Articles

Structural Basis for Broad Specificity in α-Lytic Protease Mutants[†,‡]

Roger Bone,[§,∥] Amy Fujishige,[§] Charles A. Kettner,[⊥] and David A. Agard*[,§]

Departments of Biochemistry and Biophysics and of Pharmaceutical Chemistry and The Howard Hughes Medical Institute, University of California at San Francisco, San Francisco, California 94143-0448, and Du Pont Merck Pharmaceutical Company, Du Pont Experimental Station, Wilmington, Delaware 19880-0328

Received February 25, 1991; Revised Manuscript Received June 25, 1991

ABSTRACT: Binding pocket mutants of α-lytic protease (Met 192 → Ala and Met 213 → Ala) have been constructed recently in an effort to create a protease specific for Met just prior to the scissile bond. Instead, mutation resulted in proteases with extraordinarily broad specificity profiles and high activity [Bone, R., Silen, J. L., & Agard, D. A. (1989) *Nature 339*, 191–195]. To understand the structural basis for the unexpected specificity profiles of these mutants, high-resolution X-ray crystal structures have been determined for complexes of each mutant with a series of systematically varying peptidylboronic acids. These inhibitory analogues of high-energy reaction intermediates provide models for how substrates with different side chains interact with the enzyme during the transition state. Fifteen structures have been analyzed qualitatively and quantitatively with respect to enzyme–inhibitor hydrogen-bond lengths, buried hydrophobic surface area, unfilled cavity volume, and the magnitude of inhibitor accommodating conformational adjustments (particularly in the region of another binding pocket residue, Val 217A). Comparison of these four parameters with the K_i of each inhibitor and the k_{cat} and K_m of the analogous substrates indicates that while no single structural parameter consistently correlates with activity or inhibition, the observed data can be understood as a combination of effects. Furthermore, the relative contribution of each term differs for the three enzymes, reflecting the altered conformational energetics of each mutant. From the extensive structural analysis, it is clear that enzyme flexibility, especially in the region of Val 217A, is primarily responsible for the exceptionally broad specificity observed in either mutant. Taken together, the observed patterns of substrate specificity can be understood to arise directly from interactions between the substrate and the residues lining the specificity pocket and indirectly from interactions between peripheral regions of the protein and the active-site region that serve to modulate active-site flexibility.

One of the fundamental functions of an enzyme is to provide specificity by limiting the range of substrates which are catalytically productive. Since catalysis is the result of an enzyme's ability to bind the transition state for a reaction (Wolfenden, 1972; Fersht, 1985), specificity must derive from the selective binding of the transition states for the reactions of favored substrates. In empirical terms, the selection of substrates must result from the balance between energetically favorable intermolecular interactions (hydrogen bonds, solvent exclusion from hydrophobic surfaces, electrostatics, van der Waals interactions) and energetically costly consequences of accommodating substrates in an imprecisely tailored active site (hydrogen-bond distortion or loss, cavity formation, induced conformational changes). Functional groups participating in these interactions, and estimates for their energetic contributions, can be identified by studying how catalytic efficiency is influenced by the systematic variation of substrate and enzyme structure (Fersht, 1988; Fersht et al., 1985; Estell et al., 1986). However, a thorough understanding of how the favorable and unfavorable interactions of important structural elements, contributed by both the enzyme and substrate, strike a balance to produce a particular pattern of specificity can only be achieved through direct structural analysis.

Recently the specificity of α-lytic protease, an extracellular serine protease of *Lysobacter enzymogenes* (Whitaker, 1970), has been examined by using a combination of crystallographic and mutational analyses. On small peptide substrates, sAAP-X-*p*-Na,[1] where sAAP is succinyl-Ala-Ala-Pro- and X is Ala, Val, Met, Leu, or Phe, the primary specificity of the enzyme is for Ala in the P_1[2] position, with activity falling off dramatically as the size of the P_1 substrate side chain increases (Table II; Bone et al., 1989a; Delbaere et al., 1981). Crystallographic analysis of a series of complexes between wild-type α-lytic protease and inhibitory peptide boronic acids, which mimic the transition state for substrate hydrolysis, revealed why substrates with P_1 side chains increasing in size are se-

[†] This work was supported by funds from the Howard Hughes Medical Institute and from an NSF Presidential Young Investigator grant (D. A.A.). R.B. was supported by NIH National Research Service Award GM11174-02.

[‡] The coordinates for the structures reported in this paper have been submitted to the Brookhaven Protein Data Bank.

* Author to whom correspondence should be addressed.

[§] University of California at San Francisco.

[∥] Present address: Department of Biophysical Chemistry, Merck Sharp & Dohme Research Laboratories, Box 2000 (RY80M-136), Rahway, NJ 07065.

[⊥] Du Pont Merck Pharmaceutical Co.

[1] Abbreviations: sAAP-, succinyl-Ala-Ala-Pro-; mAAP-, methoxysuccinyl-Ala-Ala-Pro-, bAP, *tert*-butyloxycarbonyl-Ala-Pro-; MeOSuc-, methoxysuccinyl-; Tris, tris(hydroxymethyl)aminomethane; *p*-Na, *p*-nitroaniline; Norleu, norleucine; rms, root mean square. The prefix boro of -boroVal- indicates that the carbonyl of the amino acid residue, in this case a valyl residue, is replaced by $-B(OH)_2$; the systematic name for valine boronic acid would be 1-amino-2-methylpropylboronic acid.

[2] In protease substrate nomenclature the residue to the N-terminal side of the scissile bond is the P_1 residue, the next residue toward the N terminus is the P_2 residue, etc. (Schecter & Berger, 1967).

0006-2960/91/0430-10388$02.50/0

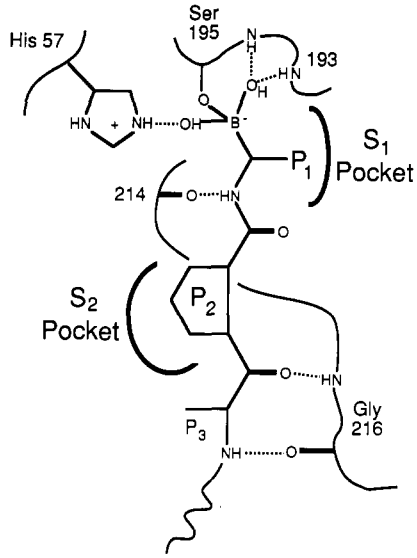

FIGURE 1: Schematic of the interactions between α-lytic protease and inhibitors. Six intermolecular hydrogen bonds stabilize the complex, as do electrostatic interactions between the negatively charged boronic acid and the positively charged histidine and hydrophobic interactions in the S_1 and S_2 specificity pockets. The primary specificity pocket of α-lytic protease (S_1) is lined by the side chains of residues Met 192, Met 213, and Val 217A.

lected against (Figure 1; Bone et al., 1989b). As the size of the P_1 side chain increases, hydrogen bonds become distorted and conformational shifts of greater magnitude are required for accommodation of the P_1 side chain. Since there is no compensating increase in the amount of hydrophobic surface area buried upon inhibitor interaction with the enzyme, binding of inhibitors with large P_1 side chains is unfavorable. It appeared that specificity was restricted to substrates with Ala in the P_1 position because of the presence in the primary specificity pocket of two Met residues (Met 192 and Met 213)[3] that block what would otherwise be a large hydrophobic pocket.

To test this hypothesis, each of the two Met residues was mutated to Ala (Bone et al., 1989a). It was reasoned that the resulting mutants would have specificity for substrates with Met in the P_1 position, which would complement the Met → Ala mutations. Remarkably, the mutagenesis resulted in two proteases having extraordinarily broad specificity, one of which retained very high activity. In both cases, activity did not decrease as expected for substrates with small side chains in the P_1 position. In addition, surprising increases in activity were observed for substrates with the largest and bulkiest P_1 residues (Leu, Phe). Initial structural analyses of these mutants suggested that one of these enzymes (Met 192 → Ala) was able to utilize conformational changes in a productive manner to accommodate a substrate with Phe in the P_1 position (Bone et al., 1989a). Furthermore, the structures of the free mutants revealed that the enzymes had unexpectedly small hydrophobic pockets.

In this work we seek to understand how these mutant proteases are able to expand to accommodate side chains such as Leu and Phe in relatively small specificity pockets and how side chains as diverse as Ala and Phe can fit well into the same specificity pocket with little loss in activity. To answer this question we have used X-ray crystallography to determine the high-resolution structures of 10 enzyme–inhibitor complexes between the mutant α-lytic proteases and a series of inhibitory peptide boronic acids. These inhibitors provide excellent mimics of the transition state for substrate hydrolysis (Bone et al., 1987, 1989b), and systematic variation of the boronic acid side chain allows structural analysis to parallel kinetic studies on enzyme specificity. When combined with previously determined crystal structures, the resulting data suggest a model for how a series of substrates interact with wild-type and mutant enzymes, how interactions differ between these enzymes, and how each enzyme changes in structure to accommodate each inhibitor.

[3] Residues in α-lytic protease are numbered by homology with chymotrypsin (Fujinaga et al., 1985) and range from 15A to 244.

Experimental Procedures

Cloned wild-type and mutant α-lytic proteases were purified from culture filtrates of *Escherichia coli* as previously described (Silen et al., 1989; Hunkapiller et al., 1973; Whitaker, 1970) and migrated as single bands when subjected to low-pH native polyacrylamide gel electrophoresis (Hames & Rickwood, 1981). sAAP-Gly-*p*-Na and sAAP-Norleu-*p*-Na were purchased from Bachem Inc. mAAP-boroAla, mAAP-boroVal, mAAP-boroNorleu, mAAP-boroLeu, and mAAP-boroPhe were synthesized as previously described (Kettner et al., 1988; Kettner & Shenvi, 1984; Bone et al., 1989b).

Values of k_{cat} and K_M for these substrates were determined from double-reciprocal plots of initial reaction velocity as a function of substrate concentration, and errors were estimated as described by Bevington (1969). Reactions, initiated by the addition of enzyme, contained substrate and 0.1 M Tris-HCl (pH = 8.0) and were monitored continuously at 410 nm for the production of *p*-nitroaniline ($\Delta\epsilon$ = 8860 M^{-1} cm^{-1}; Hunkapiller et al., 1976). The K_i values for all inhibitors, except those noted below, were determined by monitoring reaction velocity as a function of inhibitor concentration at substrate concentrations far below the K_M ([S] = $K_M/20$). Under these conditions the K_i value is the reciprocal slope of a plot of V_0/V_i as a function of inhibitor concentration, [I], where V_0 is the initial reaction velocity in the absence of inhibitor and V_i is the initial reaction velocity in the presence of inhibitor at concentration [I].

K_i values for the binding of mAAP-boroVal, mAAP-boroNorleu, and mAAP-boroLeu by Met 192 → Ala were calculated from the rate constants for association and dissociation of the complex, which were determined as described by Williams and Morrison (1979) for very slow tight-binding inhibitors. These studies were done in the presence of 0.1 mg/mL polylysine (average MW 300000, Sigma Chemical Co.) to stabilize the enzyme at very low concentrations ([E] < 4 nM). Values of K_i determined from the ratio of k_{off} to k_{on}, assuming that enzyme and inhibitor combine slowly to form the final E–I complex, were checked by calculating K_i values from the extent to which preformed E–I complexes reverse to yield free enzyme. For the binding of mAAP-boroVal and mAAP-boroPhe (Bone et al., 1989a; R. Bone, unpublished results) by Met 192 → Ala, these values matched k_{off}/k_{on}. However, for the binding of mAAP-boroNorleu and mAAP-boroLeu by Met 192 → Ala, there were discrepancies of factors of 10 and 3 in the apparent K_i values, suggesting that binding was tighter than the ratio of k_{off} to k_{on}. Such discrepancies may be indicative of an even slower isomerization step that follows slow association of the enzyme and inhibitor with apparent k_{off} values representing inhibitor dissociation from the relatively small portion of E–I complex present in unisomerized form. Therefore, the apparent K_i values for these two complexes should be considered upper estimates of the true K_i values.

Crystals of α-lytic protease and mutants were grown from 1.3 M lithium sulfate containing Tris–sulfate (20 mM, pH 8.0)

Table I: Crystallography of α-Lytic Protease–Inhibitor Complexes

structure[a]	[I][b] (mM)	resoln[c] (Å)	R^d	% reflcns $I > 3\sigma_I$[e]	rms deviations[f] bond (Å)	rms deviations[f] angle (Å)
wild type + mAAP-boroLeu	50	2.20	0.129	92	0.017	0.053
MA192 + mAAP-boroAla	150	2.12	0.136	89	0.017	0.054
MA192 + mAAP-boroVal	40	2.25	0.127	87	0.016	0.053
MA192 + mAAP-boroNorleu	50	2.15	0.138	91	0.017	0.052
MA192 + mAAP-boroLeu	25	2.10	0.136	87	0.018	0.055
MA213 + mAAP-boroAla	75	2.13	0.134	93	0.020	0.045
MA213 + mAAP-boroNorleu	25	2.10	0.132	92	0.018	0.053
MA213 + mAAP-boroLeu	75	2.05	0.136	82	0.019	0.057
MA213 + mAAP-boroPhe	100	2.25	0.132	82	0.020	0.045

[a] The boro prefix indicates that inhibitors are the α-amino boronic acid analogues of the corresponding amino acids. [b] Final inhibitor concentration during crystal soaking. [c] Highest resolution of the data. [d] Conventional crystallographic R factor. [e] Percentage of the data collected with intensities greater than $3\sigma_I$; all of the data were used in structure determination and refinement. [f] Root mean square deviations from ideal bond and angle distances.

Table II: Values of Kinetic Parameters for Wild-Type and Mutant α-Lytic Proteases[a]

P_1 residue	wild type k_{cat} (s^{-1})	wild type K_M (mM)	wild type k_{cat}/K_M (s^{-1} M^{-1})	wild type K_i (nM)	MA192 k_{cat} (s^{-1})	MA192 K_M (mM)	MA192 k_{cat}/K_M (s^{-1} M^{-1})	MA192 K_i (nM)	MA213 k_{cat} (s^{-1})	MA213 K_M (mM)	MA213 k_{cat}/K_M (s^{-1}) M^{-1})	MA213 K_i (nM)
Gly	12 (2)	35 (10)	330 (25)		0.46 (0.03)	15 (1)	32 (1.3)		3.1 (0.4)	40 (5)	78 (1.3)	
Ala	75[b]	3.6[b]	21000[b]	67[d]	37[c]	3.6[c]	10000[c]	64 (2)	34[c]	57[c]	600[c]	270 (3)
Val	13[b]	16[b]	790[b]	6.4[d]	3.4[c]	1.1[c]	3000[c]	1.3[e] (0.1)	10[c]	29[c]	340[c]	210[c]
Norleu	3.7 (0.07)	49 (10)	75 (2)	1100[b]	140 (3)	0.14 (0.01)	970000 (40000)	0.26[e] (0.01)	26 (3)	16 (2)	1600 (30)	50 (1.3)
Met	56[b]	31[b]	1800[b]		120[c]	0.33[c]	350000[c]				980[c]	
Leu	1.2[b]	290[b]	4.1[b]	2000 (20)	87[c]	0.77[c]	110000[c]	0.58[e] (0.03)			160[c]	66 (1)
Phe	0.0068[b]	17[b]	0.38[b]	540[d,f]	130[c]	0.40[c]	310000[c]	0.60[c]	47[c]	16[c]	340[c]	240 (4)

[a] The substrates used were sAAP-X-*p*-Na, where sAAP- is succinyl-Ala-Ala-Pro-, X is Gly, Ala, Val, Norleu, Met, Leu, or Phe, and *p*-Na is *p*-nitroaniline. The inhibitors used were mAAP-boroX, where mAAP- is methoxysuccinyl-Ala-Ala-Pro- and boroX is the α-amino boronic acid analogue of Ala, Val, Norleu, Leu, or Phe. For values determined in this work, standard deviations (σ; shown in parentheses after the values) were calculated as described by Bevington (1969). [b] Bone et al. (1989b). [c] Bone et al. (1989a). [d] Kettner et al. (1988). [e] Dissociation constants were taken to be k_{off}/k_{on}. [f] This inhibitor does not form a tetrahedral adduct with the enzyme (Bone et al., 1989b).

at ambient temperature (Bone et al., 1987, 1989a; Brayer et al., 1979b). To prepare enzyme–inhibitor complexes, inhibitor solutions (0.1–0.5 M) in water were added in small aliquots (0.25–1.0 μL) to vapor diffusion droplets (5 μL) containing 1–3 crystals of α-lytic protease. Generally, several aliquots of inhibitor solution were added to the same crystal drop, allowing 10–24 h of equilibration between additions, to achieve the final inhibitor concentration (Table I). Protease crystals remained undamaged by this procedure.

Data from crystals of enzyme–inhibitor complexes were collected from single crystals by using either Syntex P21 or Rigaku AFC5 automated diffractometers equipped with graphite monochromators (Stroud et al., 1974; Wycoff et al., 1967). Obtaining high-resolution data generally required that diffraction data from two or three crystals be merged. To date, all mutants and enzyme–inhibitor complexes have crystallized isomorphously in the space group $P3_221$, showing only small deviations (less than 0.15%) from the unliganded wild-type cell dimensions ($a = b = 66.3$ Å, $c = 80.1$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$). The intensities of seven check reflections were monitored in order to correct for crystal decay, which was less than 25% over the course of data collection. Corrections were also made for absorption, the effect on intensities described by Lorentzian geometry, and polarization by using standard methods and backgrounds according to Krieger et al. (1974). Data were collected by ω scan in shells of 2θ with the scan rate adjusted so that at least 81% of the reflections had intensities greater than $3\sigma_I$ over the entire data set (Table I).

Initial difference Fourier maps were computed by using structure factors and phases calculated from the refined coordinates of the complex between α-lytic protease and mAAP-boroVal (Bone et al., 1989b) from which the inhibitor had been removed. Maps were inspected and the inhibitors were placed by using the interactive graphics package FRODO

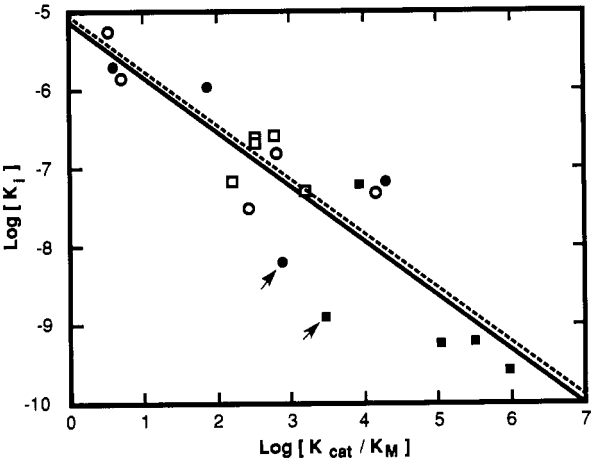

FIGURE 2: Plot of log K_i vs log (k_{cat}/K_m) for wild-type and mutant α-lytic proteases: Closed circles, wild type (Bone et al., 1989a); open circles, VA217A (Bone et al., manuscript in preparation); closed squares, MA192 (Bone et al., 1989a); open squares, MA213 (Bone et al., 1989a). Substrates/inhibitors used were sAAP-X-*p*-Na/mAAP-boroX, where X is Ala, Val, Ile (wild type only) Norleu, Leu, or Phe (except wild type) for each enzyme. The slope of the line is 0.7 with a correlation coefficient of 0.87. Without the arrowed data points, corresponding to substrates and inhibitors with Val in the P_1 position, the slope is unchanged and the correlation coefficient rises to 0.92.

(Jones, 1982), after which the coordinates were refined by the stereochemically restrained least-squares algorithm of Hendrickson and Konnert (1981), modified for multiple occupancy refinement (Smith et al., 1988), adapted for use on the FPS 264 array processor (Furey, 1984) and further modified by us. The final overall crystallographic R factors for the comparison of observed and calculated structure factors are listed in Table I. Average α-lytic protease temperature factors were

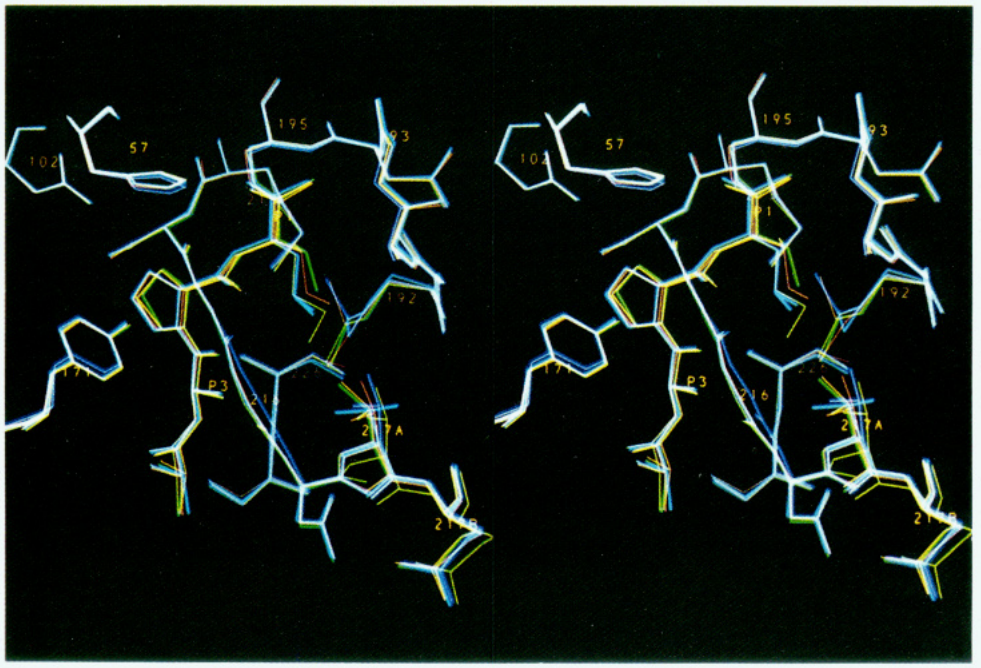

FIGURE 3: Stereo drawing of the superimposed structures of the specificity pocket regions of wild-type α-lytic protease complexes with each of five boronic acid inhibitors: purple, uncomplexed; orange, mAAP-boroAla; green, mAAP-boroVal; red, mAAP-boroIle; yellow, mAAP-boroNorleu; and blue, mAAP-boroLeu. The enzyme portion of each complexed structure was superimposed on the uncomplexed enzyme so as to minimize the rms deviation between α-carbon positions. Alternate P_1 side chains are accommodated via adjustments in the region of Val 217A and with alternate conformations of Met 192.

11 Å^2 while inhibitor temperature factors averaged 16.5 Å^2, increasing from the P_1 to the P_4 residue. Each structure also contained one sulfate and 160–180 molecules of solvent. Surface area calculations (Table III) were done with the molecular surfacing program MS (Connolly, 1983a,b). The final refined coordinates for the structures reported in this work have been submitted to the Brookhaven Protein Data Bank.

Results

The experiments undertaken in this report represent an effort to directly correlate structure with function. Toward this end, the structures of enzyme–inhibitor complexes have been determined that correspond to each enzyme–substrate complex (except Gly and Met) during the transition state. Previously determined specificity profiles have been expanded to include two new substrates (with Gly and Norleu in P_1). The question of how well inhibitors mimic the transition state is addressed, and the general features and energetics of inhibitor binding are examined followed by brief descriptions of the structures.

Kinetics. The specificity profiles of wild-type and mutant α-lytic proteases (Bone et al., 1989a) were expanded to include values of k_{cat}, K_m, and k_{cat}/K_m for sAAP-Gly-*p*-Na and sAAP-Norleu-*p*-Na (Table II). The effect of reducing the size of the P_1 side chain from Ala to Gly is a reduction in activity by a factor of 310 for Met 192 → Ala, 64 for the wild-type enzyme, and 7.7 for Met 213 → Ala. Met 192 → Ala stabilizes the transition state of the substrate with Ala in the P_1 position by 3.4 kcal/mol more than it does the substrate with a P_1 Gly, a value approaching the highest incremental binding energy observed for selection of a methyl group over a hydrogen (Fersht, 1985). The potent Ala/Gly selectivity of Met 192 → Ala cannot be simply a function of the composition or volume of the specificity pocket, else Met 213 → Ala or the wild-type enzyme would exhibit better Ala/Gly selectivity. Structural analysis of enzyme–inhibitor complexes is necessary in order to understand what structural properties give rise to such good discrimination between methyl and hydrogen substituents.

Because Norleu boronic acids have been used as analogues of Met in inhibitor studies, the activity of α-lytic proteases toward sAAP-Norleu-p-Na was examined. Though the volumes of Norleu, Met, and Leu side chains are approximately the same (Zamyatin, 1972), the activity of wild type α-lytic protease toward substrates with these side chains differs considerably (Table II). For example, wild-type α-lytic protease has 24-fold higher activity toward sAAP-Met-*p*-Na than toward the corresponding Norleu substrate. Because the activity of the enzyme toward sAAP-Norleu-*p*-Na fits the general downward trend in activity as a function of increasing P_1 side-chain bulkiness (Table II) and because the activity toward sAAP-Met-*p*-Na does not, it appears that sAAP-Met-*p*-Na is an anomalously good substrate for the enzyme. One possible explanation for these results is that C–S–C bond angles are easier to deform, allowing the Met side chain to avoid steric contacts that reduce activity toward the substrate with Norleu in the P_1 solution. In contrast, both mutant enzymes prefer slightly the Norleu substrate over the Met substrate, with Met 192 → Ala and sAAP-Norleu-*p*-Na having the highest activity of any α-lytic protease–substrate pair ($k_{cat}/K_M = 10^6$). The Met-NorLeu anomaly is abolished in the mutants because these substrate side chains are expected to bind in more extended conformations (see below) that do not require distortion of the C–S–C bond angle. A clear implication of these results is that caution must be observed in extending to Met substrates those conclusions drawn from the structures of enzyme–inhibitor complexes with Norleu boronic acids.

K_i values for inhibitors corresponding in structure to substrate are listed in Table II. For a homologous series of structurally varying transition-state analogues and substrates, a linear relationship is expected in plots of log K_i vs log (k_{cat}/K_m) (Bartlett & Marlowe, 1983; Westerik & Wolfenden, 1972). For α-lytic protease, such a linear relationship is observed between inhibition and activity (Figure 2) upon vari-

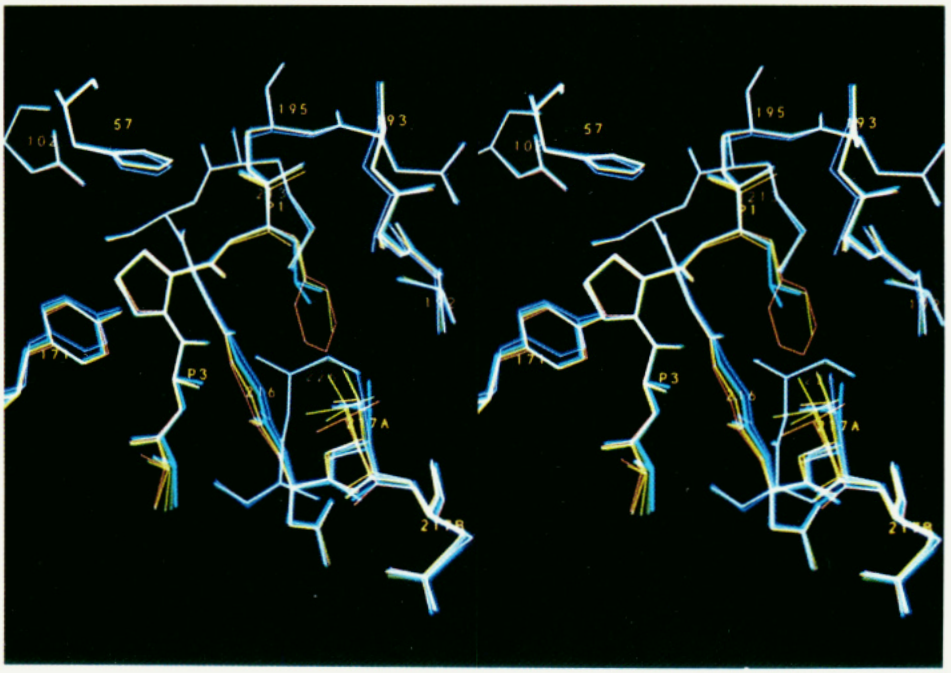

FIGURE 4: Stereo drawing of the superimposed structures of the specificity pocket regions of MA192 complexes with each of five boronic acid inhibitors: purple, uncomplexed; orange, mAAP-boroAla; green, mAAP-boroVal; yellow, mAAP-boroNorleu; blue, mAAP-boroLeu; and red, mAAP-boroPhe. The enzyme portion of each complexed structure was superimposed on the uncomplexed enzyme so as to minimize the rms deviation between α-carbon positions. Alternate P_1 side chains are accommodated via adjustments in the regions of Val 217A and Gly 216.

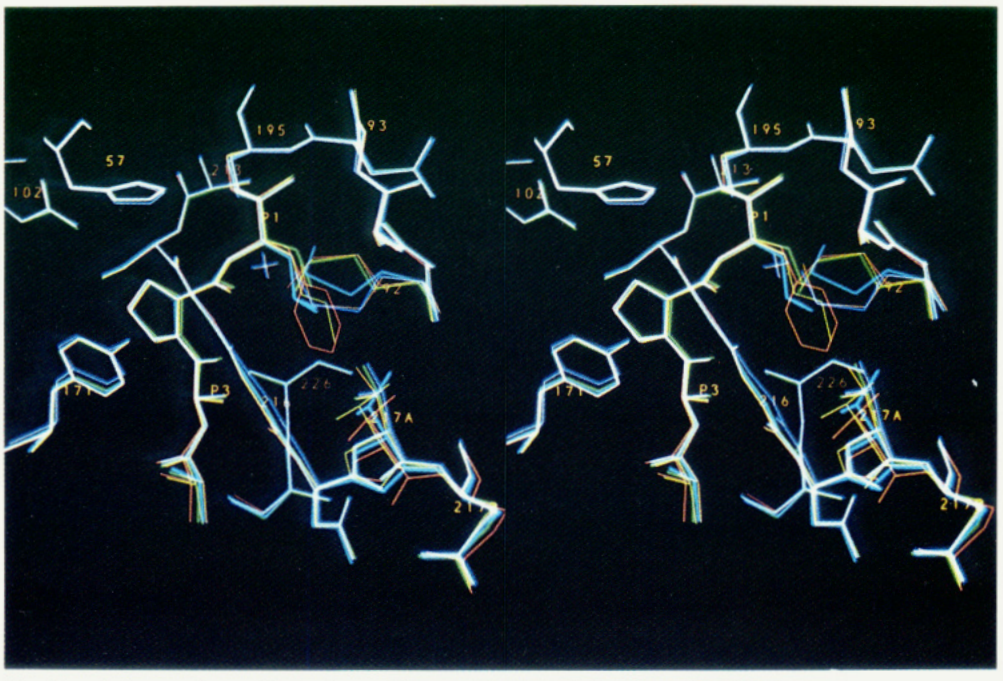

FIGURE 5: Stereo drawing of the superimposed structures of the specificity pocket regions of MA213 complexes with each of five boronic acid inhibitors: purple, uncomplexed; orange, mAAP-boroAla; green, mAAP-boroVal; yellow, mAAP-boroNorleu; blue, mAAP-boroLeu; and red, mAAP-boroPhe. The enzyme portion of each complexed structure was superimposed on the uncomplexed enzyme so as to minimize the rms deviation between α-carbon positions. Alternate P_1 side chains are accommodated via adjustments in the region of Val 217A and with alternate conformations of Met 192.

ation of both the enzyme (mutants) and substrate (alternate P_1 amino acids). The correlation coefficient for the least-squares analysis of 19 enzyme–inhibitor and enzyme–substrate pairs, which includes values for a Val 217A → Ala mutant (R. Bone, H. D. Madhani, and D. A. Agard, manuscript in preparation), is 0.87 with a slope of −0.70. If the two most deviant data points, corresponding to interaction of the wild-type enzyme and Met 192 → Ala with substrate/inhibitor pairs with a Val side chain in the P_1 position (Figure 2, arrows), are omitted, the correlation coefficient increases to 0.92 with a slope of −0.69. The slope of −0.7 indicates that enzyme–substrate interactions are more sensitive to changes in structure than corresponding enzyme–inhibitor interactions. Correlations with similar slopes have been observed for elastase (peptide aldehydes; Thompson, 1973) and chymotrypsin (trifluoroketones; Brady & Abeles, 1990). Transition-state analogue inhibitors of serine proteases appear to be very good but not exact analogues of the transient, high-energy transition state for substrate hydrolysis. An alternate possibility is that for some enzyme–substrate pairs the rate-determining step for

Table III: Structural Characterization of Wild-Type and Mutant α-Lytic Proteases

inhibitor P_1 residue mAAP-boroX	avg. H-bond length[a] (Å)	buried surface area[b] total (Å²)	buried surface area[b] hydrophobic (Å²)	cavity volume[c] (Å³)	rms deviation[d] (Å)	torsional shift from free enzyme ΔΨ217−ΔΦ217A[e] (deg)	RT ln (k_{cat}/K_m) (kcal/mol)[h]
				Wild Type			
Ala[f]	2.79	500	300	58	0.195	−6.5	5.9
Val[f]	2.88	550	350	18	0.213	0.0	4.0
Ile[f]	2.92	540	340	3.8	0.227	7.2	
Norleu[f]	2.95	540	340	10	0.230	53	2.6
Leu	2.96	540	340	16	0.232	−1.0	0.84
				MA192			
Ala	2.86	490	300	69	0.177	−6.0	5.5
Val	2.89	530	340	29	0.178	−1.0	4.7
Norleu	2.87	650	420	58	0.198	−5.3	8.1
Leu	2.92	550	350	36	0.184	3.7	6.9
Phe[g]	2.89	660	430	42	0.288	11.4	7.5
				MA213			
Ala	2.91	480	290	97	0.176	3.4	3.8
Val[g]	2.93	560	360	36	0.162	5.3	3.5
Norleu	2.90	610	380	53	0.261	3.8	4.4
Leu	2.90	550	350	27	0.224	7.2	3.0
Phe	2.92	620	400	34	0.308	46	3.5

[a] Average of the six hydrogen bonds stabilizing the enzyme inhibitor complex: Ser 195 N---Ala P_3 O, and Gly 216 O---Ala-P_3 N (Bone et al., 1987). [b] Buried surface area calculated according to Connolly (1983a,b). [c] Cavity volumes were calculated by subtracting the sum of the volumes of the side chains projecting into the specificity pocket (P_1, 192, 213, 217A) from the volume of the specificity pockets (Bone & Agard, manuscript in preparation). The volume of a molecule of solvent was subtracted from the cavity volume for the MA213 complex with mAAp-boroAla. [d] rms deviation of all atoms (N, Cα, C, O) from the unliganded wild-type enzyme structure for residues 166–232, which comprise most of the active-site region (Fujinaga et al., 1985; Bone et al., 1989a). [e] Conformational change in the region of Val 217A expressed in degrees as the deviation in torsional angles ($\Psi 217_L - \Psi 217_U$) − ($\Phi 217A_U$) where L and U refer to the liganded and unliganded forms of the enzyme. [f] Structures determined by Bone et al. (1989a). [g] Structures determined by Bone et al. (1989b). [h] For the substrate sAAP-X-*p*-Na.

the reaction may have changed. Inhibitors might not be sensitive to subtle differences in the way in which such substrates might interact with the enzyme.

General Aspects of Inhibitor Binding. Difference electron density maps clearly showed that all inhibitory boronic acids were bound by the enzyme at a single site with high occupancy. Maps calculated with Fourier coefficients ($2|F_o| - |F_c|$) showed continuous electron density stretching between the active-site Ser (195 O_γ) and the boron of the inhibitors, indicating that in each complex a covalent adduct had formed with the active-site serine. This was corroborated by the bond lengths found on refinement. On the basis of the shape of the electron density about the boron and bond angles after refinement, the geometry of the adducts was slightly distorted from tetrahedral[4] (Figures 3–5). No electron density is observed after the P_4 Ala residue for any of the inhibitors, indicating that the methoxysuccinyl group is disordered.

In qualitative terms, the intermolecular interactions that stabilize the E–I complex are essentially identical with those that stabilize the complex formed between α-lytic protease and bAP-boroVal [Figure 1; described in detail in Bone et al. (1987)]. The inhibitors make the same six hydrogen bonds with the enzyme (two in the oxyanion binding pocket with the amide groups of Ser 195 and Gly 193, one with His-57, and three with main-chain elements of residues 214–216) and have nearly the same interatomic distances between hydrogen-bonding atoms (Figures 3–5). These inhibitors also make van der Waals contacts with the side chains of Met 192, Met 213, and Val 217A and the main chains of residues 214–216 and 192–193 in the primary specificity pocket and with the side chains of Tyr 171, Phe 94, and His 57 in the P_2 binding site. Conformational changes outside of the specificity pocket, most notably in the P_2 binding site where residues 169–171 adjust away from the enzyme to accommodate the inhibitor, are very similar to those that have been characterized previously (Bone et al., 1987, 1989b). The combination of hydrogen bonds, electrostatic interactions between the positively charged histidine and the negatively charged boron (Bachovchin et al., 1988), and the removal of hydrophobic surfaces in the P_1 and P_2 specificity sites stabilizes these complexes and presumably also the transition state for substrate hydrolysis.

Where this series of structures differs is in the details of the interactions of the P_1 side chain with the specificity pocket and how these interactions affect the energetics of enzyme–inhibitor complex formation. In addition to qualitative descriptions of how the mutations have affected interactions with inhibitors (Figures 3–5), it is possible to analyze the structures in more quantitative terms. For each high-resolution structure we have calculated (Tables III and IV) the intermolecular hydrogen bond lengths (and the average), the total hydrophobic surface area buried upon complex formation (Connolly, 1983a,b), the sizes of the void volumes created in the specificity pockets upon inhibitor (cavity volume, see Table III; Connolly, 1985; R. Bone and D. A. Agard, manuscript in preparation) and the magnitude of the main-chain conformational adjustment in the region of Val 217A (measured by the sum of the changes in torsional angles Ψ_{217} and Φ_{217A} relative to the unliganded enzymes; Bone et al., 1989b; R. Bone and D. A. Agard, manuscript in preparation).

Structural Parameters and Energetics. The structural parameters reported in Table III provide a more detailed understanding of how enzyme–inhibitor interactions have changed as a result of mutation by quantifying the energetic contributions of the changes in interactions. Changes in parameters such as average hydrogen bond length have predictable energetic effects that indicate whether observed changes in structure enhance or degrade the stability of enzyme–inhibitor complexes and by what quantity. Hydrogen bonds can be considered primarily electrostatic interactions, the energy of which will fall off as the bonds become stretched or bent (Schulz & Schirmer, 1979). If a hydrogen bond is

[4] Boron bond angles of 109° would be expected in a purely tetrahedral adduct, while angles of 120° would be expected in a trigonal planar adduct.

Table IV: Intermolecular Hydrogen-Bond Distances in α-Lytic Protease–Inhibitor Complexes[a]

	Wild Type inhibitor side chain[b]				
hydrogen bond	Ala[c]	Val[c]	Ile[c]	Norleu[c]	Leu
Ser 195 N---P_1 O_1	2.97	2.97	3.03	3.14	3.20
Gly 193 N---P_1 O_1	2.45	2.59	2.55	2.54	2.71
His 57 $N\epsilon_2$---P_1 O_2	2.65	2.62	2.76	2.65	2.47
Ser 214 O---P_1 N	2.93	3.11	2.99	3.20	3.21
Gly 216 N---Ala P_3 O	2.90	2.93	3.08	3.12	3.11
Gly 216 O---Ala P_3 N	2.87	3.03	3.13	3.06	3.06
	MA192 inhibitor side chain[b]				
hydrogen bond	Ala	Val	Norleu	Leu	Phe[d]
Ser 195 N---P_1 O_1	3.00	3.00	3.03	3.02	2.92
Gly 193 N---P_1 O_1	2.46	2.63	2.59	2.51	2.60
His 57 $N\epsilon_2$---P_1 O_2	2.78	2.64	2.64	2.77	2.74
Ser 214 O---P_1 N	2.98	3.20	3.05	3.23	3.09
Gly 216 N---Ala P_3 O	2.96	2.95	3.01	3.04	3.07
Gly 216 O---Ala P_3 N	2.96	2.95	2.92	2.92	2.90
	MA213 inhibitor side chain[b]				
hydrogen bond	Ala	Val[d]	Norleu	Leu	Phe
Ser 195 N---P_1 O_1	3.01	3.04	3.03	3.11	3.06
Gly 193 N---P_1 O_1	2.46	2.44	2.49	2.41	2.48
His 57 $N\epsilon_2$---P_1 O_2	2.80	2.76	2.71	2.78	2.62
Ser 214 O---P_1 N	3.19	3.20	3.14	3.14	3.11
Gly 216 N---Ala P_3 O	3.09	3.06	2.97	3.08	3.14
Gly 216 O---Ala P_3 N	2.93	3.11	3.02	2.92	3.09

[a] All hydrogen-bond distances are given in angstroms. [b] P_1 side chain of mAAp-boroX, where the boro prefix indicates that the inhibitor is the α-amino boronic acid analogue of the corresponding amino acid. [c] Bone et al. (1989b). [d] Bone et al. (1989a).

stretched by 0.3 Å or bent by 20°, the energy of the interaction will decrease by approximately 10–15% (Pimentel & McClellan, 1960; Mitchel & Price, 1990; Reed et al., 1986; Cybulski & Scheiner, 1989). For a hydrogen bond between neutral partners the intrinsic energy of the interaction is 6 kcal/mol (Weiner et al., 1984) and a reduction of 10% would decrease inhibitor affinity by 0.6 kcal/mol. However, uncertainty in atomic positions, and therefore hydrogen-bond lengths, due to experimental error and the resolution limit of these studies is estimated to be 0.12 Å (R. Bone and D. A. Agard, manuscript in preparation). In support of this value, difference Fouriers calculated from observed structure factors from different enzyme–inhibitor complexes and calculated phases clearly indicate positional adjustments of 0.2 Å. We consider changes in hydrogen-bond lengths of 0.2 Å to be significant and well defined by our high-resolution and highly refined data.

On the basis of free energies of transfer, removing hydrophobic surfaces from solvent should contribute 25 cal/Å^2 to the free energy of inhibitor binding (Richards et al., 1977); removing an extra 55 Å^2 of hydrophobic surface from solvent should improve catalysis or binding 10-fold. The creation of void volumes upon complex formation is an unfavorable consequence of binding that will reduce binding by approximately 60 cal/Å^3 (Rashin et al., 1986). The amount of energy required to create a water-sized cavity, 11 Å^3, is estimated to be 0.66 kcal/mol and should decrease activity by a factor of nearly 3. Conformational changes resulting from inhibitor binding require energy and will detract from inhibitor affinity, although the amount of energy is uncertain.

Native Complex. In a previous study we examined how peptide boronic acids with Ala, Val, Ile, and Norleu side chains in the P_1 position were accommodated by the wild-type enzyme and found that main-chain flexibility in the region of Val 217A, involving rotations of Ψ_{217} and Φ_{217A}, was the primary conformational adjustment (Bone et al., 1989b). The magnitude of the conformational changes increases and intermolecular hydrogen bonds stretch as the size of the side chain projecting into the specificity pocket increases (Bone et al., 1989b). In this work we have extended those studies by determining the structure of the complex between α-lytic protease and mAAP-boroLeu. Rather than an adjustment in the main-chain conformation of residues 217–217A, the side chain of Val 217A rotates by 120° about the C_α–C_β bond to remove the Val methyl group from the pocket and allow the P_1 Leu side chain into the specificity pocket (Figure 3). As a consequence of the poor complementarity between the Leu side chain and the specificity pocket, Met 192 is pushed deeper into the pocket and C_ϵ occupies a position occupied by S_δ in the other complexes (Figure 3). The side-chain adjustment leaves C_ϵ and C_β in an unfavorable eclipsed conformation ($\chi_3 = 15°$), which must detract significantly from the stability of the complex; the energetic cost of maintaining a fully eclipsed Met side-chain conformation would be 3–5 kcal/mol (Jainin et al., 1978). Adjustment in the position of Met 192 forces the residues surrounding the Met, such as Ser 226, toward the body of the protein. Positioning of the inhibitor in the active site is also suboptimal as a result of the poor complementarity. The C_α and C_β of the boronic acid residue are forced away from the more optimal positions that they occupy in complexes having Ala or Val in the P_1 position (Figure 3). In addition, the hydrogen bonds between the enzyme and inhibitor become stretched[5,6] to nearly the same extent as in the complex with mAAP-boroNorleu (Tables III and IV). Despite these adjustments, no more hydrophobic surface is removed from solvent than in the other complexes (Table III).

Activity toward the substrate with Leu in the P_1 position is reduced greatly because hydrogen bonds are not optimal, the Leu side chain is not removed completely from solvent, and transition-state stabilization energy must be diverted to drive the Val 217A and Met 192 conformational changes. Because three fundamentally different conformational adjustments (the main-chain shift in the region of Val 217A, the rotation of the Val 217A side chain, and the rotation of the side chain of Met 192) are used separately or in combination to accommodate inhibitor side chains in the specificity pocket, the energetic cost of these changes must be on roughly the same scale. If this were not the case, only the lowest energy conformational change would be observed to accommodate substrates. Substrates and inhibitors with large side chains in the P_1 position are selected against by the wild-type enzyme because intermolecular hydrogen bonds are stretched, the conformation of the enzyme must be changed, and there is no compensating gain in hydrophobic interactions. Partially compensating these unfavorable results of binding is the re-

[5] The energy of hydrogen-bond formation also decreases if hydrogen bonds become shorter than optimal. In this series of enzyme–inhibitor complexes one hydrogen bond (Gly 193 N---P_1 O_1) improves by lengthening to a more optimal distance as the size of the P_1 side chain increases. However, this is compensated by a corresponding decrease in the length of another hydrogen bond (His 57 $N\epsilon_2$---P_1 O_2) to values shorter than optimum.

[6] Hydrogen bonds stretched the most include not only hydrogen bonds in the oxyanion hole but also hydrogen bonds involving residues 214 and 216. The release of strain induced by bad steric contacts is complex but appears to involve pivoting of the peripheral portion of the inhibitor about the P_1 boron. This may be due in part to the packing of P_1 side chains against the P_2 Pro of the inhibitor. In a recent comparison of the binding of transition-state analogues and product analogues, it was shown that the only hydrogen bond disrupted in the product complex was the hydrogen bond with residue 214 (Bone et al., 1991).

duction in the size of the cavity created when larger P_1 side chains interact with the specificity pocket.

Accommodation of Side Chains by Met 192 → Ala. Met 192 → Ala displays broad specificity yet maintains high activity (Bone et al., 1989a). Inhibitors with increasingly large P_1 side chains are accommodated in the specificity pocket of Met 192 → Ala in the same manner as they are in the wild-type specificity pocket; that is, the main chain of Val 217A shifts away from the center of the pocket and/or the side chain of Val 217A rotates by 120° to remove the Val methyl group from the pocket and to allow large substrate side chains access (Figure 4). The position of Gly 216 also adjusts in response to the nature of the P_1 side chain, although it does not in either wild type or Met 213 → Ala complexes with inhibitors (Figures 3–5). From this observation, it appears that one role of Met 192 is to restrict the position of Gly 216 and that its absence in Met 192 → Ala is partially responsible for the ability of the specificity pocket to shrink.

In the complex of Met 192 → Ala with mAAP-boroAla, C_β of the P_1 Ala moves further into the specificity pocket than in other complexes, the sides of the pocket (residues 192–193, 215–216) that have contracted in the free mutant enzyme (Bone et al., 1989a) remain contracted, and Val 217A moves further into the pocket. The net result is minimization of the size of the remaining cavity, reduction of the cross section of the cavity to less than the diameter of a water molecule, and exclusion of solvent. There is remarkably little increase in the cavity volume in the complex of Met 192 → Ala with mAAP-boroAla compared to the wild-type complex with the same inhibitor (Table III). Although a Met side chain, with a volume of 80 Å^3, has been removed from the pocket, the increase in cavity volume is only 11 Å^3, providing an indication of the extent to which the specificity pocket has shrunk. This shrinkage appears to be the main reason for the maintenance of high activity toward substrates with small P_1 side chains.

As the size of the P_1 side chain becomes larger, the pocket expands in volume (R. Bone and D. A. Agard, manuscript in preparation) through conformational adjustments in the region of Val 217A and Gly 216 (Table III). Smaller structural distortions are required to accommodate large side chains in the specificity pocket in Met 192 → Ala inhibitor complexes than in wild-type complexes (Table III). Also in contrast with the wild-type enzyme, there is no tendency for intermolecular hydrogen bonds to become stretched as the P_1 side chain becomes larger and the amount of hydrophobic surface area removed from solvent increases dramatically (Tables III and IV). In complexes with Met 192 → Ala, boronic acids with Norleu and Phe in the P_1 position bury 70 and 80 Å^2 more hydrophobic surface than the wild-type complex with mAAP-boroVal and 120 and 130 Å^2 more than the wild-type complex with mAAP-boroAla. This improvement in hydrophobic interactions could enhance binding or catalysis up to 200-fold. Improvements in activity and affinity on this scale are observed for the interaction of the mutant with either inhibitors or substrates (Table II). For complexes of Met 192 → Ala with inhibitors that have large P_1 side chains, the cavity volume decreases, but never to the extent of the wild-type enzyme. Because these alterations in cavity volume are quite small for Met 192 → Ala, they are unlikely to significantly influence the specificity profile. Thus, broad specificity results from improved hydrogen bonds, increased hydrophobic interactions, and attenuated conformational adjustments, all of which lead to high activity on substrates with large P_1 side chains. The ability of the specificity pocket to shrink provides for high activity on substrates with small P_1 side chains.

On the basis of the structure of the complex between mAAP-boroNorleu and Met 192 → Ala, it appears that the energetics of the conformational adjustments in the region of Val 217A are altered as compared with the wild-type enzyme. In this structure, Val 217A occupies two conformations: in one conformation the Val 217A side chain is rotated into the specificity pocket and in the other the side chain is rotated out of the pocket (Figure 4). The two conformations, each at about 50% occupancy, are clearly indicated by patterns of positive and negative electron density in initial difference maps and in difference maps calculated after refinement with one or the other of the conformations at full occupancy. The conformation of the main chain ($\Phi_{217} = 76°$, $\Psi_{217A} = -64°$) is almost exactly the same as in the wild-type complex with the Ala boronic acid ($\Phi_{217} = 74°$, $\Psi_{217A} = -66°$). The Val side-chain rotation is not observed in wild-type complexes with boroIle and boroNorleu inhibitors, which require a large and costly main-chain shift in order to be accommodated (Bone et al., 1989b), but it is observed here where the main-chain conformation seems optimal. From these observations, it appears that the energy required to rotate Val 217A out of the specificity pocket is less in the mutant than in the wild-type enzyme. This may not be surprising since elimination of Met 192 from the specificity pocket by mutation eliminates hydrophobic interactions between Val 217A and Met 192 that presumably participate in the stabilization of the native conformation.

Accommodation of Side Chains by Met 213 → Ala. Met 213 → Ala exhibits almost no selectivity for the P_1 side chain of substrates (Bone et al., 1989a) or inhibitors, and the activity of the enzyme has been reduced by approximately an order of magnitude relative to the wild-type enzyme. As the P_1 side chain increases in size, the same combination of main-chain flexibility in the region of Val 217A and rotation of the Val 217A side chain out of the specificity pocket is responsible for accommodating the P_1 side chains of inhibitors (Figure 5). However, in contrast with Met 192 → Ala, Met 213 → Ala also uses three different Met 192 conformations to adjust the shape of the specificity pocket (Figure 5) to complement the shape of the P_1 inhibitor side chain. Despite this seeming ability to tailor the shape of the specificity pocket, the Met 213 → Ala specificity pocket still shows poor complementarity to substrate side chains, as evidenced by the tendency of the boronic acid C_β to be pushed away from the center of the specificity pocket as the P_1 side chain increases in size. For both Met 213 → Ala and the wild-type enzyme, this tendency appears to be an indicator of poor P_1 side-chain complementarity (Figures 3–5). The specificity pocket of Met 213 → Ala appears to be shallow rather than deep, and although it can tolerate large side chains, it cannot accommodate them as well as Met 192 → Ala.

Solvent is sequestered in the active site in both the free mutant (Bone et al., 1989a) and the complex with the Ala boronic acid (Figure 5). Only a hydrogen bond with the ring nitrogen of Trp 147 stabilizes the molecule of solvent in the enzyme–inhibitor complex. This is the only α-lytic protease complex examined to date in which a molecule of solvent is observed to be buried in the specificity pocket on inhibitor binding. Evidently the presence of Met 192 prevents the type of shrinkage and solvent exclusion that occurs in the Met 192 → Ala complex with mAAP-boroAla from also occurring in Met 213 → Ala complexes. Because the specificity pocket is unable to shrink, the void volume created in the complex of Met 213 → Ala with mAAP-boroAla is 40 Å^3 greater than in the corresponding wild-type complex. The increased cavity

volume in this complex destabilizes the enzyme–inhibitor complex by an estimated 2.3 kcal/mol (Rashin et al., 1986) and is sufficient to explain the reduced activity and affinity of the enzyme toward Ala substrates and inhibitors.

There is no tendency for hydrogen bonds to become stretched as the size of the P_1 side chain of inhibitors increases, although the average hydrogen-bond distances appear to be slightly longer than either wild type or Met 193 → Ala intermolecular hydrogen-bond lengths (Tables III and IV). Hydrophobic interactions improve as the boronic acid side chain increases in size, but not to the extent observed in Met 192 → Ala complexes. When Norleu and Phe side chains interact with the Met 213 → Ala specificity pocket, 40 and 30 Å^2 less hydrophobic surface is buried, equivalent to approximately 1 kcal/mol less binding energy, than when inhibitors with these side chains interact with Met 192 → Ala. This must partially account for the decreased activity of Met 213 → Ala as compared with Met 192 → Ala toward substrates with large P_1 side chains. Another factor limiting the activity and affinity of Met 213 → Ala for substrates and inhibitors with large P_1 side chains is that side-chain accommodation requires conformational changes nearly as large as in the wild type–mAAP-boroNorleu complex (Table III; Figure 5). In addition, it must also take energy to alter the conformation of Met 192 in these complexes from its conformation in the free mutant enzyme.

Discussion

Quality of Structural Models. Qualitatively, the enzyme–inhibitor complexes studied in this work appear to be excellent analogues of high-energy reaction intermediates (Bone et al., 1987). However, the question of how closely these complexes resemble the actual high-energy, anionic, tetrahedral intermediates that are formed during substrate hydrolysis should be addressed. The correlation between log K_i and log (k_{cat}/K_M), based on data for 19 enzyme–substrate pairs varying over 5 orders of magnitude in activity, is quite good (Figure 2; correlation coefficient = 0.87) and improves significantly (correlation coefficient = 0.92) if two outlying points (wild-type and Met 192 → Ala enzymes paired with Val in the P_1 position) are omitted. These two deviant points, the scatter of the observations about the line, and the slope of the plot (–0.7) suggest that some aspects of the transition state are not mimicked properly by the inhibitors.

The breakdown in transition-state analogy is likely to be a consequence of subtle differences between the charge distributions of the bound inhibitors (negative charge on boron) and the true high-energy intermediates (negative charge on oxygen), which are reinforced by the steric consequences of accommodating branched side chains in the specificity pocket (collision with Met 213; Bone et al., 1989b). Consequently, hydrogen-bond lengths and electrostatic interactions are expected to be slightly different in substrate complexes during the transition state than in boronic acid complexes. Differences in interatomic distances of only 0.1–0.3 Å would translate into changes in interaction energy of 0.5–2.5 kcal/mol (Cybluski & Scheiner, 1989; Mitchel & Price, 1990; Reed et al., 1986; Schulz & Schirmer, 1979; Weiner et al., 1984) and would be sufficient to account for the deviation of even the most scattered data from the correlation between transition-state stabilization and inhibitor affinity. Therefore, the structures of substrate complexes during the transition state are not expected to differ significantly from the structures of boronic acid complexes in surface area buried, the extent of conformational changes, or the sizes of void volumes created upon inhibitor binding. In addition, relative changes in hydrogen-bond lengths in boronic acid complexes should be representative of changes that occur in the structures of the corresponding transition states for substrate hydrolysis. With these limitations in mind, conclusions regarding the specificity profiles for inhibitor binding by these three enzymes should also apply to the specificity profiles for substrate hydrolysis.

Structural Basis for Specificity. Comparison of the structures of the enzyme–inhibitor complexes studied in this work provide a qualitative picture of how each enzyme interacts with substrates (Figures 3–5). More detailed analysis of the energetic components of inhibitor binding (Tables III and IV) has led to an increased understanding of the structural and energetic basis for the substrate and inhibitor specificity profiles for these three enzymes. However, it is important to note that no single factor correlates well with inhibitor affinity or substrate activity. Nor does specificity correlate with properties of the substrates, as has been observed for some other proteases (Dorovska et al., 1972; Estell et al., 1986).

The wild-type specificity profile (Bone et al., 1989a), for either inhibitors or substrates, appears to be driven for the most part by steric exclusion and limited flexibility. The unfavorable consequences of allowing large side chains into the specificity pocket (stretched hydrogen bonds and induction of large conformational changes, Table III) appear to be the primary determinants of the specificity profile.

Mutation of Met 192 to Ala creates an enzyme with an extraordinarily broad specificity profile because the contribution of unfavorable interactions to binding have been reduced. Hydrogen bonds do not become stretched as the size of the P_1 side chain increases, and the conformational changes required to accommodate larger side chains have been greatly attenuated (Tables III and IV). The energy required to drive rotation of the side chain of Val 217A out of the specificity pocket has been reduced by mutation. In addition, hydrophobic interactions improve significantly as the size of the P_1 side chain increases, enhancing inhibitor affinity and substrate activity. Activity toward substrates with small side chains is retained because of the ability of the specificity pocket to shrink in the absence of Met 192, thereby reducing the size of the residual cavity in the enzyme–substrate complexes sufficiently to exclude solvent.

The Met 213 → Ala mutation leads to a flexible enzyme that, like Met 192 → Ala, can accommodate large variations in P1 side chain size but has lost activity as a result of mutation (Bone et al., 1989a). There appears to be no dominant factor responsible for the reduced activity; rather, a combination of factors are involved. The creation of large void volumes reduces activity toward P_1 Ala and Val substrates while poor complementarity, less than optimal hydrophobic interactions, and costly conformational changes reduce activity toward substrates with large side chains. Thus, is seems that Met 213 plays a significant role in either directly or indirectly effecting the precise positioning of substrates or inhibitors for optimal interactions with the enzyme.

While the trend in enzyme–inhibitor affinity appears to correlate with trends in average hydrogen-bond length and in the magnitude of side-chain-accommodating conformational changes for the wild-type enzyme, it does not for either Met 192 → Ala or Met 213 → Ala. Similarly, buried hydrophobic surface area may be an indicator of relative affinity for Met 192 → Ala but is not for either of the other enzymes. For Met 213 → Ala, inhibitor affinity does not change dramatically as the P_1 side chain is varied, but large variations are observed in hydrophobic interactions, cavity volumes, and conformational changes. The specificity profiles of these enzymes cannot

be understood in terms of only one factor such as the number of hydrogen bonds or the hydrophobicity of the P_1 side chain of the substrate. Rather, each specificity profile results from the summation of contributions from factors promoting complex formation (hydrogen bonding, hydrophobic interactions, electrostatic interactions) and factors discouraging complex formation (conformational changes, steric clashes leading to hydrogen-bond distortion, creation of void volumes).

Discrimination of CH_3- and H- by Met 192 → Ala. The specificity profiles of each enzyme have been expanded to include values for a substrate with Gly in the P_1 position. Of the three enzymes, Met 192 → Ala is most effective at selecting for Ala in preference to Gly at the P_1 position of substrates. Elimination of the methyl group destabilizes the transition state by 3.4 kcal/mol for amide hydrolysis catalyzed by Met 192 → Ala (Table II) and approaches the highest discrimination between methyl and hydrogen substituents observed (Fersht et al., 1980). The specificity pocket of Met 192 → Ala shrinks considerably in the absence of ligands (Bone et al., 1989a) and when accommodating an Ala side chain in the P_1 position (Figure 4). It seems unlikely that it would be able to collapse further. The ability of the enzyme to select against a Gly substrate is therefore a consequence of both the rigidity of the specificity pocket, at least in the sense of not having the capacity to shrink any further, and its oversized nature relative to the preferred substrate side chain (Ala). In the Met 213 → Ala active site the size of the specificity pocket does not shrink, which leads to the sequestering of solvent when Met 213 → Ala interacts with the Ala substrate. This results in a dramatic decrease in the Ala-Gly selectivity ($\Delta\Delta G$ = 1.2 kcal/mol), primarily through a reduction in activity toward the Ala substrate. Selectivity of the wild-type enzyme is intermediate between those of Met 192 → Ala and Met 213 → Ala ($\Delta\Delta G$ = 2.5 kcal/mol), perhaps because the specificity pocket of the wild-type enzyme has some ability to shrink and reduce size of the cavity when it interacts with the P_1 Gly substrate. These results suggest that optimum selectivity for a large nonpolar substrate substituent over a smaller one occurs when the specificity pocket is somewhat larger than the substituent that it selects for and cannot adjust to accommodate the smaller substituent.

Broad Specificity and Conserved Conformational Changes. Although mutation of either Met 192 or Met 213 to Ala results in an enzyme that can efficiently hydrolyze substrates having Met just prior to the scissile bond, both mutant proteases exhibit extraordinarily broad specificity. Broad specificity in these two mutants appears to be a primary consequence of conformational flexibility in the region of Val 217A. It is the ability of the main-chain and side-chain torsional angles in this and adjacent residues to adjust that allows the specificity pocket to expand dramatically in volume and accommodate side chains larger than Met or Norleu. In addition, adjustments in these residues participate in the shrinkage of the pocket that allows Met 192 → Ala to retain complementarity with substrates having smaller side chains.

The changes in conformation responsible for the accommodation of substrates by the wild-type specificity pocket are conserved in the mutant enzymes, although the energy required for these conformational transitions may differ from that of the wild-type enzyme. That each enzyme uses basically the same repertoire of responses as the primary means of accommodating substrate side chains in the specificity pocket suggests that the behaviors of mutant enzymes should be predictable if the wild-type enzyme is sufficiently well characterized in structural terms. Furthermore, the presence of this flexibility in the wild-type enzyme must be a factor promoting broadened specificity. If the enzyme were less flexible in the region of Val 217A, it should exhibit higher selectivity for Ala in the P_1 position.

Conformational flexibility in the region of Val 217A is very likely derived from the secondary structure of an eight-residue Ω loop (Leszcynski & Rose, 1986) beginning with Val 217A. This Ω loop appears to be able to readily absorb conformational adjustments that occur in the immediately preceding residues. Remaining unclear is how much energy each of these conformational adjustments requires and which intramolecular interactions determine the energy cost. The changes in main-chain torsional angles in the region of Val217A do not represent transitions from inherently low-energy Φ and ψ angular values to high-energy ones (Brant et al., 1967), so the energetics of these changes must be the result of interresidue interactions.

From the work presented here, it is now clear that the patterns of substrate specificity observed for α-lytic protease result not just from the direct interactions of side chains of Met 192, Met 213, and Val 217A with the P_1 side chain of the substrate. Conformational changes are central participants in substrate selection, and the energy requirements of these changes must be due in large part to interactions with surrounding residues. Therefore, the observed specificity profiles can be seen to arise indirectly from interactions that occur between peripheral regions of the protein and the active-site region. These interactions dictate the degree of flexibility available to the active-site residues by setting the energetic price for changing the conformation of the enzyme. One of the remaining challenges is to understand the energetic and structural origins of the interactions that act to modulate active-site flexibility.

Acknowledgments

We acknowledge Dr. Robert M. Stroud for allowing us the use of his diffractometer and Jim Mace for providing assistance with the kinetic analysis.

References

Bachovchin, W. W., Wong, W. Y. L., Farr-Jones, S., Shenvi, A. B., & Kettner, C. A. (1988) *Biochemistry 27*, 7689–7697.

Bartlett, P. A., & Marlowe, C. K. (1983) *Biochemistry 22*, 4618–4624.

Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, pp 92–118, McGraw-Hill, New York.

Bone, R., Shenvi, A. B., Kettner, C. A., & Agard, D. A. (1987) *Biochemistry 27*, 7609–7614.

Bone, R., Silen, J. L., & Agard, D. A. (1989a) *Nature 339*, 191–195.

Bone, R., Frank, D., Kettner, C. A., & Agard, D. A. (1989b) *Biochemistry 28*, 7600–7609.

Bone, R., Sampson, N. S., Bartlett, P. A., & Agard, D. A. (1991) *Biochemistry 30*, 2263–2272.

Brady, K., & Abeles, R. H. (1990) *Biochemistry 29*, 7608–7617.

Brant, D. A., Miller, W. G., & Flory, P. J. (1967) *J. Mol. Biol. 23*, 47.

Brayer, G. D., Delbaere, L. T. J., & James, M. N. G. (1979) *J. Mol. Biol. 131*, 743–775.

Chothia, C. (1984) *Annu. Rev. Biochem. 53*, 537.

Connolly, M. L. (1983a) *J. Appl. Crystallogr. 16*, 548–558.

Connolly, M. L. (1983b) *Science 221*, 709–713.

Connolly, M. L. (1985) *J. Am. Chem. Soc. 107*, 1118–1124.
Cybulski, S. M., & Scheiner, S. (1989) *J. Phys. Chem. 93*, 6565–6574.
Delbaere, L. T. J., Brayer, G. D., & James, M. N. G. (1981) *Eur. J. Biochem. 120*, 289–294.
Dorovska, V. N., Varfolomeyev, S. D., Kazanskaya, N. F., Klyosov, A. A., & Martinek, K. (1972) *FEBS Lett. 23*, 122–124.
Estell, D. A., Graycar, T. P., Miller, J. V., Powers, D. B., Burnier, J. P., Ng, P. G., & Wells, J. A. (1986) *Science 233*, 659–663.
Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., Freeman, San Francisco, CA.
Fersht, A. R. (1988) *Biochemistry 27*, 1577–1580.
Fersht, A. R., Shindler, J. S., & Tsui, W.-C. (1980) *Biochemistry 19*, 5520.
Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., & Winter, G. (1985) *Nature (London) 314*, 235–238.
Fujinaga, M., Delbaere, L. T. J., Brayer, G. D., & James, M. N. G. (1985) *J. Mol. Biol. 183*, 479–502.
Hames, B. D., & Rickwood, D., Eds. (1981) *Gel Electrophoresis of Proteins*, IRL, Oxford, England.
Hendrickson, W. A., & Konnert, J. (1981) in *Biomolecular Structure, Function, Conformation, and Evolution* (Srinivasam, R., Ed.) Vol. 1, pp 43–47, Pergamon, Oxford, England.
Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. R., & Richards, J. H. (1973) *Biochemistry 12*, 4732–4743.
Hunkapiller, M. W., Forgac, M. D., & Richards, J. H. (1976) *Biochemistry 15*, 5581–5588.
Jones, T. A., (1982) in *Computational Crystallography* (Sayre, D., Ed.) pp 303–317, Oxford University Press, Oxford, England.
Kettner, C. A., & Shenvi, A. B. (1984) *J. Biol. Chem. 259*, 15106–15114.
Kettner, C. A., Bone, R., Agard, D. A., & Bachovchin, W. W. (1988) *Biochemistry 27*, 7682–7688.
Krieger, M., Chambers, J. L., Cristoph, G. G., & Stroud, R. M. (1974) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. 30*, 740–748.
Leszcynski, J. F., & Rose, G. D. (1986) *Science 234*, 849.
Mitchel, J. B. O., & Price, S. L. (1990) *J. Comput. Chem. 11*, 1217–1233.
Pimentel, G. C., & McClellan, A. L. (1960) *The Hydrogen Bond*, Freeman & Co., London.
Rashin, A. A., Iofin, M., & Honig, B. (1986) *Biochemistry 25*, 3619–3625.
Reed, A. E., Weinhold, F., Curtiss, L. A., & Pochatko, D. J. (1986) *J. Chem. Phys. 84*, 5687–5703.
Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.
Russell, A. J., Thomas, P. G., & Fersht, A. R. (1987) *J. Mol. Biol. 193*, 803–813.
Schecter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157.
Schulz, G. E. and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York.
Silen, J. L., Frank, D., Fujishige, A., Bone, R., & Agard, D. A. (1989) *J. Bacteriol. 171*, 1320–1325.
Smith, J. L., Corfield, P. W. R., Hendrickson, W. A., & Low, B. W. (1988) *Acta Crystallogr., Sect. A 44*, 357–368.
Stroud, R. M., Kieg, L. J., & Dickerson, R. E. (1974) *J. Mol. Biol. 83*, 185.
Thompson, R. C. (1973) *Biochemistry 12*, 47–51.
Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765–784.
Westerik, J. O., & Wolfenden, R. (1972) *J. Biol. Chem. 247*, 8195–8197.
Whitaker, D. R. (1970) *Methods Enzymol. 19*, 599–613.
Williams, J. W., & Morrison, J. F. (1979) *Methods Enzymol. 63*, 437–467.
Wolfenden, R. (1972) *Acc. Chem. Res. 5*, 10–18.
Wycoff, H. W., Dorscher, R., Tsernoglou, D., Inagami, T., Johnson, L., Hrdman, K. D., Allewell, N. M., Kelly, D. M., & Richards, F. M. (1967) *J. Mol. Biol. 27*, 563–578.
Zamyatnin, A. A. (1972) *Prog. Biophys. Mol. Biol. 24*, 107–123.