusual reactivity is not recapitulated in protease reactions. The hydrolysis of anilides, including *p*-nitroanilides, by chymotrypsin displays no dependence on the pK_a of the leaving group (most likely due to general base catalysis). Moreover, peptide bonds are *more* reactive than anilides with leaving groups of comparable pK_a.[110] Therefore, it is unlikely that the results cited above can be attributed to the unusual reactivity of the leaving groups. These observations further demonstrate that the rate-determining step of an amide substrate should not be automatically assigned to acylation.

Several methods have been used to measure the values of the individual rate constants k_1, k_{-1}, k_2, and k_3. Surprisingly, although pre-steady state and single turnover experiments are the best methods for determining individual rate constants, they have only rarely been applied to serine proteases, and usually confined to the analysis of ester hydrolysis (for examples, see refs 208−211). A recent study has applied these methods to the analysis of the hydrolysis of peptide *p*-nitroanilides by elastase.[202] The individual rate constants are more commonly determined by characterizing the hydrolysis of analogous amide/ester pairs.[59,199,200] Deacylation is usually rate-limiting in the hydrolysis of esters, so that $k_{cat,ester} = k_3$. While this fact is often taken for granted, it can be substantiated if the reactions of a set of analogous esters with different leaving groups have the same k_{cat}; this observation suggests that a common step is rate-limiting, i.e., deacylation. Likewise, an analogous amide substrate must form the same acylenzyme, so that k_3 must be the same as in the ester reaction. The values of K_s and k_2 can then be determined from the equations in Figure 12 (with the assumption that $k_{-1} \gg k_2$). The values of K_s, k_2, and k_3 can also be determined using acyltransfer reactions under conditions where the added nucleophile does not form a complex with the acylenzyme.[94,212] This method is especially useful in evaluating the hydrolysis of substrates where $k_2 \geq k_3$. The analysis of the viscosity dependence of k_{cat}/K_m can yield values of k_1 and the ratio of k_{-1}/k_2.[189] This method must be applied with caution because it involves changing solvent, with attendent changes in dielectric constant and the possibility of other nonspecific effects, and therefore requires careful attention to controls. Individual rate constants have also been determined by analyzing the temperature dependence of the Michaelis−Menten parameters.[191] This method cannot identify the rate-limiting step and generally operates under the assumption that $k_3 \gg k_2$. Using these various methods, values of k_2 of 10^2-10^4 s^{-1} for good ester substrates and 10−1000 s^{-1} for good amide substrates have been reported. The value of k_3 is usually ~ 50 s^{-1} for good substrates. Obviously, the values of k_2 and k_3 can be much lower if the substrate is not optimal. The values of k_1 and k_{-1} are typically 10^6-10^8 M^{-1} s^{-1} and 5−500 s^{-1}, respectively.

V. Substrate Discrimination during Catalysis

A. The Hallmarks of Serine Protease Specificity

Specificity is defined by how substrates compete for an enzyme, and is measured by the value of k_{cat}/K_m. The P1/S1 interaction controls specificity, as illustrated with trypsin, where the value of k_{cat}/K_m varies over 10^5-fold as the P1/S1 interaction is optimized (Table 3). This increase results mainly from an increase in k_{cat} (10^4-fold) rather than K_m (10-fold). Interactions at the S2−Sn sites also strongly influence specificity, as illustrated with elastase where the value of k_{cat}/K_m increases 10^5-fold as the additional P residues are added (Table 3). Again, the increase results mainly from an increase in k_{cat} (10^3-fold) rather than K_m (100-fold). Similarly, P′/S′ interactions increase k_{cat}/K_m, once again via an increase in k_{cat} (Table 3). The effect of the P′/S′ interactions can be most dramatically illustrated by the inactivation of chymotrypsin by carbonate esters, where

Table 3. The Hydrolysis of Substrates by Rat Trypsin and Porcine Pancreatic Elastase

substrate	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_m (μM)
	Rat Trypsin[b]		
Suc-Ala-Ala-Pro-Arg−AMC	1.3×10^6	91	70
Suc-Ala-Ala-Pro-Lys-AMC	1.2×10^6	120	100
Suc-Ala-Ala-Pro-Tyr−AMC	3	6.3×10^{-4}	200
Suc-Ala-Ala-Pro-Phe−AMC	8	6.3×10^{-3}	780
	Porcine Pancreatic Elastase		
Ac-Ala−NH_2[c]	<0.05	<0.008	160 (K_i)
Ac-Pro-Ala−NH_2[c]	0.07	0.007	100
Ac-Ala-Pro-Ala−NH_2[d]	25	0.1	4.2
Ac-Pro-Ala-Pro-Ala−NH_2[d]	3700	9.3	2.5
Ac-Ala-Pro-Ala−Ala-NH_2[d]	520	2.4	4.6
Ac-Pro-Ala-Pro-Ala−Ala-NH_2[d]	25,500	37	1.4
	Human Leukocyte Elastase[e]		
Cbz-Val−pNP	480		
Cbz-Phe−pNP	65		
Boc-Ala-Ala-Val−SBzl	3.0×10^6		
Boc-Ala-Ala-Phe−SBzl	<10		

[a] "−" denotes cleaved bond; AMC, aminomethylcoumaride; pNP, *p*-nitrophenyl; SBzl, thiobenzyl. [b] Data from ref 218. [c] Data from ref 260. [d] Data from ref 261. [e] Data from ref 215.

Table 4. Inhibition of Chymotrypsin by Carbonate Esters[a]

leaving group X	k_{on} (M^{-1} s^{-1})
OCH_3	9
O-Ala-OCH_2CH_3	200
O-Ala-Leu-OCH_3	2,400
O-Ala-Leu-Arg-OCH_3	>23,000

[a] Data from ref 213.

interactions with S1′−S3′ sites increase the rate of inactivation >2000-fold (Table 4).[213]

Importantly, the S1 site and the remote binding sites act in concert to determine serine protease specificity.[200,214,215] While the value of k_{cat}/K_m increases dramatically with peptide length, this increase requires the correct P1−S1 interaction.[52,216] This phenomenon can be illustrated with the hydrolysis of *p*-nitrophenylesters by human leukocyte elastase (HLE) (Table 3).[215] HLE displays little specificity in the hydrolysis of single residue esters, with the P1−Val substrate preferred by only 7-fold over the P1−Phe, even though the S1 site of HLE should not accommodate large residues. The addition of two peptide residues increases the hydrolysis of the P1−Val substrate by $>10^3$, but *decreases* the hydrolysis of the Phe-containing substrate. Likewise, increasing peptide length dramatically increases the value of k_{cat}/K_m in the trypsin-catalyzed hydrolysis of substrates containing P1−Arg or Lys, but not P1−Phe.[200] Similar results are observed for other serine proteases.

B. Specificity Is Determined by the Binding and Acylation Steps

As noted above, k_{cat}/K_m depends only on the binding and acylation steps k_1, k_{-1}, and k_2 (Figure 12). From an intuitive perspective, if two substrates compete for protease, the competition is "won" when acylation occurs and deacylation is irrelevant. Since the increase in the values of k_{cat}/K_m is derived from an increase in k_{cat} rather than a decrease in K_m, specificity must be primarily determined by acylation. This conclusion is borne out when the steady-state kinetic parameters are deconvoluted into K_s and k_2 (Table 5). The value of k_2 varies by more than $10^3\times$ between good and poor substrates while only 10−100-fold of substrate discrimination derives from substrate affinity.[52,200,202,216] It is especially noteworthy that long peptide substrates bind with approximately the same affinity as short substrates.

Table 5. The Specificity of Binding, Acylation, and Deacylation[a]

substrate	k_{cat}/K_m (M^{-1} s^{-1})	K_s (μM)	k_2 (s^{-1})	k_3 (s^{-1})
	Human Leukocyte Elastase[b]			
MeOSuc-Val−pNA	75	788	0.06	10
MeOSuc-Pro-Val−pNA	575	722	0.42	11
MeOSuc-Ala-Pro-Val−pNA	56,000	810	43	13
	Rat Trypsin II			
Z-Arg−AMC[c]	4000	96	0.7	65
Tos-Gly-Pro-Arg−AMC[c]	1×10^7	≥50	≥1300	56
Z-Lys−AMC[c]	1500	220	0.32	130
D-Val-Leu-Lys−AMC[c]	30,000	6500	280	18
Ac-Phe-NH_2[d]	0.2	1000	0.0002	30
Suc-Ala-Ala-Pro-Phe−AMC[d]	6	1000	0.007	60

[a] Mechanistic rate constants as defined in Figure 12. "−" denotes cleaved bond; pNA, *p*-nitroanilide; AMC, aminomethylcoumaride. [b] Data from ref 52. [c] Data from ref 243. [d] Data from ref 200.

Although k_3 is not a component of k_{cat}/K_m, it should not be completely ignored. If the value of k_3 is very small, the acylenzyme will be very stable, substrate turnover will be very slow and the enzyme will be essentially inactivated. The deacylation step is therefore a critical determinant of substrate turnover. The deacylation step also varies among substrates, although much more dramatic changes in the substrate are required to affect k_3. Interestingly, while the values of k_2 for P1−Arg containing substrates are 100−10^5-fold great than P1−Phe, the values of K_s are <10-fold less and the values of k_3 are similar (Table 5). Thus, the preferences of the binding, acylation and

deacylation steps are different even though these steps occur on the same enzyme surface.

C. The Contribution of Substrate Association to Specificity

Inspection of the algebraic expression for k_{cat}/K_m reveals another important mechanistic fact: as the value of k_2 increases, k_{-1} will become insignificant, $k_{cat}/K_m = k_1$ and the reaction will be diffusion controlled. If the value of k_1 was determined by the simple collision of substrate and protease, then k_1 would be similar for all substrates, i.e., $\sim 10^8$ M^{-1} s^{-1}. However, values of k_1 of 10^6-10^7 M^{-1} s^{-1} have been reported, which suggest that substrate association is more complicated than a simple collision. Obviously, the substrate/protease interface has many hydrogen bonding and van der Waals interactions which seem unlikely to form in a single step. Some evidence for isomerization steps has been reported.[52,56] One might expect that the value of k_1 should decrease with peptide length since more interactions must form. The value of k_{-1} should also decrease, since more interactions would have to be disrupted when the substrate dissociates. If the values of k_1 and k_{-1} decrease in parallel with peptide length, then the overall affinity will remain constant. This idea can be illustrated with the inhibition of chymotrypsin by trifluoroketones. Ac-Leu-Phe-TFK is a more potent inhibitor of chymotrypsin than Ac-Phe-TFK ($K_i = 0.5$ and 4.5 nM, respectively; values are corrected for free ketone concentration).[49] Ac-Leu-Phe-TFK has a lower value of k_1 than Ac-Phe-TFK (4×10^6 and 3×10^7 M^{-1} s^{-1}, respectively), and a lower value of k_{-1} (0.002 and 0.13 s^{-1}, respectively). The structure of the chymotrypsin does not change appreciably in the two inhibitor complexes, suggesting the kinetics of Ac-Leu-Phe-TFK inhibition are slower because more enzyme−inhibitor contacts must form.[74]

When a reaction is diffusion controlled, the conformation and solvation of the substrate and protease will determine the rate of substrate association, and hence of specificity. These considerations illustrate why the structures of both free enzyme *and free substrate* must be known to assess the basis of specificity. Unfortunately, the structure of the substrate in solution is rarely considered in any enzymatic reaction. Several puzzling observations about serine protease specificity might be explained by the solution conformation of peptide substrates. For example, the P3 residue appears to be a major determinant of specificity even when it makes few contacts with the enzyme. Large P3 residues may favor a more extended peptide conformation that would favor association. The surprising observation that the substitution of ester bonds at the P4−P3 and P3−P2 amide bonds has little effect on the value of k_{cat}/K_m in trypsin reactions[46] might also be explained by changes in solution conformation and/or solvation.

D. Implications for the Specificity of Ester Hydrolysis

The specificity of ester hydrolysis may be dominated by substrate association. Once bound in the active site cleft, the intrinsic reactivity of an ester ensures acylation and progression through the catalytic cycle. This condition almost certainly holds for *p*-nitrophenylesters, where acylation is so facile that the reaction is concerted. Thus, the specificity of ester hydrolysis will underestimate the potential of a protease to discriminate between peptide substrates. The validity of this idea is illustrated in the hydrolysis of "inverse substrates". Trypsin hydrolyzes *p*-amidinophenyl acetate, where the leaving group binds in the S1 site, almost as well as normal ester substrates ($k_{cat}/K_m = 10^5$ versus 10^7 M^{-1} s^{-1}) (Figure 13).[217] Similarly, the facile reaction of *p*-nitrophenyl guanidinobenzoate with chymotrypsin can also be explained by an inverse orientation, where the *p*-nitrophenyl leaving group binds in the S1 site and the guanidino group interacts with S1′.

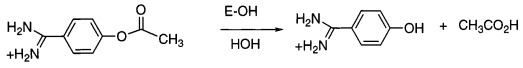


Figure 13. Inverse substrates of trypsin.

When the reaction is diffusion controlled, the rate-limiting transition state will not involve formation of the tetrahedral intermediate as generally supposed; instead the rate-limiting step will be formation of the E·S complex. The general features of this transition state might include the canonical substrate conformation as well as interactions at the various subsites. Many investigators have assumed that protein inhibitors, which are locked in the canonical conformation, form complexes that mimic ground-state E·S complexes. However, the values of K_i for protein inhibitors correlate with values of k_{cat}/K_m for the analogous substrates, a feature of transition state analogues but not ground-state inhibitors.[38] The canonical conformation may be a high energy conformation that activates the scissile bond and thus may be the transition state-like feature of protein inhibitors.

VI. Redesigning the Trypsin–Chymotrypsin–Elastase Paradigm

Textbooks often illustrate enzyme specificity with the pancreatic serine proteases trypsin, chymotrypsin, and elastase: trypsin, with Asp189, cleaves after P1 Arg/Lys residues; chymotrypsin, with Ser189, cleaves after P1 Phe/Tyr/Trp; and elastase, with Val216 and Thr226, cleaves after P1 Ala/Val. The trypsin−chymotrypsin−elastase paradigm suggests that protease specificity is determined by a few structural elements. However, the obvious mutations fail to transfer the specificity of one protease into another.

A. Redesigning Amidase Activity

Switching these elements from one framework to the next should be sufficient to change specificity. However, the substitution of Asp189 with Ser does not convert trypsin into protease with specificity for P1−Phe (Table 6).[200,218,219] Trypsin does acquire chymotrypsin-like specificity when the residues of the S1 site and two surface loops are replaced with their

Table 6. Mechanistic Kinetic Parameters for Amide Hydrolysis by Chymotrypsin, Trypsin, and Trypsin Mutants[a]

substrate	k_{cat}/K_m (M^{-1} s^{-1})	K_s (M)	k_2 (s^{-1})	k_3 (s^{-1})
	Chymotrypsin			
Ac-Phe-NH_2	18	0.023	0.43	60
Suc-AlaAlaProPhe-AMC	4.9×10^5	2.5×10^{-4}	12	52
Suc-AlaAlaProPhe-pNA	5.6×10^5	1.5×10^3	850	52
	Trypsin			
Ac-Phe-NH_2	0.2	1×10^{-3}	2×10^{-4}	30
Suc-AlaAlaProPhe-AMC	6.0	1.1×10^{-3}	0.007	36
Suc-AlaAlaProPhe-pNA	9	≥ 0.25	≥ 0.2	36
	D189S			
Ac-Phe-NH_2	0.1	0.17	0.018	39
Suc-AlaAlaProPhe-AMC	38	2.8×10^{-3}	0.10	33
Suc-AlaAlaProPhe-pNA	21	0.15	0.29	33
	Tr → Ch[S1+L1+L2]			
Ac-Phe-NH_2	0.3	0.065	0.018	33
Suc-AlaAlaProPhe-AMC	800	>0.006	>4	37
Suc-AlaAlaProPhe-pNA	1.7×10^3	0.011	20	37
	Tr → Ch[S1+L1+L2+Y172W]			
Ac-Phe-NH_2	0.7	≥0.1	≥0.08	38
Suc-AlaAlaProPhe-AMC	9.3×10^3	≥0.006	≥60	63
Suc-AlaAlaProPhe-pNA	8.3×10^4	5×10^{-4}	41	63

[a] Data from ref 220.

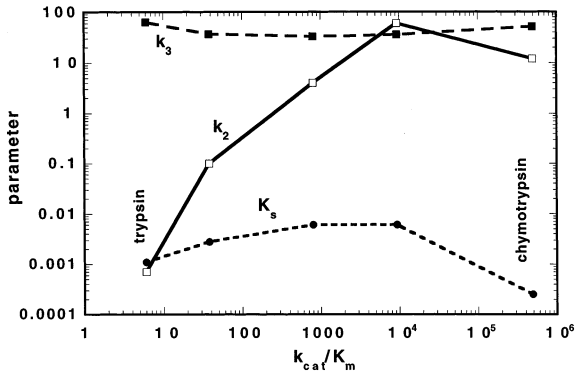


Figure 14. The transfer of P1−Phe specificity into trypsin. The values of K_s, k_2, and k_3 are plotted versus $k_{cat}K_m$ for the hydrolysis of Suc-Ala-Ala-Pro-Phe−AMC by trypsin, Asp189Ser trypsin, Tr → Ch[S1+L1+L2], Tr → Ch[S1+L1+L2+Y172W] and chymotrypsin. The acylation rate k_2 increases as k_{cat}/K_m increases, but the values of K_s and k_3 remain constant.

counterparts from chymotrypsin. These substitutions are Asp189Ser, Gln192Met, insertion of Thr219, Ile138Thr, the 185 loop (residues 185−188) and the 220 loop (residues 221−224). This mutant protease, Tr → Ch[S1+L1+L2], has 100× more activity than D189S, and has 0.5% of the activity of chymotrypsin (Table 6). An additional substitution, Tyr172Trp, increases activity to within 15% of chymotrypsin.[220] The installation of chymotrypsin-like activity into the trypsin framework is almost entirely the result of acceleration of the acylation step, which increases by ~10^4-fold (Figure 14). Deacylation is comparable in all of the enzymes. However, none of the mutant proteases bind substrates well.

Although these mutations focused on the S1 site, the increase in acylation originates in the interactions of the more remote sites. For chymotrypsin, the value of k_2 is ~400-fold larger for a long substrate than a short substrate. In contrast, for D189S trypsin, the values of k_2 for the hydrolysis of short and long substrates are similar (Table 6). Thus, unlike chymotrypsin, D189S trypsin cannot utilize binding in the remote sites to increase acylation. However, in Tr → Ch[S1+L1+L2] and Tr → Ch[S1+L1+L2+Y172W], acylation increases by >10^3-fold between short and long substrates. Thus, the 185 and 220 loops mediate the translation of remote binding interactions into catalysis.

Why is Asp189Ser trypsin a poor nonspecific protease? First, the substitution of Asp189 with Ser deforms the S1 site and the activation domain (the region that rearranges during zymogen activation).[219,221] This is the all too typical protein engineering result: mutation leads to unanticipated defects in structure. Interestingly, BPTI induces D189S to assume a trypsin-like conformation, masking these structural defects[222] and underscoring the need to evaluate both free and bound enzymes and substrates when assessing the structural basis of specificity. The substitution of Asp189 with His deforms the S1 site even in the presence of BPTI, further illustrating the inherent fragility of the S1 site.[223] These observations suggest that mutation of Asp189 might be sufficient to change specificity if only the S1 site was more structurally sound.

How do the 185 and 220 loops and Tyr172Trp substitution influence catalysis? The Tyr172Trp mutation helps rebuild the S1 site in the presence of Ser189. The S1 site of Tr → Ch[S1+L1+L2+Y172W], unlike D189S trypsin or Tr → Ch[S1+L1+L2], is sufficiently structured to allow proflavin binding. The structures of the chloromethyl ketone-inactivated complexes of these proteases also suggest that the Tyr172Trp mutation stabilizes the S1 site.[16] Overall, the structures of both Tr → Ch[S1+L1+L2] and Tr → Ch[S1+L1+L2+Y172W] resemble chymotrypsin, but the 185 and 220 loops are disordered in both mutants. Tr → Ch[S1+L1+L2+Y172W] is more structured than Tr → Ch[S1+L1+L2], in keeping with its higher activity. The substitution of the 185 and 220 loops are probably required to construct an S1 site containing Ser189.

The fragility of the S1 site is not the whole story. The substitution of the 185 and 220 loops have a subtle effect on the polypeptide binding site: the conformation of Gly216 changes so that the carbonyl points toward the P3 NH in a better orientation for the hydrogen bond. The function of the polypeptide binding site may be mediated by the conformation of Gly216, and therefore by the 185 and 220 loops.[15] The loops may also modulate the dynamic properties of the active site cleft. Gly216 and the analogous 220 loop also appear to be important determinants of alpha lytic protease specificity, where strong evidence exists for the involvement of dynamic processes (see below).[224,225]

The S1 site, 185 and 220 loops, and residue 172 do not define a universal set of structural determinants of specificity in serine proteases. This set of substitutions fails to install trypsin-like specificity into chymotrypsin,[226] nor do they install an elastase-like specificity into trypsin.[227] These results suggest that a different structural solution is required for each

specificity. In addition, alternate solutions to the trypsin-to-chymotrypsin conversion may exist. Stabilization of the activation domain would also stabilize the S1 site. Bacterial serine proteases do not use the zymogen activation mechanism of mammalian serine proteases, which suggests that other strategies exist for stabilizing the S1 site. This strategy is probably the mechanism operating a metal-activated Asp189Ser trypsin variant. The substitution of Asn143 and Glu151 with His creates a metal binding site in trypsin.[228] This metal site can also interact with a P2′ His residue, thus installing specificity for His at the S2′ site. These mutations also increase the activity of Asp189Ser trypsin by 300-fold. Asn143 and Glu151 are also part of the activation domain; metal binding will stabilize the activation domain, and by extension, the S1 site, which accounts for the increase in activity.

Unlike the S1 site, the S1′ and S2′ sites can be redesigned easily. A single substitution, Lys60Glu, is sufficient to create a trypsin variant with a ~100-fold preference for Arg at S1′.[229] The construction of a metal binding site at S2′ as described above was sufficient to create specificity for P2 His.[223] Why are the S1 sites of trypsin and chymotrypsin so difficult to redesign, especially when compared with the S1′ site? Part of the answer is undoubtedly in topology of the sites. The S1 site is a deep pocket which surrounds the P1 residue. A cavity is an intrinsically unstable structure and less able to tolerate substitutions. In contrast, the S1′ site is a shallow groove. Moreover, the S1 site of trypsin is part of the activation domain of trypsin. The S1 site has two conformations, the ordered structure of the mature protease and an inactive, zymogen-like conformation. It is not surprising that the Asp189Ser mutation causes complete deformation of the S1 site, pushing the conformational equilibrium to the zymogen-like form.[221,230] In effect, the S1 site is "built to collapse". In contrast, the structure of the S1′ site remains constant during zymogen activation, and is inherently more structurally sound than the S1 site. The stability of the framework is obviously an important determinant of the ability to redesign an enzyme.

A well-ordered S1 site may not be an absolute requirement for high protease activity, so long as the conversion to a catalytically competent structure does not invoke a huge energetic penalty. Tr → Ch[S1+L1+L2] is almost as active as chymotrypsin although its S1 site is deformed. However, the rigid S1 site of chymotrypsin is almost certainly a requirement for survival in its physiological niche—the protease-rich duodenum. It is worth remembering that enzymes evolve to optimize all aspects of their physiological functions, including catalysis, expression, translation, folding, and degradation.

B. Redesigning Esterase Activity

In contrast to amide hydrolysis, the esterase specificity of the S1 site can be easily re-engineered. The Asp189Ser mutation is sufficient to change the specificity of trypsin to that of chymotrypsin in ester hydrolysis.[200] Trypsin has been converted into an elastase-like esterase, but this enzyme has no measurable amidase activity.[227] Similarly, the preference of granzyme B for acidic P1 residues can be changed to Phe with the substitution of Arg226 with Gly (unfortunately, this work did not utilize purified enzymes, and therefore must be included with caution).[231] The ease with which esterase activity can be manipulated derives from the much greater intrinsic reactivity of esters as discussed above.

VII. How Do Remote Interactions Translate into Catalysis?

The operation of the catalytic machinery in the immediate vicinity of the scissile bond appears very straightforward: Asp102, His57, and Ser195 and the oxyanion hole are aligned to facilitate proton transfers and stabilize the charges that develop as the reaction proceeds. It is also clear that a network of hydrogen bonds propagates outward from the catalytic triad, so that changes in bonding order and charge distribution at the scissile bond will propagate to more remote enzyme−substrate interactions, and vice versa. Interactions far from the scissile bond, in the S1 site, polypeptide site, S2−Sn or S1′−Sn′, can dramatically and cooperatively accelerate the hydrolysis of substrate. Such interactions must optimize the function of the catalytic machinery.

The influence of remote binding interactions on the catalytic triad is illustrated by trifluoroketone inhibition. The His57 Nδ1 protons have low-field chemical shifts, elevated pK_a's, low fractionation factors and slow rates of D_2O exchange.[149,150,165] For a series of inhibitors varying in the P2 residue, these properties correlate with the affinity of the inhibitor as well as k_{cat}/K_m for the analogous substrate.[150,151,163,165,232] Of course, the big question is how do these remote interactions translate into catalysis. Most remote interactions do not increase substrate affinity but contribute exclusively to catalysis. This observation suggests that these interactions introduce strain into the substrate, which penalizes binding affinity but accelerates catalysis. This strain can take many forms, including desolvation of enzyme and substrate, locking the substrate in the canonical conformation, interactions of the carbonyl with the oxyanion hole, damping of protein motions, conformational changes and/or geometric distortion. What follows are some possible mechanisms, in no particular order. These mechanisms are not mutually exclusive, and it is likely that several are operating.

A. Remote Interactions Align the Substrate in the Active Site

One proposal is that remote binding interactions align the substrate relative to the catalytic triad, placing the substrate in catalytic register.[15] This mechanism is particularly attractive when considering the protease subsites immediately adjacent to the scissile bond, i.e., S2, S1, and S1′. These interactions dock the scissile bond next to Ser195 and the oxyanion hole. Unfortunately, the alignment mechanism becomes less attractive as the interactions get farther from the site of chemical transformation. Once the P1/S1 and P2/S2 interactions are formed, the sub-

strate is tethered in position; there would seem to be little gained for interactions beyond the S2 site. Another mechanism would appear necessary to harness the binding energy at these more remote sites.

Any explanation for how distal interactions accelerate protease catalysis must also explain the how these interactions can act "in trans". For example, amines or guanidinium compounds accelerate the hydrolysis of substrates containing P1-Gly or Ala by trypsin.[233,234] This observation suggests that the P1−S1 interaction does not simply dock the substrate in the active site cleft. Ethylamine could trigger a subtle conformational change that activates the enzyme. The P1/S1 interaction may prevent nonproductive binding modes, thus indirectly improving the alignment of the substrate in the active site. However, the alignment mechanism does not appear to be sufficient to explain the effects of the remote binding interactions on catalysis.

B. Remote Interactions Are Optimized in the Transition State

Enzymes are often described as templates for the transition state of a reaction. Conventional views of the transition state of protease reactions focus on the site of bond making/breaking and formation of the tetrahedral adduct. However, if remote interactions increase catalysis, then remote interactions must also be part of the transition state. Therefore, a complete description of the transition state must include the P/S and P′/S′ interactions in addition to the tetrahedral adduct, catalytic triad and oxyanion hole. The most dramatic changes in structure are involved in the bond making and breaking at the scissile bond, while smaller changes result from the changes in geometry in the transformation from planar substrate to tetrahedral transitions state. These small changes can propagate into significant displacements down the peptide chain, optimizing the remote binding site interactions in the transition state.[235] Such changes could provide a mechanism for the rate acceleration. Support for this mechanism can be found in the inhibition of elastase by peptidyl aldehydes.[47] While the P4 residue increases the affinity of an alcohol inhibitor by 10-fold, the affinity of the corresponding aldehyde inhibitor increases by 100-fold (Table 7). This observation demonstrates that the P4−S4 interaction is optimized in a tetrahedral adduct. Similarly, the hydrogen bond between the P1−NH and the Ser214 carbonyl appears to be stronger in transition state analogue complexes than in protein inhibitor complexes (Table 8).[45,236] These observations suggest that the interactions at the remote sites are optimized in the transition state.

Table 7. Inhibition of Porcine Pancreatic Elastase by Peptidyl Aldehydes[a]

inhibitor	K_i (μM)
Ac-Ala-Pro-Ala-CH_2OH	7000
Ac-Pro-Ala-Pro-Ala-CH_2OH	600
Ac-Ala-Pro-Ala-CHO	62
Ac-Pro-Ala-Pro-Ala-CHO	0.8

[a] Data from ref 47.

Table 8. Optimization of Hydrogen Bonds in Transition State Complexes[a]

hydrogen bond	boroVal	Pval-Lac (Pval)	ovomucoid
Ser195 N−P1 O1	2.9	2.8	3.1
Gly193 N−P1 O1	2.6	2.6	2.6
His57 Nϵ2−P1 O2	2.7	5.5 (2.9)[b]	
Ser214 O−P1 N	3.0	3.0	3.6
Gly216 N−Ala P3 O	3.0	3.0	2.9
Gly216 O−Ala P3N	2.9	2.9	3.0
His57 Nϵ2−Ser195 Oγ	3.0	7.1 (3.0)	2.6
Leu41 O−P2′ N		2.8	2.9
Leu41 N−P2′ O		3.0	3.1

[a] X-ray crystal structures of alpha lytic protease complexes with Boc−Ala−Pro−boro-Val,[37] Boc−Ala−Ala−Pro−Pval−Lac−Ala, where Pval is the phosphonic acid analogue of Val and Lac is lactate [45] and Boc−Ala−Ala−Pro−Pval (in parentheses) and ovomucoid third domain[236] are compared. Data from ref 45. [b] In the Boc−Ala−Ala−Pro−Pval−Lac−Ala, His57 has rotated away from the inhibitor, presumably due to the absence of a negative charge in the neutral phosphonate. The Nϵ1 hydrogen bond is present in the charged phosphonate.

C. Remote Interactions Induce a Conformational Change that Favors Catalysis

One oft-invoked mechanism for utilizing remote binding interactions is induced fit, i.e., remote interactions induce a conformational change that aligns the catalytic machinery, activating the enzyme. However, no conformational change is observed when chymotrypsin or trypsin bind substrates or inhibitors, and the alignment of the catalytic triad does not change. Thus, a classical induced fit mechanism is not operative and serine proteases approach the ideal of the enzyme as a template for the transition state (examples of serine proteases that do utilize an induced fit mechanism will be discussed below). However, subtle conformational changes are observed when apoenzyme and inhibitor complexes are carefully compared. For example, when trypsin binds a peptidyl boronic acid, small adjustments are observed in ~50 of the 224 residues,[237] and similar adjustments are also observed in boronic acid structures of alpha lytic protease.[37] These structural changes are small and seem unlikely to involve a significant energetic barrier. Rather, this structural plasticity is probably required to accommodate product association/dissociation and the different transition states of the protease reaction.

If good substrates induce conformational changes that increase the reactivity of the catalytic triad and oxyanion hole, then disruptions of the catalytic machinery should have a more deleterious effect on the hydrolysis of good substrates. Unfortunately, there are little data on this point because these mutations decrease activity to the point where only good substrates can be monitored. The substitution of Asp102 with Asn appears to have a more deleterious effect on the hydrolysis of peptide substrates than other substrates, but this difference is not observed when the pH-independent reaction is considered.[79] Mutations of His57 do change specificity, but this is expected because His57 forms part of both the S2 and S1′ sites (Figures 7). Unfortunately, it will be difficult to correlate such subtle structural differences to catalysis.

D. The Remote Interactions Shield the Catalytic Triad from Solvent

Many enzymes, even proteases, have motile flaps that close over the active sites, preventing solvent access to the active site. Such desolvation provides a simple mechanism to achieve large rate accelerations.[262,264] The remote portions of the peptide substrate could perform this task for serine proteases. In particular, the P2 and P1′ residues will shield His57 from bulk solvent. This effect can be illustrated by the exchange of the protons of His57 with solvent in trifluoroketone complexes. The rate of exchange decreases markedly with better P2 residues (from 280 to 12 s^{-1}; these rates are 100× greater than the rate of inhibitor dissociation, indicating that proton exchange depends on the diffusion of water into the complex).[150] Thus, the P2 residue blocks water access to the catalytic triad. Interactions at more remote sites could similarly restrict solvent access as well as lower the dielectric constant of the active site. Similarly, the "trans" phenomenon could be explained by the expulsion of water from the active site, or perhaps by neutralizing the negative charge of Asp189, favoring catalysis by lowering the dielectric constant in the vicinity of the scissile bond. Note that the tight packing of enzyme and substrate can cause steric compression that will also facilitate proton transfer.[23,163,238]

E. Remote Interactions Couple Catalysis to Motion of the Protease Structure

Protein dynamics is the wild card in protease catalysis and specificity, as is true of enzyme catalysis as a whole. The idea that reaction dynamics are coupled to protein motions has long intrigued biochemists. Such motions might be used to mechanically stress the peptide bond, disrupting resonance stabilization and activating the carbonyl for nucleophilic attack.[239] Likewise, the attack of Ser195 on the scissile bond might be linked to movement of the surrounding residues. Less dramatically, dynamic movement could simply be important for substrate association and product dissociation. The location of the active site between the two beta barrels places the substrate at junction of domain motions in chymotrypsin-like proteases.

The serine proteases provide a unique opportunity to address this question: if such protein motion is required for catalysis, then similar motions should be observed in chymotrypsin-like and subtilisin-like proteases, as well as the other serine protease families. On the other hand, it is difficult to see how protein dynamics could be maintained in four different structures. It may be possible to monitor protein motions during serine protease catalysis. Such experiments best utilize reversible reactions, which has thwarted their application to protease reactions. However, proteases also catalyze the exchange of water into the C-terminus of a product carboxylic acid. This reaction could be exploited to investigate protease motions during catalysis.

VIII. Specificity via Induced Fit and Allostery

The above discussion focused on specificity that derived from a relatively rigid active site cleft as found in chymotrypsin, trypsin, and elastase. Serine protease specificity has also been ascribed to classical induced fit mechanisms and allosteric interactions. The following will provide examples of these phenomena with complement protease Factor D, the bacterial enzyme α-lytic protease and the coagulation protease thrombin.

A. Induced Fit as a Mechanism for Narrow Specificity

It is important to recognize that an effective induced fit mechanism requires more than the simply placing a conformational change in the reaction coordinate (Figure 15).[240,241] An enzyme that can rapidly equilibrate between inactive and active conformations will be no more specific than an enzyme that does not undergo a conformational change—the equilibrium between inactive and active forms simply lowers the fraction of active enzyme. This must be the case if the chemical transformation is rate-limiting for the reaction; the substrate binding steps will be in equilibrium, and no advantage is gained from the conformational change.

1. Rigid enzyme

$$E \xrightleftharpoons{S_1} E{\bullet}S_1 \xrightarrow{slow} E + P_1$$

$$E \xrightleftharpoons{S_2} E{\bullet}S_2 \xrightarrow{slow} E + P_2$$

$$specificity = \frac{(k_{cat}/K_m)_{S1}}{(k_{cat}/K_m)_{S2}}$$

2. Enzyme with a conformational change, $K_c^* << 1$

$$\begin{array}{ccc} E^* + S_1 & \xrightleftharpoons{K_{s1}} & E^*{\bullet}S_1 \\ K_c^* \uparrow\downarrow & & \uparrow\downarrow \\ E + S_1 & \xrightleftharpoons{} & E{\bullet}S_1 \xrightarrow[slow]{k_{c1}} E + P_1 \end{array}$$

$$\begin{array}{ccc} E^* + S_2 & \rightleftharpoons & E^*{\bullet}S_2 \\ K_c^* \uparrow\downarrow & & \uparrow\downarrow \\ E + S_2 & \xrightleftharpoons{K_{s2}} & E{\bullet}S_2 \xrightarrow[slow]{k_{c2}} E + P_2 \end{array}$$

$$specificity = \frac{(k_{cat}/K_m)_{S1} \times (1/K_c^*)}{(k_{cat}/K_m)_{S2} \times (1/K_c^*)} = \frac{(k_{cat}/K_m)_{S1}}{(k_{cat}/K_m)_{S2}}$$

3. Induced fit enzyme, two active conformations

$$\begin{array}{ccccc} E^* + S_1 & \xrightleftharpoons{K_{s1}} & E^*{\bullet}S_1 & \xrightarrow{k_{c1}} & E^* + P_1 \\ K_c^* \uparrow\downarrow & & & & \\ E + S_2 & \xrightleftharpoons{K_{s2}} & E{\bullet}S_2 & \xrightarrow{k_{c2}} & E + P_2 \end{array}$$

Figure 15. The effect of a conformational change on specificity. Simply adding a conformational change to an enzyme does not provide specificity relative to an enzyme that does not undergo a conformational change. As shown in Schemes 1 and 2, the conformational change simply decreases the amount of active enzyme form. However, with the added constraint that the conformational change is rate-limiting, and the good substrate accelerates the conformational change, specificity can be achieved because the binding steps are no longer at equilibrium. If both conformations are active as in Scheme 3, broad specificity can be attained.

This principle is illustrated in the trypsin/trypsinogen system, where mutations that alter the equilibrium between active and inactive conformations do not change specificity.[242] The Ile16−Asp194 salt bridge can be destabilized by the substitution of Ile16 with smaller residues, favoring the inactive zymogen-like conformation. Such mutations decrease the hydrolysis of good and poor substrates equivalently, including both amides and esters.[243] In addition, mutations can be introduced that activate trypsinogen by stabilizing the mature protease conformation.[244] The equilibrium between inactive and active conformations can be determined by measuring the affinity of BPTI, which induces the mature protease conformation.[242] The decrease in values of k_{cat}/K_m correlates with the decrease in affinity of BPTI, as expected for the mechanism of Figure 15.

An induced fit mechanism can impart specificity when the conformational change is rate-limiting. In this case, a good substrate can accelerate the rate of the conformational change, and thus accelerate the reaction. The binding steps will not be at equilibrium because the substrate will rapidly react to products. Further discrimination can be achieved if chemistry is slow for the poor substrate.[241]

B. Factor D: An Induced Fit Protease

Such an induced fit mechanism can account for the remarkable specificity of Factor D. Unlike most serine proteases, Factor D circulates in plasma in its mature form, and therefore must have extraordinary specificity. Factor D has only one natural substrate, the complement complex C3bB.[245] Remarkably, k_{cat}/K_m for hydrolysis of C3bB is 10^6 M^{-1} s^{-1}, comparable to the hydrolysis of a good peptide substrate by trypsin. However, Factor D has little esterase activity ($k_{cat}/K_m = \sim 300$ M^{-1} s^{-1} compared to 10^7 M^{-1} s^{-1} for trypsin[197]) and cannot hydrolyze peptides containing the C3bB cleavage site.[197] The structure of Factor D does not resemble an active serine protease, which explains these puzzling observations.[246] The S1 site, S2−S4 sites and catalytic triad are deformed due to the presence of Ser94, Thr214, and Ser215 at positions usually occupied by Trp/Tyr, Ser, and Trp/Phe, respectively. Substitution of these residues returns the catalytic triad to a normal conformation, but only increases the esterase activity 10×.[247] The binding of C3bB must repair these defects and induce Factor D to assume the usual active conformation.

C. Alpha Lytic Protease: Induced Fit as a Mechanism for Broad Specificity

Induced fit also provides a mechanism for expanding specificity.[248] Imagine an enzyme has several conformations of comparable energies, each of which accommodates a different substrate or, more precisely, the transition state for a different substrate. Each substrate will induce the appropriate conformation for catalysis, and the enzyme can hydrolyze several substrates with comparable values of k_{cat}/K_m. Such an induced fit mechanism provides the structural plasticity necessary for broad specificity.

Table 9. Specificity of Wild Type and M192A Alpha Lytic Protease[a]

	k_{cat}/K_m M^{-1} s^{-1}	
substrate	wild type	M192A
Suc-Ala-Ala-Pro-Ala−pNA	2.1×10^4	1×10^4
Suc-Ala-Ala-Pro-Val−pNA	790	3×10^3
Suc-Ala-Ala-Pro-Met−pNA	1800	3.5×10^5
Suc-Ala-Ala-Pro-Leu−pNA	4.1	1.1×10^5
Suc-Ala-Ala-Pro-Phe−pNA	0.38	3.1×10^5

[a]Data from ref 249.

The broad specificity of the Met192Ala mutant of alpha lytic protease can be understood in terms of this structural plasticity (although induced fit is probably not an appropriate term since the catalytic machinery is always aligned). Alpha lytic protease prefers substrates with small aliphatic residues at P1. The S1 site is small and relatively rigid, occluded by Met192. Substitution of Met192 with Ala would be expected to create a large hydrophobic S1 site similar to chymotrypsin and the X-ray crystal structure of the unliganded mutant protease shows that the S1 site is enlarged. However, unlike chymotrypsin, Met192Ala hydrolyzes P1−Ala and Phe substrates with comparable values of k_{cat}/K_m (Table 9).[249] Inhibitor complexes show that the structure of the S1 site of Met192Ala is not a simple hydrophobic pocket, but adapts to accommodate different P1 residues.[251] Therefore, the S1 site displays structural plasticity wherein each substrate induces the appropriate conformation for hydrolysis. The mutation appears to have a profound effect on the dynamic properties of the S1 site. In wild-type alpha lytic protease, the walls of the S1 site are coupled and move together as a unit so that the dimensions of the S1 site are maintained.[252] However, the Met192 mutation disrupts this coupling, allowing the two walls of the S1 site to move independently and the dimensions of the S1 site to vary.[250,253,254] Thus, the mutant enzyme can sample many different conformations of the S1 site, and hydrolyze many different substrates.

Thrombin provides a second example of how specificity can be expanded by the ability to access multiple conformations. Thrombin's role in the formation of a blood clot is well-known.[265,266] Less appreciated is thrombin's anticoagulation function. The substrates of the coagulation pathway (fibrinogen, factor V, factor VIII, and factor XIII) contain neutral or positively charged P3 and P3′ residues, while the substrates of the anticoagulation pathway (e.g., protein C) contain Asp at these positions. Thrombin requires two conformations to process such different subtrates.[255,256] The procoagulant conformation, or "fast" form, hydrolyzes fibrinogen $\sim 10^3$ times faster than protein C. In contrast, the "slow" form has a 50-fold preference for protein C.[255] Importantly, Na^+ is an allosteric effector of this conformational switch, binding specifically to the fast form at a single site.[257] The fast and slow forms are approximately equal at physiological Na^+ concentrations. Na^+ is an allosteric effector of several other serine proteases, including Factor Xa, activated protein C, and Factor IXa; Na^+ regulation correlates with the presence of

Tyr or Phe at position 225, while Na^+ independence correlates with Pro225.[258] The Na^+ binding site lies between the 180 and 220 loops.[259] Substitution of Trp215 and Glu217 abrogates Na^+ binding, stabilizing the slow form while Glu39, Trp60d, Glu192, Asp221, and Asp222 form an "allosteric core" that connects the Na^+ binding site to the active site cleft.

IX. How Serine Proteases Work: Summary and Prospects

Serine proteases are probably the most thoroughly investigated enzyme system. While the function of the catalytic triad and oxyanion hole can be rationalized in terms of electrostatic stabilization of charges developing in the transition state, how remote binding interactions facilitate catalysis remains a mystery. Catalysis and specificity are not simply controlled by a few residues, but are rather a property of the entire protein framework, controlled via the distribution of charge across a network of hydrogen bonds and perhaps also by the coupling of domain motion to the chemical transformation. Serine proteases present a unique opportunity to test these ideas.

X. Useful Web Sites

Rose, T., and Di Cera, E. Serine Protease Home Page. http://www.biochem.wustl.edu/~protease/index.html.

Rawlings, N. D., O'Brien, E. A., and Barrett, A. J. 2002 MEROPS: The Protease Database. *Nucleic Acids Res. 30*, 343. http://merops.iapc.bbsrc.ac.uk/.

XI. References

(1) Blow, D. M. *Trends Biochem. Sci.* **1997**, *22*, 405−408.
(2) Dodson, G.; Wlodawr, A. *Trends Biochem. Sci.* **1998**, *23*, 347−352.
(3) Rawlings, N. D.; Barrett, A. J. *Nucleic Acids Res.* **2000**, *28*, 323−325.
(4) Neurath, H. *Science* **1984**, *224*, 350−357.
(5) Johnson, D. E. *Leukemia* **2000**, 1695−1703.
(6) Joseph, K.; Ghebrehiwet, B.; Kaplan, A. P. *Biol. Chem.* **2001**, *382*, 71−75.
(7) Coughlin, S. R. *Nature* **2000**, *407*, 258−264.
(8) Barros, C.; Crosby, J. A.; Moreno, R. D. *Cell Biol. Int.* **1996**, *20*, 33−39.
(9) Davidson, D. J.; Higgins, D. L.; Castellino, F. J. *Biochemistry* **1990**, *29*, 3585−3590.
(10) Sim, R. B.; Laich, A. *Biochem. Soc. Trans.* **2000**, *28*, 545−50.
(11) Collen, D.; Lijnen, H. R. *CRC Crit. Rev. Oncol. Hematol.* **1986**, *4*, 249−301.
(12) LeMosy, E. K.; Hong, C. C.; Hashimoto, C. *Trends Cell Biol.* **1999**, *9*, 102−107.
(13) Van den Steen, P. E.; Opdenakker, G.; Wormald, M. R.; Dwek, R. A.; Rudd, P. M. *Biochim. Biophys. Acta* **2001**, *1528*, 61−73.
(14) Selvarajan, S.; Lund, L. R.; Takeuch, T.; Craik, C. S.; Werb, Z. *Nat. Cell Biol.* **2001**, *3*, 267−275.
(15) Perona, J. J.; Craik, C. S. *Protein Sci.* **1995**, *4*, 337−360.
(16) Perona, J. J.; Hedstrom, L.; Rutter, W. J.; Fletterick, R. J. *Biochemistry* **1995**, *34*.
(17) Czapinska, H.; Otlewski, J. *Eur. J. Biochem.* **1999**, *260*, 571−595.
(18) Blow, D. M. In *The Enzymes*; 3rd ed.; Boyer, P. D., Ed.; Academic Press: Boca Raton, 1971; Vol. 3.
(19) Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331−358.
(20) Patthy, L. *Methods Enzymol.* **1993**, *222*, 10−21.
(21) Bachovchin, W. W. *Magn. Reson. Chem.* **2001**, *39*, S199−S213.
(22) Gandour, R. D. *Bioorg. Chem.* **1981**, *10*, 169−176.
(23) Gandour, R. D.; Nabulsi, N. A. R.; Fronczek, F. R. *J. Am. Chem. Soc.* **1990**, *112*, 7816−7817.
(24) McGrath, M. E.; Vasquez, J. R.; Craik, C. S.; Yang, A. S.; Honig, B.; Fletterick, R. J. *Biochemistry* **1992**, *31*, 3059−3064.
(25) Epstein, D. M.; Abeles, R. H. *Biochemistry* **1992**, *31*, 11216−11223.
(26) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1994**, *241*, 83−93.
(27) Ash, E. L.; Sudmeier, J. L.; DAy, R. M.; Vincent, M.; Torchilin, E. V.; Haddad, K. C.; Bradshaw, E. M.; Sanford, D. G.; Bachovchin, W. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10371−10376.
(28) Matthews, D. A.; Alden, R. A.; Birktoft, J. J.; Freer, S. T.; Kraut, J. *J. Biol. Chem.* **250**, *250*, 7120−7126.
(29) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157−162.
(30) Knowles, J. R. *J. Theor. Biol.* **1965**, *9*, 213−228.
(31) Dorovskaya, V. N.; Varfolomeyev, S. D.; Kaznskaya, N. F.; Klyosov, A. A.; Martinek, K. *FEBS Lett.* **1972**, *23*, 122.
(32) Huber, R.; Kukla, D.; Bode, W.; Schwager, P.; Bartels, K.; Deisenhofer, J.; Steigemann, W. *J. Mol. Biol.* **1974**, *89*, 73−101.
(33) Kreiger, M.; Kay, L. M.; Stroud, R. M. *J. Mol. Biol.* **1974**, *83*, 209−230.
(34) Shotton, D. M.; Watson, H. C. *Nature* **1969**, *25*, 811−816.
(35) Waugh, S. M.; Harris, J. L.; Fletterick, R.; Craik, C. S. *Nat. Struct. Biol.* **2000**, *7*, 762−765.
(36) Huber, H.; Bode, W. *Acct. Chem. Res.* **1978**, *11*, 114−122.
(37) Bone, R.; Shenvi, A. B.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1987**, *26*, 7609−7614.
(38) Laskowski, M.; Qasim, M. A. *Biochim. Biophys. Acta* **2000**, *1477*, 324−337.
(39) Wilmouth, R. C.; Clifton, I. J.; Robnson, C. V.; Roach, P. L.; Aplin, R. T.; Westwood, N. J.; Hajdu, J.; Schofield, C. J. *Nat. Struct. Biol.* **1997**, *4*, 456−461.
(40) Bode, W.; Huber, R. *Eur. J. Biochem.* **1992**, *204*, 433−451.
(41) Dixon, M. M.; Matthews, B. W. *Biochemistry* **1989**, *28*, 7033−7038.
(42) Blanchard, H.; James, M. N. G. *J. Mol. Biol.* **1994**, *241*, 574−587.
(43) Bode, W.; Huber, R. *Eur. J. Biochem.* **1992**, *204*, 433−451.
(44) Read, R. J.; James, M. N. G. Elsevier Science Publishing Co., Inc.: New York, 1986.
(45) Bone, R.; Sampson, N. S.; Bartlett, P. A.; Agard, D. A. *Biochemistry* **1991**, *30*, 2263−2272.
(46) Coombs, G. S.; Rao, M. S.; Olson, A. J.; Dawson, P. E.; Madison, E. L. *J. Biol. Chem.* **1999**, *274*, 24074−24079.
(47) Thompson, R. C. *Biochemistry* **1973**, *12*, 47−51.
(48) Kettner, C. A.; Bone, R.; Agard, D. A.; Bachovchin, W. W. *Biochemistry* **1988**, *27*, 7682−7688.
(49) Brady, K.; Abeles, R. H. *Biochemistry* **1990**, *29*, 7608−7617.
(50) Schellenberger, V.; Braune, K.; Hofmann, H. J.; Jakubke, H. D. *Eur. J. Biochem.* **1991**, *199*, 623−636.
(51) Bode, W.; Turk, D.; Karshikov, A. *Protein Sci.* **1992**, *1*, 426−471.
(52) Stein, R. L.; Strimpler, A. M.; Hori, H.; Powers, J. C. *Biochemistry* **1987**, *26*, 1301−1305.
(53) Bode, W.; Meyer, E., Jr.; Powers, J. C. *Biochemistry* **1989**, *28*, 1951−1963.
(54) Lu, D.; Futterer, K.; Korolev, S.; Zheng, X.; Tan, K.; Waksman, G.; Sadler, J. E. *J. Mol. Biol.* **1999**, *292*, 361−373.
(55) Schellenberger, V.; Turck, C. W.; Rutter, W. J. *Biochemistry* **1994**, *33*, 4251−4257.
(56) Hess, G. P.; McConn, J.; Ku, E.; McConkey, G. *Philos. Trans. R. Soc. London* **1970**, *B 257*, 89−104.
(57) Fersht, A. R. *J. Mol. Biol.* **1972**, *64*, 497−509.
(58) Brayer, G. D.; Delbaere, L. T.; James, M. N. *J. Mol. Biol.* **1979**, *131*, 743−775.
(59) Zerner, B.; Bond, R. P. M.; Bender, M. L. *J. Am. Chem. Soc.* **1964**, *86*, 3674−3679.
(60) Bender, M. L.; Killhefer, J. V. *Crit. Rev. Biochem.* **1973**, *1*, 149.
(61) Steitz, T. A.; Shulman, R. G. *Annu. Rev. Biophys. Bioeng.* **1982**, *11*, 419−444.
(62) Polgar, L. *Mechanism of Protease Action*; CRC Press: Boca Raton, 1989.
(63) Jansen, E. F.; Nutting, M.-D. F.; Balls, A. K. *J. Biol. Chem.* **1949**, *179*, 201.
(64) Schaffer, N. K.; May, S. C., Jr.; Summerson, W. H. *J. Biol. Chem.* **1953**, *202*, 67.
(65) Oosterbaan, R. A.; Kunst, P.; van Rotterdam, J.; Cohen, J. A. *Biochem. Biophys. Acta* **1958**, *27*, 556.
(66) Schoellman, G.; Shaw, E. *Biochemistry* **1963**, *2*, 252.
(67) Blow, D. M.; Birktoft, J. J.; Hartley, B. S. *Nature* **1969**, *221*, 337−40.
(68) Shotton, D. M.; Watson, H. C. *Nature* **1970**, *25*, 811−816.
(69) Wright, C. S.; Alden, R. A.; Kraut, J. *Nature* **1969**, *221*, 235−42.
(70) Drenth, J.; Hol, W. G.; Jansonius, J. N.; Koekoek, R. *Eur. J. Biochem.* **1972**, *26*, 177−81.
(71) Henderson, R. *J. Mol. Biol.* **1970**, *54*, 341−354.
(72) Powers, J. C. *Methods Enzymol.* **1977**, *46*, 197−208.

(73) Bone, R.; Frank, D.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1989**, *28*, 8, 7600−7609.
(74) Brady, K.; Wei, A.; Ringe, D.; Abeles, R. H. *Biochemistry* **1990**, *29*, 7600−7607.
(75) Weiner, H.; White, W. N.; Hoare, D. G.; Koshland, D. E., Jr. *J. Am. Chem. Soc.* **1966**, *88*, 3851−3859.
(76) Nakagawa, T.; Bender, M. L. *Biochemistry* **1970**, *9*, 259.
(77) Henderson, R. *Biochem. J.* **1971**, *124*, 13−18.
(78) West, J. B.; Hennen, W. J.; Lalonde, J. L.; Bibbs, J. A.; Zhong, Z.; Meyer, E. F., Jr.; Wong, C.-H. *J. Am. Chem. Soc.* **1990**, *112*, 5313−5320.
(79) Corey, D. R.; Craik, C. S. *J. Am. Chem. Soc.* **1992**, *114*, 1784−1790.
(80) Craik, C. S.; Roczniak, S.; Largman, C.; Rutter, W. J. *Science* **1987**, *237*, 909−913.
(81) Sprang, S.; Standing, T.; Fletterick, R. J.; Stroud, R. M.; Finer-Moore, J.; Xuong, N.-H.; Hamlin, R.; Rutter, W. J.; Craik, C. S. *Science* **1987**, *237*, 905−909.
(82) Carter, P.; Wells, J. A. *Nature* **1988**, *332*, 564−568.
(83) Carter, P.; Wells, J. A. *Science* **1987**, *237*, 394−399.
(84) Carter, P.; Abrahmsen, L.; Wells, J. A. *Biochemistry* **1991**, *30*, 6142−6148.
(85) Dall'Acqua, W.; Halin, C.; Rodrigues, M. L.; Carter, P. *Protein Eng.* **1999**, *12*, 981−7.
(86) Corey, D. R.; Willett, W. S.; Coombs, G. S.; Craik, C. S. *Biochemistry* **1995**, *34*, 11521−11527.
(87) Bryan, P.; Pantoliano, M. W.; Quill, S. G.; Hsiao, H.-Y.; Poulos, T. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3743−3745.
(88) Braxton, S.; Wells, J. A. *J. Biol. Chem.* **1991**, *266*, 11797−11800.
(89) Wells, J. A.; Cunningham, B. C.; Graycar, T. P.; Estell, D. A. *Philos. Trans. R. Soc. London Ser. A* **1986**, *317*, 415−423.
(90) Carter, P.; Wells, J. A. *Proteins Struct. Func. Genet.* **1990**, *7*, 335−342.
(91) Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. A. *Nature* **1987**, *328*, 551−554.
(92) Hartley, B. S.; Kilbey, B. A. *Biochem. J.* **1954**, *56*, 288−297.
(93) Fastrez, J.; Fersht, A. R. *Biochemistry* **1973**, *12*, 2025−2034.
(94) Bender, M. L.; Clement, G. E.; Gunter, C. R.; Kezdy, F. J. *J. Am. Chem. Soc.* **1964**, *86*, 3697−3703.
(95) Fersht, A. R.; Blow, D. M.; Fastrez, J. *Biochemistry* **1973**, *12*, 2035−2041.
(96) Stein, R. L.; Strimpler, A. M. *Biochemistry* **1987**, *26*, 2238−2242.
(97) Schellenberger, V.; Jakubke, H.-D. *Biochim. Biophys. Acta* **1986**, *869*, 54−60.
(98) Schellenberger, V.; Turck, C. W.; Hedstrom, L.; Rutter, W. J. *Biochemistry* **1993**, *32*, 4349−4353.
(99) Kurth, T.; Ullman, D.; Jakubke, H.-D.; Hedstrom, L. *Biochemistry* **1997**, *36*, 10098−10104.
(100) Dixon, M. M.; Brennan, R. G.; Matthews, B. W. *Int. J. Biol. Macromol.* **1991**, *13*, 89−96.
(101) Harel, M.; Su, C. T.; Frolow, F.; Silman, I.; Sussman, J. L. *Biochemistry* **1991**, *30*, 5217−5225.
(102) Wilmouth, R. C.; Edman, K.; Neutze, R.; Wright, P. A.; Clifton, I. J.; Schneider, T. R.; Schofield, C. J.; Hajdu, J. *Nat. Struct. Biol.* **2001**, *8*, 689−694.
(103) Bender, M. L. *Chem. Rev.* **1960**, *60*, 53−113.
(104) Bender, M. L.; Thomas, R. J. *J. Am. Chem. Soc.* **1961**, *83*, 4183.
(105) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; McGraw-Hill: New York, 1969.
(106) Ba−Saif, S.; Luthra, A. K.; Williams, A. *J. Am. Chem. Soc.* **1989**, *111*, 2647−2652.
(107) Stefanidis, D.; Cho, S.; Dhe-Paganon, S.; Jencks, W. P. J. *J. Am. Chem. Soc.* **1993**, *115*, 1650−1656.
(108) Hengge, A. C.; Hess, R. A. *J. Am. Chem. Soc.* **1994**, *116*, 11256−11263.
(109) Cleland, W. W.; Hengge, A. C. *FASEB J.* **1995**, *9*, 1585−1594.
(110) Fastrez, J.; Fersht, A. R. *Biochemistry* **1973**, *12*, 1067−1074.
(111) Inward, P. W.; Jencks, W. P. *J. Biol. Chem.* **1965**, *240*, 1986.
(112) Markley, J. L.; Travers, F.; Balny, C. *Eur. J. Biochem.* **1981**, *120*, 477.
(113) Hunkapiller, M. W.; Forgac, M. D.; Yu, E. H.; Richards, J. H. *Biochem. Biophys. Res. Commun.* **1979**, *87*, 25−31.
(114) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. *Acta Cryst.* **1983**, *B39*, 480−490.
(115) Deisenhofer, J.; Steigemann, W. *Acta Crystallogr.* **1975**, *B31*, 238−250.
(116) Meyer, E.; Cole, G.; Radhakrishnan, R.; Epp, O. *Acta Crystallogr. B* **1988**, *44 (Pt 1)*, 26−38.
(117) Presnell, S. R.; Patil, G. S.; Mura, C.; Jude, K. M.; Conley, J. M.; Bertrand, J. A.; Kam, C. M.; Powers, J. C.; Williams, L. D. *Biochemistry* **1998**, *37*, 17068−17081.
(118) Yennawar, N. H.; Yennawar, H. P.; Farber, G. K. *Biochemistry* **1994**, *33*, 7326−7336.
(119) Sampson, N. S.; Bartlett, P. A. *Biochemistry* **1991**, *30*, 2255−2263.
(120) Powers, J. C.; Odake, S.; Oleksyszyn, J.; Hori, H.; Ueda, T.; Boduszek, B.; Kam, C. *Agents Actions Suppl.* **1993**, *42*, 3−18.
(121) Poulos, T. L.; Alden, R. A.; Freer, S. Y.; Birktoft, J. J.; Kraut, J. *J. Biol. Chem.* **1976**, *251*, 1097.
(122) James, M. N. G.; Brayer, G. D.; Delbaere, L. T. J.; Sielecki, A. R.; Gertler, A. *J. Mol. Biol.* **1980**, *139*, 423.
(123) Scott, A. I.; Mackenzie, N. E.; Malthouse, J. P. G.; Preimrose, W. V.; Fagerness, P. E.; Brisson, A.; Qi, L. Z.; Bode, W.; Carter, C.; Jang, Y. J. *Tetrahedron* **1986**, *42*, 3269.
(124) Stevens, C. L.; Chang, C. H. *J. Org. Chem.* **1962**, *27*, 4392−4396.
(125) Goyu, B.; Rouessac, F. *Bull. Soc. Chim. Fr.* **1978**, 590−592.
(126) Kreutter, K.; Steinmetz, A. C.; Liang, T. C.; Prorok, M.; Abeles, R. H.; Ringe, D. *Biochemistry* **1994**, *33*, 13792−13800.
(127) Prorok, M.; Albeck, A.; Foxman, B. M.; Abeles, R. H. *Biochemistry* **1994**, *33*, 9784−9790.
(128) Bachovchin, W. W.; Wong, W. Y. L.; Farr-Jones, S.; Shenvi, A. B.; Kettner, C. A. *Biochemistry* **1988**, *27*, 7689−7697.
(129) Ortiz, C.; Tellier, C.; Williams, H.; Stolowich, N. J.; Scott, I. A. *Biochemistry* **1991**, *30*, 10026−10034.
(130) Kurinov, I. V.; Harrison, R. W. *Protein Sci.* **1996**, *5*, 752−758.
(131) Neidhart, D.; wei, Y.; Cassidy, C.; Lin, J.; Cleland, W. E.; Frey, P. A. *Biochemistry* **2001**, *40*, 2439−2447.
(132) Tsilikounas, E.; Kettner, C. A.; Bachovchin, W. W. *Biochemistry* **1993**, *32*, 12651−12655.
(133) Tsilikounas, E.; Kettner, C. A.; Bachovchin, W. W. *Biochemistry* **1992**, *31*, 12839−12846.
(134) Caplow, M.; Jencks, W. P. *J. Biol. Chem.* **1962**, *1*, 883−893.
(135) Gupton, B. F.; Carroll, D. L.; Tuhy, P. M.; Kam, C. M.; Powers, J. C. *J. Biol. Chem.* **1984**, *259*, 4279−4287.
(136) Robillard, G. T.; Powers, J. C.; Wilcox, P. E. *Biochemistry* **1972**, *11*, 1773−1784.
(137) Chase, T. J.; Shaw, E. *Biochem. Biophys. Res. Commun.* **1967**, *29*, 508.
(138) Bender, M. L.; Schonbaum, G. R.; Zerner, B. *J. Am. Chem. Soc.* **1961**, *84*, 2540−2550.
(139) Bernhard, S. A.; Tashjian, Z. H. *J. Am. Chem. Soc.* **1965**, *87*, 1806.
(140) Tonge, P. J.; Carey, P. R. *Biochemistry* **1989**, *28*, 6701−6709.
(141) Stoddard, B. L.; Bruhnke, J.; Porter, N.; Ringe, D.; Petsko, G. A. *Biochemistry* **1990**, *29*, 4871−4879.
(142) Westkaemper, R. B.; Abeles, R. H. *Biochemistry* **1983**, *22*, 3256−3264.
(143) Ringe, D.; Mottonen, J. M.; Gelb, M. H.; Abeles, R. H. *Biochemistry* **1986**, *25*, 5633−5638.
(144) Huntington, J. A.; Read, R. J.; Carrell, R. W. *Nature* **2000**, *407*, 923−926.
(145) Hunkapiller, M. W.; Smallcombe, S. H.; Witaker, D. R.; Richards, J. H. *J. Biol. Chem.* **1973**, *248*, 8306−8308.
(146) Hunkapiller, M. W.; Smallcombe, S. H.; Whitaker, D. R.; Richards, J. H. *Biochemistry* **1973**, *12*, 4732−4743.
(147) Robillard, G.; Schulman, R. G. *J. Mol. Biol.* **1974**, *86*, 541−558.
(148) Bachovchin, W. W. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 7948−7951.
(149) Liang, T.; Abeles, R. H. *Biochemistry* **1987**, *26*, 7603−7608.
(150) Lin, J.; Westler, W. M.; Cleland, W. W.; Markley, J. L.; Frey, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14664−14668.
(151) Lin, J.; Cassidy, C. S.; Frey, P. A. *Biochemistry* **1998**, *37*, 11940−11948.
(152) Kossiakoff, A. A.; Spencer, S. A. *Nature* **1980**, *288*, 414−416.
(153) Gerlt, J. A.; Gassman, P. G. *Biochemistry* **1993**, *32*, 11943−11952.
(154) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887−1890.
(155) Warshel, A.; Papazyan, A.; Kollman, P. A. *Science* **1995**, *269*, 102−106.
(156) Guthrie, J. P. *Chem. Biol.* **1996**, *3*, 163−170.
(157) Shan, S. O.; Loh, S.; Herschlag, D. *Science* **1996**, *272*, 97−101.
(158) Ash, E. L.; Sudmeier, J. L.; De Fabo, E. C.; Bachovchin, W. W. *Science* **1997**, *278*, 1128−1132.
(159) Shan, S.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14474−14479.
(160) Bach, R. D.; Dmitrenko, O.; Glukhovtsev, M. N. *J. Am. Chem. Soc.* **2001**, *123*, 7134−7145.
(161) Hibbert, F.; Emsley, J. *Adv. Phys. Org. Chem.* **1990**, *26*, 255−279.
(162) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−1930.
(163) Cassidy, C. S.; Lin, J.; Frey, P. A. *Biochemistry* **1997**, *36*, 4576−4584.
(164) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. *J. Biol. Chem.* **1998**, *273*, 25529−25532.
(165) Cassidy, C. S.; Lin, J.; Frey, P. A. *Biochem. Biophys. Res. Commun.* **2000**, *273*, 789−792.
(166) Halkides, C. J.; Wu, Y. Q.; Murray, C. J. *Biochemistry* **1996**, *35*, 15941−15948.
(167) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629−3639.
(168) Corey, D. R.; McGrath, M. E.; Vasquez, J. R.; Fletterick, R. J.; Craik, C. S. *J. Am. Chem. Soc.* **1992**, *114*, 4905−4907.
(169) Stratton, J. R.; Pelton, J. G.; Kirsch, J. F. *Biochemistry* **2001**, *40*, 10411−10416.

(170) O'Sullivan, D. B.; O'Connell, T. P.; Mahon, M. M.; Koenig, A.; Milne, J. J.; Fitzpatrick, T. B.; Malthouse, J. P. *Biochemistry* **1999**, *38*, 6187−6194.
(171) Brady, K.; Liang, T.-C.; Abeles, R. H. *Biochemistry* **1989**, *28*, 9066−9077.
(172) Bachovchin, W. W. *Biochemistry* **1986**, *25*, 7751−7759.
(173) Dutler, H.; Bizzozero, S. A. *Acc. Chem. Res.* **1989**, *22*, 322−327.
(174) Kidd, R. D.; Sears, P.; Huang, D. H.; Witte, K.; Wong, C. H.; Farber, G. K. *Protein Sci.* **1999**, *8*, 410−417.
(175) Federov, A.; Shi, W.; Kicksa, G.; Federov, W.; Tyler, P. C.; Furneaux, R. H.; Hanson, J. C.; Gainsford, G. J.; Larese, J. Z.; Schramm, V. L.; Almo, S. C. *Biochemistry* **2001**, *40*, 853−860.
(176) Perona, J. J.; Craik, C. S.; Fletterick, R. J. *Science* **1993**, *261*, 620−622.
(177) Singer, P. T.; Smalas, A.; Carty, R. P.; Mangel, W. F.; Sweet, R. M. *Science* **1993**, *259*, 669−673.
(178) Singer, P. T.; Smalas, A.; Carty, R. P.; Mangel, W. F.; Sweet, R. M. *Science* **1993**, *261*, 621−622.
(179) Tonge, P. J.; Carey, P. R. *Biochemistry* **1990**, *29*, 10723−10727.
(180) Tonge, P. J.; Pusztai, M.; White, A. J.; Wharton, A. J.; Wharton, C. W.; Carey, P. R. *Biochemistry* **1991**, *30*, 4790−4795.
(181) Tonge, P. J.; Carey, P. R. *Biochemistry* **1992**, *31*, 9122−9125.
(182) Whiting, A. K.; Petricolas, W. L. *Biochemistry* **1994**, *33*, 552−561.
(183) White, A. J.; Wharton, C. W. *Biochem. J.* **1990**, *270*, 627−637.
(184) Bender, M. L.; Hamilton, G. A. *J. Am. Chem. Soc.* **1962**, *84*, 2570.
(185) Chang, T. K.; Chiang, Y. W.; Guo, H.-X.; Kresge, A. J.; Mathew, L.; Powell, M. F.; Wells, J. A. *J. Am. Chem. Soc.* **1996**, *118*, 8802−8807.
(186) Riechmann, L.; Kasche, V. *Biochem Biophys Res Commun* **1984**, *120*, 686−691.
(187) Riechmann, L.; Kasche, V. *Biochim Biophys Acta* **1985**, *830*, 164−172.
(188) Gololobov, M. Y.; Stepanov, V. M.; Voyushina, T. J.; Aldercreutz, P. *Eur. J. Biochem.* **1993**, *217*, 955−963.
(189) Brouwer, A. C.; Kirsch, J. F. *Biochemistry* **1982**, *21*, 1302−1307.
(190) Stone, S. R.; Betz, A.; Hofsteenge, J. *Biochemistry* **1991**, *30*, 9841−9848.
(191) Ayala, Y. M.; Di Cera, E. *Protein Sci.* **2000**, *9*, 1589−1593.
(192) Bauer, C.-A. *Biochim. Biophys. Acta* **1976**, *438*, 495−502.
(193) Coombs, G. S.; Dang, A. T.; Madison, E. L.; Corey, D. R. *J. Biol. Chem.* **1996**, *271*, 4461−4467.
(194) Wohl, R. C.; Summaria, L.; Robbins, K. G. *J. Biol. Chem.* **1980**, *255*, 2005−2013.
(195) Gibbs, C. S.; Coutre, S. E.; Tsiang, M.; Li, W.-X.; Jain, A. K.; Dunn, K. E.; Law, V. S.; Mao, C. T.; Matsumara, S. Y.; Mejza, S. J.; Paborsky, L. R.; Leung, L. L. K. *Nature* **1995**, *378*, 413−416.
(196) Kam, C.-M.; McRae, B. J.; Harper, J. W.; Niemann, M. A.; Volankis, J. E.; Powers, J. C. *J. Biol. Chem.* **1987**, *262*, 3444−3451.
(197) Taylor, F. R.; Bixler, S. A.; Budman, J. I.; Wen, D.; Karpusas, M.; Ryan, S. T.; Jaworski, G. J.; Safari-Fard, A.; Pollard, S.; Whitty, A. *Biochemistry* **1999**, *38*, 2849−2859.
(198) Christensen, U.; Ipsen, H.-H. *Biochim. Biophys. Acta* **1979**, *5569*, 177−183.
(199) Stein, R. L.; Viscarello, B. R.; Wildonger, R. A. *J. Am. Chem. Soc.* **1984**, *106*, 796−798.
(200) Hedstrom, L.; Szilagyi, L.; Rutter, W. J. *Science* **1992**, *255*, 1249−1253.
(201) Wang, E. C. W.; Hung, S.-H.; Cahoon, M.; Hedstrom, L. *Protein Eng.* **1997**, *10*, 405−411.
(202) Stein, R. L. *J. Am. Chem. Soc.* **1985**, *107*, 5767−5775.
(203) Brenner, S.; Fuller, R. S. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 922−926.
(204) Bravo, D. A.; Gleason, J. B.; Sanchez, R. I.; Roth, R. A.; Fuller, R. S. *J. Biol. Chem.* **1994**, *269*, 25830−25837.
(205) Rockwell, N. C.; Fuller, R. S. *Biochemistry* **2001**, *40*, 3657−3665.
(206) Procter, P.; Gensmantel, N. P.; Page, M. I. *J. Chem. Soc., Perkin Trans. 2* **1982**, 1185−1192.
(207) Pollack, R. M.; Bender, M. L. *J. Am. Chem. Soc.* **1970**, *92*, 7190−7194.
(208) Himoe, A.; Brandt, K. G.; DeSa, R. J.; Hess, G. P. *J. Biol. Chem.* **1969**, *244*, 3483−3493.
(209) Bender, M. L.; Kezdy, F. J.; Gunter, C. R. *J. Am. Chem. Soc.* **1964**, *86*, 3714−3721.
(210) Bernhard, S. A.; Gutfreund, H. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *53*, 1238−1243.
(211) Ascenzi, P.; Menegatti, E.; Guarneri, M.; Antonini, E. *Mol Cell Biochem* **1983**, *56*, 33−8.
(212) Berezin, I. V.; Kazanskaya, N. F.; Klyosov, A. A. *FEBS Lett.* **1971**, *15*, 121−124.
(213) Baggio, R.; Shi, Y.-Q.; Wu, Y.-Q.; Abeles, R. H. *Biochemistry* **1996**, *35*, 3351−3353.
(214) Harper, J. W.; Cook, R. R.; Roberts, C. J.; McLaughlin, B. J.; Powers, J. C. *Biochemistry* **1984**, *23*, 2995−3002.
(215) Stein, R. L. *Arch. Biochem. Biophys.* **1985**, *236*, 677−680.
(216) Thompson, R. C.; Blout, E. R. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *67*, 1734−1740.
(217) Tanizawa, K.; Kanaoka, Y.; Lawson, W. B. *Acc. Chem. Res.* **1987**, *20*, 337−343.
(218) Graf, L.; Jancso, A.; Szilagyi, L.; Hegyi, G.; Pinter, K.; Naray-Szabo, G.; Hepp, J.; Medzihradsky, K.; Rutter, W. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4961−4965.
(219) Hedstrom, L. *Biol. Chem.* **1996**, *377*, 465−471.
(220) Hedstrom, L.; Perona, J.; Rutter, W. J. *Biochemistry* **1994**, *33*, 3, 8757−8763.
(221) Szabo, E.; Bocskei, Z.; Naray-Szabo, G.; Graf, L. *Eur. J. Biochem.* **1999**, *263*, 20−26.
(222) Perona, J. J.; Hedstrom, L.; Wagner, R. L.; Rutter, W. J.; Craik, C. S.; Fletterick, R. *Biochemistry* **1994**, *33*, 3252−3259.
(223) Willett, W. S.; Gillmore, S. A.; Perona, J. J.; Fletterick, R. J.; Craik, C. J. *Biochemistry* **1995**, *34*, 2172−2180.
(224) Mace, J. E.; Wilk, B. J.; Agard, D. A. *J. Mol. Biol.* **1995**, *251*, 116−34.
(225) Mace, J. E.; Agard, D. A. *J. Mol. Biol.* **1995**, *254*, 720−736.
(226) Venekei, I.; Szilagyi, L.; Graf, L.; W. J., R. *FEBS Lett.* **1996**, *379*, 143−147.
(227) Hung, S.-H.; Hedstrom, L. *Protein Eng.* **1998**, *11*, 669−673.
(228) Willett, W. S.; Brinen, L. S.; Fletterick, R. J.; Craik, C. J. *Biochemistry* **1996**, *35*, 5992−5998.
(229) Kurth, T.; Grahn, S.; Thormann, M.; Ullman, D.; Hofman, H.-J.; Jakubke, H.-D.; Hedstrom, L. *Biochemistry* **1998**, *37*, 7.
(230) Hedstrom, L.; Farr-Jones, S.; Kettner, C.; Rutter, W. J. *Biochemistry* **1994**, *33*, 8764−8769.
(231) Caputo, A.; James, M. N. G.; Powers, J. C.; Hudig, D.; Bleackley, R. C. *Nat. Struct. Biol* **1994**, *1*, 364−367.
(232) Tobin, J. B.; Whitt, S. A.; Cassidy, C. S.; Frey, P. A. *Biochemistry* **1995**, *34*, 6919−24.
(233) Inagami, T.; Murachi, T. *J. Biol. Chem.* **1964**, *239*, 1395−1401.
(234) Ascenzi, P.; Menegatti, E.; Guarneri, M.; Bortolotti, F.; Antonini, E. *Biochemistry* **1982**, *21*, 2483−2490.
(235) Robertus, J. D.; Kraut, J.; Alden, R. A.; Birktoft, J. J. *Biochemistry* **1972**, *11*, 4293−4318.
(236) Read, R. J.; Fujinaga, M.; Sielecki, A. R.; James, M. N. G. *Biochemistry* **1983**, *22*, 4420−4433.
(237) Katz, B. A.; Finer-Moore, J.; Mortezaei, R.; Rich, D. H.; Stroud, R. M. *Biochemistry* **1995**, *34*, 8264−8280.
(238) Satterthwait, A. C.; Jencks, W. P. *J. Am. Chem. Soc.* **1974**, *96*, 7018−7031.
(239) Dufton, M. J. *FEBS Lett.* **1990**, *271*, 9−13.
(240) Fersht, A. R. *Enzyme Structure and Mechanism*; 2nd ed.; W. H. Freeman and Co.: New York, 1985.
(241) Herschlag, D. *Bioorg. Chem.* **1988**, *16*, 62−96.
(242) Pasternak, A.; White, A.; Jeffrey, C.; Medina, N.; Cahoon, M.; Ringe, D.; Hedstrom, L. *Protein Sci.* **2001**, *10*, 1331−1342.
(243) Hedstrom, L.; Lin, T.-Y.; Fast, W. *Biochemistry* **1996**, *35*, 4515−4523.
(244) Pasternak, A.; Liu, X.; Lin, T.-Y.; Hedstrom, L. *Biochemistry* **1998**, *37*, 16201−16210.
(245) Volankis, J. E.; Narayana, S. V. L. *Protein Sci.* **1996**, *5*, 553−564.
(246) Narayana, S. V.; Carson, M.; el-Kabbani, O.; Kilpatrick, J. M.; Moore, D.; Chen, X.; Bugg, C. E.; Volanakis, J. E.; DeLucas, L. J. *J. Mol. Biol.* **1994**, *235*, 695−708.
(247) Kim, S.; Narayana, S. V.; Volanakis, J. E. *J. Biol. Chem.* **1995**, *270*, 24399−24405.
(248) Post, C. B.; Ray, J. W. J. *Biochemistry* **1995**, *34*, 15881−15885.
(249) Bone, R.; Silen, J. L.; Agard, D. A. *Nature* **1989**, *339*, 191−195.
(250) Davis, J. H.; Agard, D. A. *Biochemistry* **1998**, *37*, 7696−7707.
(251) Bone, R.; Fujishige, A.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1991**, *30*, 10388−10398.
(252) Rader, S. D.; Agard, D. A. *Protein Sci.* **1997**, *6*, 1375−1386.
(253) Ota, N.; Agard, D. A. *Protein Sci.* **2001**, *10*, 1403−1414.
(254) Miller, D. W.; Agard, D. A. *J. Mol. Biol.* **1999**, *286*, 267−278.
(255) Dang, Q. D.; Vindigni, A.; Di Cera, E. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5977−5981.
(256) Di Cera, E.; Canwell, A. M. *Ann. N. Y. Acad. Sci.* **2001**, *936*, 133−146.
(257) Wells, C. M.; Di Cera, E. *Biochemistry* **1992**, *31*, 11721−11730.
(258) Dang, Q. D.; Di Cera, E. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10653−10656.
(259) Guinto, E. R.; Vindigni, A.; Ayala, Y. M.; Dang, Q. D.; Di Cera, E. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11185−11189.
(260) Thompson, R. C.; Blout, E. R. *Biochemistry* **1973**, *12*, 57−65.
(261) Thompson, R. C.; Blout, E. R. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *67*, 1734−1740.
(262) Cannon, W. R.; Benkovic, S. J. *J. Biol. Chem.* **1998**, *273*, 26257−26260.
(263) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035−27038.
(264) Jencks, W. P. *Adv. Enzymol.* **1975**, *43*, 219−410.
(265) Di Cera, E.; Dang, Q. D.; Ayala, Y. M. *Cell Mol. Life Sci.* **1997**, *53*, 701−730.
(266) Esmon, C. T. *Biochim. Biophys. Acta* **2000**, *1477*, 349−360.
(267) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.

CR000033X