Episelection: Novel $K_i \sim$ Nanomolar Inhibitors of Serine Proteases Selected by Binding or Chemistry on an Enzyme Surface[†,‡]

Bradley A. Katz,[§] Janet Finer-Moore,[∥] Reza Mortezaei,[⊥] Daniel H. Rich,[⊥] and Robert M. Stroud*[,§,∥]

Arris Pharmaceutical Corporation, 385 Oyster Point Boulevard, Suite 4, South San Francisco, California 94080, Department of Biochemistry and Biophysics and Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0448, and School of Pharmacy and Department of Chemistry, University of Wisconsin—Madison, 425 North Charter Street, Madison, Wisconsin 53706

Received January 18, 1995; Revised Manuscript Received March 29, 1995[⊗]

ABSTRACT: A novel class of mechanism-based inhibitors of the serine proteases is developed using epitaxial selection. Tripeptide boronates esterified by an alcohol or alcohols at the boron retain the tight binding to trypsin-like enzymes associated with transition-state analogs and incorporate additional groups that can be utilized for selectivity between proteases. Formed by reaction of a series of alcohols with the inhibitor boronate oxygen(s), the most structurally compatible alcohol-derivatized inhibitors are either selected by binding to the enzyme (epitaxial selection) or assembled by epitaxial reaction on the enzyme surface. Mass spectrometry of the derivatized boronates and X-ray crystallography of the complexes identify the chemical structures and the three-dimensional interactions of inhibitors generated. This scheme also engineers novel, potent ($K_i \sim 7$ nM), and more specific inhibitors of individual serine proteases, by derivitizations of compounds obtained by epitaxial selection.

The proposal of Pauling (1946) that enzymes evolved to interact most strongly with transition states is validated both by structures of enzymes complexed with transition-state analogs and by catalytic antibodies generated against such analogs (Zhou et al., 1994). In a multienzyme family whose members share the same catalytic mechanism, inhibitors that are transition-state analogs, although tight-binding family-specific inhibitors, are often compromised as pharmaceutical leads because they are not selective for individual enzymes within the family. However, if an already tight binding motif can be further engineered to generate selectivity toward individuals in the enzyme family at sites peripheral to the active site and primary substrate binding sites, the resulting compounds become potential prospects for drug development strategies.

With the goal of developing potent and selective inhibitors of serine proteases, we sought to define modes of binding of high-affinity inhibitors. Here, we report on one of these classes and its development toward a tunable, selective class of inhibitors. Serine proteases are therapeutic targets, since they play pivotal roles in numerous diverse biological processes (Stroud, 1974) including blood coagulation (Davie et al., 1991), T-cell activation (Flentke et al., 1991), bacterial pathogenesis (Bachovchin et al., 1990), fibrinolysis (Collen & Lijnen, 1991), complement activation (Frank & Fries, 1989), neuropeptide modulation (Franconi et al., 1989), muscle contraction (Sekizawa et al., 1989), extracellular matrix catabolism (Vartio et al., 1981), and submucosal gland secretion (Sommerhoff et al., 1989). Specific roles are postulated in mast cell degranulation and hence airway responses in allergy (Tam & Caughey, 1990). Accordingly, members of this enzyme family are targets for a variety of diseases including arterial thrombosis (Kettner et al., 1993), asthma (Powers, 1986), anaphylaxis (Powers, 1986), emphysema (Hance & Crystal, 1975), and disorders of the immune response (Flentke et al., 1991).

The catalytic power of serine proteases is coded in four essential structural characteristics: binding of the main chain of the protein substrate, the catalytic triad, the oxyanion hole, and the specificity pockets. Binding of the main chain of natural peptide or protein substrates of serine proteases proceeds by formation of antiparallel β-sheet hydrogen bonds between substrate and enzyme near the active site. In trypsin-like proteases, active site residues Asp102, His57, and Ser195 [chymotrypsinogen residue numbering scheme of Hartley and Kauffman (1966)] participate in catalyzing peptide bond cleavage. The OH of Ser195 acts as the nucleophile and interacts with His57 which is oriented by the buried Asp102 to present the correct tautomer, with its lone pair, ready to accept a proton during catalysis (Blow et al., 1969; Stroud et al., 1971; Kossiakoff & Spencer, 1981). Ser195 attacks the carbonyl of a substrate scissile peptide bond, forming a tetrahedral transition-state-like intermediate in which the negatively charged oxygen of the tetrahedral intermediate is stabilized in the oxyanion hole by hydrogen bond donors from the backbone amides of Gly193 and Ser195 (Kraut, 1977; Stroud et al., 1977).

Serine proteases related to trypsin comprise an enzyme class unified by a common mechanism of peptide bond hydrolysis, but distinguished within the class by selectivity for three-dimensional determinants on the target substrates. A key determinant of specificity involving the substrate side chain prior to the scissile bond cleaved by trypsin-like

[†] Supported by Arris Pharmaceuticals, by National Institutes of Health Grant GM24485 to R.M.S., and by a grant from Arris Pharmaceuticals to D.H.R.

[‡] Coordinates for the structures have been deposited in the Brookhaven Protein Data Bank (file names 1btx, 1btw, 1bty, 1btz).

* To whom correspondence should be addressed.

[§] Arris Pharmaceutical Corp.

[∥] University of California at San Francisco.

[⊥] University of Wisconsin—Madison.

[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1995.

0006-2960/95/0434-8264$09.00/0

proteases is the P1 pocket [nomenclature of Schecter and Berger (1967)]. In trypsin, the side chain of Asp189, at the base of the P1 pocket, confers a preference for sequences containing Lys or Arg at S1 of the substrate polypeptide chain, which form salt bridges with Asp189 upon binding (Krieger et al., 1974; Mangel et al., 1990). However mutagenesis to convert trypsin into chymotrypsin (Hedstrom et al., 1992, 1994a,b; Perona et al., 1994) demonstrates that the specificity features of each different serine protease also involve the several (primarily three) loops that surround the P1 binding pocket. These structural and mechanistic features have been exploited in the understanding and development of various classes of potent transition-state-like protease inhibitors including organofluorophosphates (Stroud et al., 1974), peptidyl chloromethyl ketones (James et al., 1980), aldehydes (Delbaere & Brayer, 1985), phosphonate esters (Bone et al., 1991a), and boronic acids (Bone et al., 1989, 1991b; Bachovchin et al., 1988; Tsilikounas et al., 1992).

Our first goal was to elucidate the interactions responsible for the tight binding of a tripeptidylboronate transition-state analog to trypsin-like serine proteases by crystallography. We discovered a novel compound in which a stereoselective esterification of one of the boronate oxygens had occurred by reaction with an alcohol (methanol) from the crystallization medium. Subsequently, we sought to develop a basis for selectivity by utilizing the enzyme surface to "select" from the components in solution, or to facilitate selective reactions, and thereby incorporate additional inhibitor substituents, following incubation with reacting species, in this case alcohols. Peptidylboronate ester derivatives complexed with trypsin were prepared by incubation of the enzyme and the pinanediol diester of the peptidylboronate **8** with an excess of an alcohol, followed by crystallization for structural analysis. Selection and incorporation by the enzyme surface yielded boronate esters in which either or both of the available boronate oxygens is (are) linked to substituent R groups to yield additional binding substituent(s). We report the design, synthesis, activity, and chemical and crystal structures of a series of transition-state analog inhibitors prepared by this epitaxial selection route.

EXPERIMENTAL PROCEDURES

Synthesis

Synthesis of the Parent Tripeptide Boronate (Scheme 1). Proton nuclear magnetic resonance (^{1}H-NMR)[1] spectra were recorded on a Bruker AM 300 (300 MHz), and chemical shifts are reported as δ units (ppm) relative to tetramethylsilane, the internal standard. Coupling constants (J) are given in hertz (Hz), and multiplicities are described as singlet (s), doublet (d), triplet (t), quadruplet (q), broad singlet (bs), massive (mass), and multiplet (m). High-resolution fast atom bombardment (HRFAB) mass spectra were carried out by the Midwest Center for Mass Spectrometry (University of Nebraska, Lincoln). High-resolution electronic ionization (HREI) mass spectra were carried out at the Department of Chemistry, University of Wisconsin, Madison.

4-Bromobutylboronate Ester of (+)-Pinanediol (2). 4-Bromo-1-butene (13.5 g, 0.1 mol) was hydroborated by adding a 1.0 M solution of catecholborane in 100 mL of THF and heating for 4 h at 100 °C under a nitrogen atmosphere. After evaporation of solvent and distillation (bp = 123−128 °C at 0.3 mm), 11 g of 4-bromobutylboronate catechol **1** were obtained and reacted with (+)-pinanediol (7.3 g, 43 mmol) in THF at 0 °C for 30 min. After additional stirring for 30 min at room temperature, the solvent was evaporated and hexane was added to remove catechol as a crystalline solid. Purification by flash chromatography (EtOAc:hexane = 1:9) gave 10 g of the desired ester as a liquid (overall yield = 39%): R_f = 0.54 (EtOAc:hexane = 1:9); $[\alpha]_{RT}$ = +63.6°, c = 0.83 (EtOAc); HREI (m/e) calculated = 316.1038, found = 316.1048; ^{1}H-NMR ($CDCl_3$) 4.29 (1H, dd, J = 8.7 and 1.9), 3.41 (2H, t, J = 6.9), 2.34 (1H, m), 2.22 (1H, m), 2.04 (1H, t, J = 5.2), 1.88 (4H, m), 1.57 (2H, m), 1.38 (3H, s), 1.22 (3H, s), 1.09 (1H, d, J = 10.8), 0.84 (3H, s), 0.90−0.80 (2H, m).

5-Bromo-1-chloropentylboronate Ester of (+)-Pinanediol (3). In a typical homologation process, 676 mL of methylene chloride was dissolved in 10 mL of freshly distilled THF, and the mixture was cooled to −100 °C. A 2.0 N n-butyllithium solution (4.6 mL, 9.2 mmol) was added very slowly dropwise down the side of the flask before contact with the reaction mixture. Compound **2** (2.92 g, 9.2 mmol) was then dissolved in 10 mL of THF. The resulting solution was cooled to its freezing point and added slowly to the reaction mixture. After being stirred for 20 min at −100 °C, a 1.0 M solution of zinc chloride (4.6 mL, 4.6 mmol) was cooled to 0 °C and was added in two equal portions to the reaction mixture, which was then slowly warmed to room temperature and stirred overnight. After evaporation of the solvent, the residue was dissolved in hexane and washed with water and then with brine. After being dried over magnesium sulfate, the solvent was evaporated to give 3.1 g (92% yield) of the desired compound as a liquid. An analytical sample was purified by flash chromatography (Et_2O:hexane = 1:9): R_f = 0.39 (Et_2O:hexane = 1:9); $[\alpha]_{RT}$ = +27.3°, c = 1.45 ($CHCl_3$); HREI (m/e) calculated = 364.0803, found = 364.0806; ^{1}H-NMR ($CDCl_3$) 4.37(1H, dd, J = 8.9 and 2.0), 3.47 (1H, m), 3.41 (2H, t, J = 6.8), 2.45−2.20 (2H, mass), 2.09 (1H, m), 1.98−1.83 (6H, mass), 1.77−1.51 (3H, mass), 1.42 (3H, s), 1.16 (1H, dd, J = 11 and 1.6), 0.95−0.80 (5H, mass includes 3H, s at 0.84).

1-Amino-5-bromopentylboronate Ester of (+)-Pinanediol Hydrochloride (4). Compound **3** (3.1 g, 8.5 mmol) was dissolved in 10 mL of THF and cooled to −78 °C. A 1.0 M solution of lithium hexamethyldisilazane (9 mL, 9 mmol) in THF was added dropwise. The mixture was warmed to room temperature and was stirred overnight under argon. After evaporation of the solvent and addition of hexane, undissolved material was removed by centrifugation. Evaporation of hexane under vacuum yielded a crude oil which was dissolved in 10 mL of hexane. The solution was cooled to −78 °C, and 8.5 mL of 4 N HCl in dioxane was added dropwise. The solution was slowly warmed to room temperature and was stirred overnight. The solvent was evaporated, hexane was added, and undissolved material was removed after centrifugation. Compound **4** (3.1 g) was

[1] Abbreviations: NMR, nuclear magnetic resonance; THF, tetrahydrofuran; EtOAc, ethyl acetate; DMSO, dimethyl sulfoxide; MeOH, methanol; DMF, dimethylformamide; PTSA, p-toluenesulfonic acid; Et_2O, diethyl acetate; BOC, *tert*-butoxycarbonyl; DIP-trypsin, diisopropylfluorophosphate-inhibited trypsin IR, infrared; Tris, tris(hydroxymethyl)aminomethane; $|F_o|$ and $|F_c|$, observed and calculated structure factor amplitudes; α_c, calculated phases; rms, root mean square; BPTI, bovine pancreatic trypsin inhibitor. The prefix "boro" of boroLys indicates that the carbonyl of lysine is replaced by $B(OH)_2$.

obtained as a white solid after evaporation of hexane in 88% yield: $[\alpha]_{RT} = +14.6°$, $c = 2.2$ (MeOH); HRFAB (*m/e*) [M + 1] calculated = 344.1396, found = 344.1400; ^{1}H-NMR ($CDCl_3$) 8.29 (3H, bs), 4.40 (1H, d, $J = 8.7$), 3.42 (2H, t, $J = 6.8$), 2.96 (1H, bs), 2.40−2.15 (2H, mass), 2.10−1.52 (9H, mass), 1.42 (3H, s), 1.28 (3H, s), 1.15 (1H, d, $J = 11$), 0.83 (3H, s).

*(BOC)Ala-Val-NH[$(CH_2)_4$Br]CHB$O_2C_{10}H_{16}$ (**5**) and the Corresponding Boronic Acid (**6**).* (BOC)Ala-Val-OH (415 mg, 1.4 mmol) was dissolved in 3 mL of methylene chloride and cooled to −20 °C. *N*-Methylmorpholine (160 mL, 1.44 mmol) was added followed by isobutyl chloroformate (164 mL, 1.44 mmol). The mixture was stirred for 15 min. A solution of **4** (500 mg, 1.31 mmol) in 3 mL of methylene chloride was added followed by addition of *N*-methylmorpholine (160 mL, 1.44 mmol). After being stirred for 1 h at −20 °C and 1 h at room temperature, the solvent was evaporated and ethyl acetate was added. After extraction, the organic layer was washed with 5% citric acid, followed by a saturated solution of sodium bicarbonate and brine, and then was dried over magnesium sulfate and concentrated to give 500 mg of crude product. After flash chromatography on silica gel, 200 mg of ester **5** (25% yield) and 140 mg of the corresponding acid **6** (22%) were obtained.

Characterization of **5**: $R_f = 0.57$ (EtOAc); $[\alpha]_{RT} = -40.9$, $c = 0.6$ ($CHCl_3$); ^{1}H-NMR ($CDCl_3$) 7.01 (1H, mass), 6.59 (1H, d, $J = 8.8$), 5.03 (1H, d, $J = 5.5$), 4.34 (1H, m), 4.28 (1H, d, $J = 6.9$), 4.11 (1H, m), 3.39 (2H, d, $J = 6.8$), 2.91 (1H, mass), 2.40−2.07 (3H, mass), 2.00 (1H, t, $J = 5.2$), 1.92−1.75 (4H, mass), 1.68−1.25 (23H mass, includes 9H, s, at 1.44, 3H, s, at 1.38, 3H, d, $J = 6.4$ at 1.37, and 3H, s, at 1.27), 0.93 (3H, d, $J = 9.0$), 0.91 (3H, d, $J = 9.2$), 0.84 (3H, s).

Characterization of **6**: $R_f = 0.41$ (EtOAc); $[\alpha]_{RT} = -40.2°$, $c = 0.5$ ($CHCl_3$); ^{1}H-NMR ($CDCl_3$) 6.69(1H, d, $J = 8.6$), 6.43 (1H, bs), 5.00 (1H, d, J = 5.8), 4.20 (1H, dd, $J = 8.3$ and 6.0), 4.13 (1H, t, $J = 6.8$), 3.39 (2H, t, $J = 6.3$), 3.25 (2H, m), 2.27 (2H, mass), 1.87 (2H, m), 1.59−1.33 (16H, mass, includes 9H, s, at 1.45 and 3H, d, $J = 6.0$ at 1.39), 0.94 (3H, d, $J = 6.8$), 0.89 (3H, d, $J = 6.8$).

*(BOC)Ala-Val-NH[$(CH_2)_4N_3$]CHB$O_2C_{10}H_{16}$ (**7**).* Compound **5** (200 mg, 0.32 mmol) was dissolved in 3 mL of dry DMF. Sodium azide (43 mg, 0.95 mmol) was added, and the reaction mixture was heated for 4 h at 100 °C under argon. After cooling, 10 mL of ethyl acetate was added, and the organic layer was washed three times with 5 mL of water. After being dried over magnesium sulfate and evaporation of solvent, 181 mg of **7** was obtained (98% yield): $R_f = 0.57$ (EtOAc); $[\alpha]_{RT} = -46.3°$, $c = 0.33$ ($CHCl_3$); HREI (*m/e*) calculated = 576.3813, found = 576.3817; ^{1}H-NMR ($CDCl_3$) 6.97 (1H, bs), 6.50 (1H, d, $J = 8.0$), 4.98 (1H, bs), 4.40−4.20 (2H, mass), 4.12 (1H, m), 3.25 (2H, t, $J = 6.9$), 2.92 (1H, bs), 2.40−2.10 (3H, mass), 2.01 (1H, m), 1.95−1.50 (6H, mass), 1.44 (10H, includes 9H, s), 1.40 (6H, m, includes 3H, s), 1.27 (5H, m, includes 3H, s), 0.93 (6H, m), 0.84 (3H, s). The IR spectrum of this compound shows a strong peak at 2107 cm^{-1} indicative of the azide functionality.

*(BOC)Ala-Val-boroLys-$C_{10}H_{16}$ PTSA (**8**) (Pinanediol Diester of the Tripeptide Boronate).* Compound **7** (200 mg, 34 mmol) was dissolved in 3 mL of methanol, and anhydrous *p*-toluenesulfonic acid (58 mg, 0.34 mmol) was added. The solution was hydrogenated over 10% Pd/C, and the reaction was monitored by thin-layer chromatography. After 2 h, no starting material remained. The reaction solution was filtered over Celite, and solvent was removed to yield 227 mg (92% yield) of the pinanediol diester of the tripeptide boronate as a white solid: $[\alpha]_{RT} = -53.0°$, $c = 1.95$ (MeOH); ^{1}H-NMR (DMSO-d_6) 9.00 (1H, s), 7.76 (1H, d, $J = 8.8$), 7.54 (1H, bs), 7.48 (2H, d, $J = 8.0$), 7.12 (2H, d, $J = 8.0$), 7.03 (1H, d, $J = 7.5$), 4.20 (1H, t, $J = 8.5$), 4.08 (2H, m), 2.73 (1H, t, $J = 7.8$), 2.43 (1H, bs), 2.29 (3H, s), 2.22 (1H, m), 2.01 (2H, mass), 1.83 (2H, m), 1.64 (1H, d, $J = 13.4$), 1.57−1.28 (19H, mass, includes 9H, s, at 1.38), 1.25 (3H, s), 1.23 (3H, s), 1.16 (3H, d, $J = 7.0$), 0.86 (6H, m), 0.81 (3H, s).

*(BOC)Ala-Val-NH[$(CH_2)_4N_3$]BO_2H_2 (**9**).* This compound was prepared by the same procedure described for preparation of **7** in 92% yield: $R_f = 0.41$ (EtOAc); $[\alpha]_{RT} = -44.3°$, $c = 1.71$ ($CHCl_3$); ^{1}H-NMR ($CDCl_3$) 6.74 (1H, d, $J = 6.0$), 6.53 (1H, bs), 5.08 (1H, bs), 4.19 (2H, m), 3.26 (4H, m), 2.26 (1H, bs), 1.68−1.32 (19H, m, includes 9H, s, at 1.45 and 3H, d, $J = 7.0$ at 1.37), 0.94 (3H, d, $J = 6.8$), 0.90 (3H, d, $J = 6.8$). The IR spectrum of this compound indicates a strong peak at 2110 cm^{-1} indicative of the azide functionality.

*(BOC)Ala-Val-boroLys PTSA (**10**) (Tripeptide Boronate (p-Toluenesulfonate Salt).* Compound **8** was hydrogenated using the procedure described for hydrogenation of **7** to give **10** in 78% yield: $[\alpha]_{RT} = -41.5°$, $c = 1.1$ (MeOH); ^{1}H-NMR (DMSO-d_6) 7.90 (2H, mass), 7.48 (3H, m, includes 2H, d, $J = 7.9$), 7.12 (3H, m, includes 2H, d, J = 7.9), 4.08 (1H, dd, $J = 8.6$ and 7.3), 3.99 (1H, m), 3.03 (2H, m), 2.84 (1H, m), 2.29 (3H, s), 1.91 (1H, m), 1.62−1.20 (19H, mass includes 9H, s, at 1.38), 1.16 (3H, d, $J = 7.0$), 0.83 (6H, m).

Methods

Assay of the Tripeptide Boronate. The activity of bovine trypsin (10.5 nM, lot 58H4064 from Worthington Biochemicals, Freehold, NJ 07728) was measured by spectrophotometric assay using tosyl-L-Gly-Pro-Arg-*p*-nitroanilide (lot 056-180 from Sigma Chemical Co., St. Louis, MO 63178) as colorometric substrate. The buffer system was 50 mM Tris, pH 8.2, 100 mM NaCl, 0.05% Tween-20, and 10% DMSO. The inhibitor was agitated for 1 h for complete solubilization before the assay. Eight concentrations of inhibitor were used and the K_i values extracted using the program BatchK_i written by Kuzmic (University of Wisconsin, Madison) and based on published algorithms (Cha, 1975; Bajzer & Prendergast, 1992; Marquardt, 1963).

Cocrystallization of Trypsin−Tripeptide Boronate. Orthorhombic needle-shaped crystals of benzamidine-inhibited bovine trypsin were grown by the batch method of Stroud et al. (1974). Trypsin (150.0 mg) was dissolved in 5.0 mL of 20 mM Tris, pH 8.15, and 1.0 mg/mL benzamidine. $MgSO_4 \cdot 7H_2O$ (1.02 g) was added and dissolved, and the resulting solution was centrifuged, filtered through a 0.45-mm syringe filter, and sealed in depression wells. Microscopic crystals appeared after 3−4 weeks and grew to data-collection size after 5−6 months.

Crystals of trypsin inhibited with the tripeptide boronate, added as its pinanediol diester *p*-toluenesulfonate salt, were grown in hanging drops by seeding with trypsin−benzamidine crystals. Two orthorhombic crystal forms of trypsin−tripeptide boronate, one with needle-shaped morphology in space group $P2_12_12_1$, $a = 54.8$ Å, $b = 58.7$ Å, $c = 67.6$ Å

Table 1: Crystallography of Trypsin–Benzamidine, Trypsin–Tripeptide Boronate, and Trypsin–Tripeptide Boronate Esters

	inhibitors						
	boronic acid		tripeptide boronate ester			benzamidine	
	xtl1	xtl2	MeOH	EtOH	propanediol	xtl1	xtl2
parameters[a]							
space group	2	2	1	1	2	1	1
no. of atoms (including disorder)	1963	1963	1914	1942	1910	4114	4095
no. of waters (including disorder)	191	191	205	199	176	225	229
no. of discretely disordered groups[b]	19	19	21	27	18	23	18
no. of discretely disordered waters	2	2	3	3	3	2	5
no. of side chains with refined occs[c]	14	14	8	7	14	12	13
no. of waters with refined occs	0	0	1	0	0	90	91
diffraction statistics							
resolution[e] (Å)	7.0–1.5	7.0–2.0	8.0–2.0	8.0–1.7	7.0–1.8	8.0–1.5	7.0–1.7
no. of observations			24120	15280			63142
no. of reflections[d]	24945	15820	8204	7977	12248	24173	17697
F/σ cutoff[e]	3.5	2.5	0	0	3.3	2.4	0
R_{merge}[f] (%)			10.1	8.0			
R_{cryst}[g] (%)	18.5	15.7	13.4	14.0	16.4	16.1	15.4
completeness[e] (%)	56.0	74.2	51.9	33.3	48.2	69.2	69.4
rms deviations[h]							
bond lengths (Å)	0.016	0.015	0.014	0.013	0.017	0.016	0.016
bond angles (deg)	2.9	2.9	2.9	3.0	3.3	3.7	3.7
torsion angles (deg)	26.1	26.2	26.5	26.5	26.4	26.7	26.3

[a] Restrained, isotropic temperature factors were refined for all structures. Hydrogen atoms were included in the refinement of the trypsin–benzamadine structures. Space group 1 is $P2_12_12_1$, $a = 54.8$ Å, $b = 58.7$ Å, $c = 67.6$ Å (Chambers & Stroud, 1977); space group 2 is $P2_12_12_1$, $a = 63.74$ Å, $b = 63.54$ Å, $c = 68.93$ Å. [b] Not including waters. [c] Also includes inhibitor groups. Density for all side-chain atoms or for terminal atoms in these groups was weak or absent, and temperature factors were high. Discretely disordered groups are not included in this category. Occupancies for poorly defined atoms were refined. In addition to the occupancies for the side chains totaled above, the occupancies for the main-chain atoms of loop residues 146–149 were refined (the main-chain atoms within each of these residues were constrained to be the same). [d] Refers to refinement limits. [e] Only eight reflections had $F/\sigma < 3.0$. [f] $R_{merge} = \sum_h\sum_i|I(h)_i - \langle I(h)\rangle|/\sum_h\sum_i I(h)_i$, where $I(h)_i$ is the ith observation of the intensity of reflection. [g] $R_{cryst} = \sum(F_o - F_c)/\sum F_o$. [h] Root mean square deviations from ideal bond lengths and bond angles.

(Chambers & Stroud, 1977), and one with a larger unit cell $P2_12_12_1$, $a = 63.74$ Å, $b = 63.54$ Å, $c = 68.93$ Å, and tetragonal-bipyramidal morphology (Mangel et al., 1990), were obtained, depending on the $MgSO_4$ concentration in the wells.

Preparation of Boronate Esters. Trypsin inhibited by tripeptide boronate esters was crystallized by incubating trypsin, and tripeptide boronate was added as its pinanediol diester *p*-toluenesulfonate salt in the presence of 7.5% alcohol (by volume) for 1–2 h. Solid $MgSO_4{\cdot}7H_2O$ was then added, and hanging drops were set up and seeded with both crystal forms of trypsin–tripeptide boronate to produce both crystal forms of trypsin–tripeptide boronate esters. Crystals were mounted directly from the hanging drops or alternatively transferred to synthetic mother liquor containing 442 mg/mL $MgSO_4{\cdot}7H_2O$, 100 mM Tris, pH 8.2, and 1.0% alcohol (by volume). The complexes crystallized by this method include trypsin–tripeptide boronate and its methanol, ethanol, and 1,3-propanediol derived esters.

Mass Spectrometry. Single crystals of trypsin–tripeptide boronate methyl ester were dissolved in a solution of 50% methanol:50% water (v:v) and 0.5% acetic acid. Residual salt from the crystallization medium was removed with a micro-desalting column. Samples were immediately run on an Electrospray mass spectrometer (Finegan).

Crystal Structure Determination. X-ray diffraction data from single crystals of trypsin–benzamidine, trypsin–tripeptide boronate, and trypsin–tripeptide boronate esters were collected on a multiwire IPC area detector coupled to a three-circle goniometer (Siemens Corp., Madison, WI) mounted on a rotating target tube (Rigaku) with Franks focusing optics (Franks, 1955). Data were corrected to produce integrated intensities using the program XDS (Kabsch, 1988) and scaled together using programs Rotavata/Agrovata (CCP4 programs, Daresbury, England). Typical settings were crystal to detector distance = 10 cm, $2\theta = 28°$, and oscillation widths = 4′ in ω. Alternatively, data were collected on an image plate detector (Rigaku RAXIS II) mounted on an 18-kW generator (Rigaku) run at 60 kV and 50 mA using a beam collimated to 0.3 mm and using 1.0–1.5° wide frames in ω. Data collection statistics are summarized in Table 1.

A structure of trypsin with no inhibitor or waters at or near the expected binding regions of the peptidylboronate moiety was extracted from the most highly refined coordinates for trypsin ($R = 15.2\%$ to 1.34 Å resolution) from our crystal structure of monoisopropylphosphoryltrypsin (Chambers & Stroud, 1977; Finer-Moore et al., 1992) in the Brookhaven Protein Data Bank (Bernstein et al., 1977, Entry Code 4PTP). Initial electron density maps calculated with coefficients ($|F_o| - |F_c|$) or ($2|F_o| - |F_c|$) and phases (α_c) for each trypsin–tripeptide boronate or tripeptide boronate ester showed strong density for all non-hydrogen atoms of the tripeptide boronate, except for those in the N-terminal BOC group (which appeared in later maps). For trypsin–tripeptide boronate esters, these maps also indicated additional density connected to one or both of the boronate oxygen atoms.

Models incorporating these features were built using FRODO (Jones, 1982), Quanta (molecular modeling program licensed from Molecular Simulations, 16 New England Executive Park, Burlington, MA 01803), or Insight (molecular modeling program licensed from Biosym Technologies, Inc., 10065 Barnes Canyon Road, San Diego, CA 92121)

and fully refined with difference Fourier methods (Chambers & Stroud, 1977) and with X-PLOR (Brunger, 1990) until residual density peaks were below ~1/3 of that for a carbon atom. Deviations of bond lengths from ideality in the final structures were typically 0.015 Å. Refinement statistics are given in Table 1. To analyze the geometry about the boron atom or ester groups more completely, the bond lengths and bond angles associated with atoms in these groups were unconstrained in some refinement cycles. The results of the unrestrained refinement together with boron bond length and bond angle data gleaned from small molecule crystal structures in the literature were used to develop appropriate bond length and bond angle parameters for further restrained refinements. The occupancies of the esterifying alcohol-derived group(s) were refined. For the trypsin−tripeptide boronate methyl ester (the first of this series of boronate ester complexes analyzed), an alternate model was also refined in which the density extending from the O1 boronate oxygen (in the oxyanion hole) was modeled as a water molecule. Comparison of the resulting maps with those obtained when the density was modeled as a methyl ester group further established that there was indeed a methyl ester at O1 and not a water hydrogen bonded to it.

Several models for the structure of the trypsin−tripeptide 1,3-propanediol ester were refined and analyzed. In several models the geometry about the boron was distorted from tetrahedral toward trigonal planar similarly to the geometry seen for other trypsin−tripeptide boronate complexes in this study. In one of these the boron coordinated to Cα of boroLys, and Oγ of Ser195 was coordinated by both oxygens of 1,3-propanediol in a cyclic diester and in the other by one of the 1,3-propanediol oxygens (O1, occupying the oxyanion hole) and one hydroxyl group.

The water structure for trypsin−inhibitor complexes was determined and refined as for other trypsin structures (Finer-Moore et al., 1992) by several iterations of difference Fourier analysis (Chambers & Stroud, 1977) followed by refinement with X-PLOR. Waters were included on the basis of significant density in both ($|F_o| - |F_c|$), α_c and ($2|F_o| - |F_c|$), α_c maps in stereochemically reasonable locations as defined by Finer-Moore et al. (1992), and waters with refined temperature factors exceeding 55 Å^2 were generally rejected.

Structure Comparisons. Comparisons were made with other reference trypsin crystal structures:[2] inhibitor-free structures of trypsin (Brookhaven Data Bank Entry Codes 1TPO and 2PTN; Marquart et al., 1983) and two structures of trypsin−benzamidine [Krieger et al., 1974; one recently refined to $R = 0.16$ at 1.5 Å resolution in this study and one independently refined by Bode and Schwager (1975) (3PTB)]. The coordinates of these reference structures were averaged to produce a "consensus" structure with rms deviations onto which the structure of trypsin−tripeptide boronate complexes were superimposed using the common core backbone atom pairs identified.

[2] Trypsin−inhibitor complexes: bovine trypsin−BPTI (2PTC), bovine trypsinogen−BPTI (2TGP; Marquart et al., 1983), bovine trypsinogen−BPTI−IleVal; bovine trypsinogen−porcine pancreatic secretory trypsin inhibitor (1TGS; Bolognesi et al., 1982); bovine α-chymotrypsin−third domain of turkey ovomucoid inhibitor (1CHO; Fujinaga et al., 1987); porcine kallikrein A−BPTI (2KAI; Chen et al., 1983); *Streptomyces griseus* proteinase B−potato inhibitor (4SGB; Greenblatt et al., 1989); *S. griseus* proteinase B−third domain of turkey ovomucoid inhibitor (3SGB; Read et al., 1983).

Positional differences and corresponding standard deviations between all pairs of atoms involving the trypsin−tripeptide boronate complex and the consensus structure were calculated. To probe for correlated structural differences between the complex and the consensus structure, vectors corresponding to the average differences between corresponding segments were calculated using a "moving window" that included each backbone target atom plus three atoms before and after it (Katz & Kossiakoff, 1990; Weber, 1987). For each residue the largest of these main-chain vectors was displayed on the corresponding main-chain atom. For each non-disordered residue average residue displacement vectors were also calculated and displayed at their centers of masses. All vectors are scaled by a factor of 10 and displayed as double-headed arrows in Figures 6 and 7 of superimposed structures. Arrows are also color coded and thickened to represent their statistical significance (calculated with respect to the five reference structures).

RESULTS

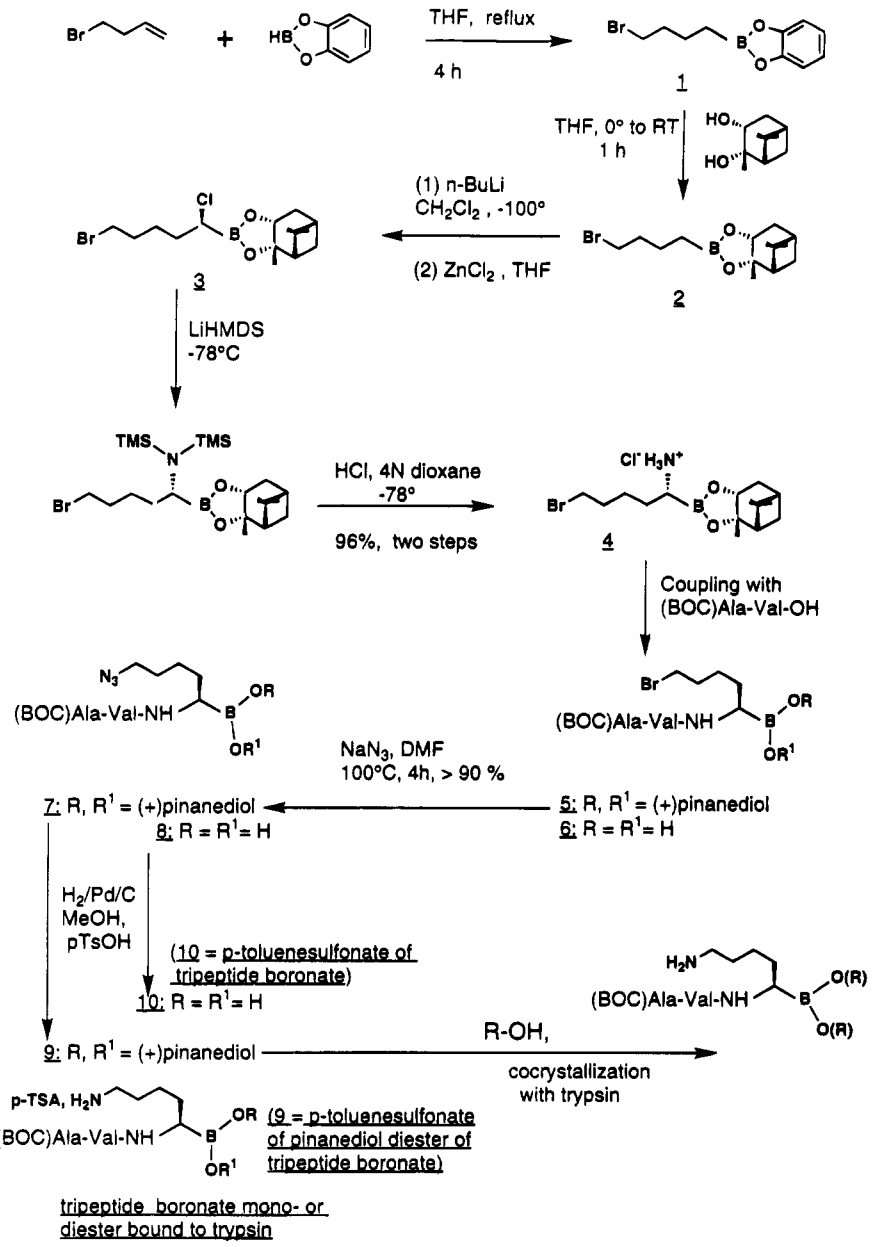
Synthesis and Assay of (BOC)Ala-Val-Lys-boronate (K_i = 7 nM). Synthesis of the tripeptide boronate was as described by Kettner and Shenvi (1984) and Kettner et al. (1988) for a homolog (Scheme 1). Thus 4-bromobutylboronate catechol **1** was prepared by hydroboration of 4-bromo-1-butene with catecholborane. The transesterification of **1** in the presence of (+)-pinanediol afforded chiral boronic ester **2**, which was transformed to the (1*S*)-5-bromo-1-chloropentylboronate ester of pinanediol **3** using Matteson's stereospecific homologation (Matteson & Majumber, 1983). Substitution of chloride by an amino group was achieved by treatment of **3** with lithium bis(trimethylsilyl)amine at −78 °C, followed by removal of trimethylsilyl groups at low temperature with HCl in dioxane to give compound **4**. The coupling reaction between **4** and (BOC)Ala-Val-OH yielded peptide boronic ester **5** and the corresponding peptide boronic acid **6**, which were separated by flash chromatography. Compound **5** was transformed to **7** by replacement of bromide with an azido group to provide intermediate **8**, which was hydrogenated in the presence of *p*-toluenesulfonic acid to give **9** (tripeptide boronate pinanediol diester *p*-toluenesulfonate salt). The same last two steps were applied to peptide boronic acid **6** to synthesize **10**, (BOC)Ala-Val-boroLys PTSA (tripeptide boronate *p*-toluenesulfonate).

The tripeptide boronate **10** is a potent trypsin inhibitor with a K_i determined to be 7.0 nM.

Boronate Geometry. All the tripeptide boronates or boronate esters reported here have features in common as follows. The peptidylboronate moiety is covalently bound to Ser195 in a tetrahedral complex resembling the transition state formed during the deacylation step of the proteolysis reaction. Figure 1 shows the ($|F_o| - |F_c|$), α_c map of the parent, unesterified trypsin−tripeptide boronate complex superimposed onto its refined structure.

The geometry about the boron is distorted tetrahedral. In Table 2, the bond lengths and bond angles involving boron obtained in the refinement of several trypsin−tripeptide boronate complexes are compared to the corresponding averages in seven α-lytic protease−peptidylboronate structures [Bone et al., 1989, and Brookhaven Data Bank (Bernstein et al., 1977) Entry Codes 1P08 and 1P10 by Bone and Agard]. In both systems the B−Oγ195 bond is

Scheme 1



significantly longer than the other B−O bonds. In trypsin−tripeptide methyl ester, since a methyl group is bound to O1, the B−O1 bond length might be expected to be similar to the B−Oγ195 bond length, but it is significantly shorter. In Table 2 bond lengths and angles involving boron in the peptidylboronate enzyme complexes are also compared with those in some small molecule crystal structures. The boron−oxygen bond lengths compare favorably with those in small molecule tetrahedral complexes; the boron−carbon bond lengths are somewhat longer in the peptidylboronate enzyme complexes (Table 2).

In the structures of trypsin−tripeptide boronate, of trypsin−tripeptide boronate esters, and of peptidylboronate−α-lytic proteases in Table 2, the geometry about the boron is about midway between the tetrahedral and trigonal planar adduct, with Oγ195 as the pyramidal capping ligand. In a purely tetrahedral complex the distance from the boron to the trigonal plane defined by O1, O2, and Cα would be about 0.95 Å. But in trypsin−tripeptide boronate and derived ester complexes the corresponding distance is 0.39(5) Å; in α-lytic protease complexes it is 0.37(10) Å (Table 2).

This geometry about the boron in the trypsin−tripeptide boronate complexes is very similar to that in the four-coordinate boron unit in β-HBO_2. The latter compound has three short B−O bond lengths, 1.45(1) Å, whereas the former series exhibit bond lengths of 1.49(1) Å and 1.47(1) Å for B−O1 and B−O2, respectively. The longer boron−oxygen bond length (1.55 Å) in the four-coordinate boron unit in β-HBO_2 involving an uncharged capping water molecule is the same as the B−Oγ195 bond length [1.56(1) Å] in the peptidylboronate enzyme complexes. The distortion from the tetrahedral to trigonal planar adduct in the four-coordinate boron unit of β-HBO_2 is in the direction of the B−H_2O vector while in the peptidylboronate enzyme complexes it is toward the B−Oγ195 vector, and the displacements from corresponding trigonal planes are similar (Table 2).

In trypsin−tripeptide boronate and trypsin−tripeptide boronate esters, in addition to the normally observed

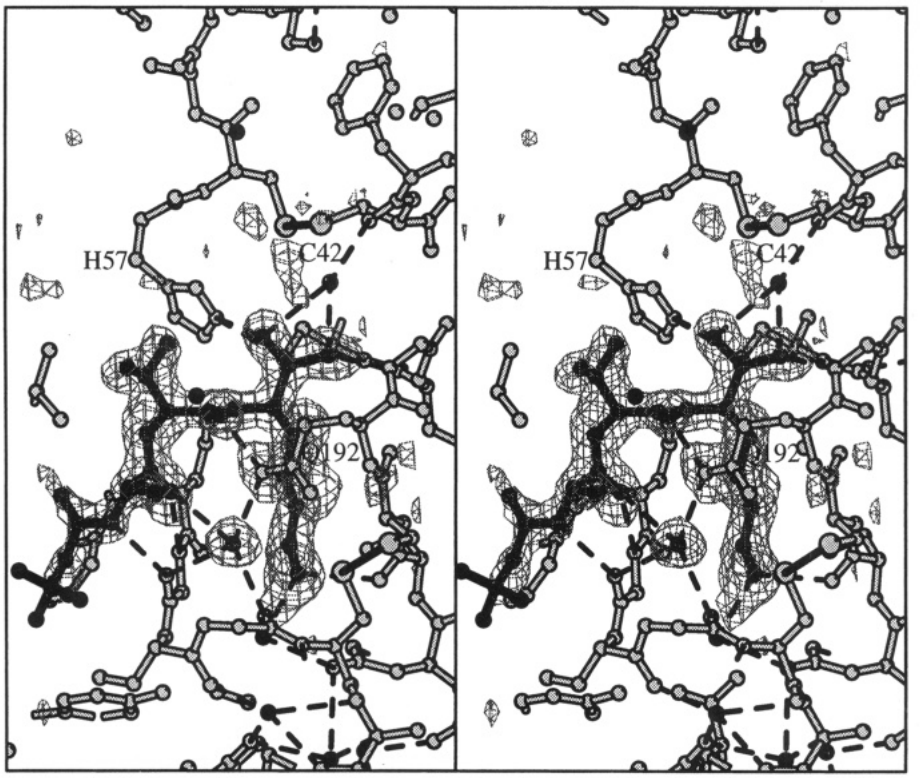


FIGURE 1: Refined structure of trypsin−tripeptide boronate superimposed on the original ($|F_o| - |F_c|$), α_c map calculated with F_c's derived from the coordinates of MIP−trypsin (Chambers & Stroud, 1977; Finer-Moore et al., 1992) with the inhibitor removed. Hydrogen bonds, denoted by dashed lines, show a well-ordered water molecule which is the donor to the carbonyls of Gly216 and Gly219 and the carbonyl of Ala at P3 and the acceptor from the NH_2 of Gln192. The side chain of Gln192 becomes ordered on binding the peptidylboronate. The density for the *tert*-butyl group of the inhibitor appeared in difference maps during refinement of the structure. Visualization by MOLSCRIPT (Kraulis, 1991).

hydrogen bond from $N_{\delta 1}57$ to the Asp102 side chain, the $N\epsilon 2$ atom of His57 also forms short hydrogen bonds to the $O\gamma$ atom of Ser195 (2.86 Å), as well as to the boronate O2 atom (2.87 Å) (Table 3). The hydrogen-bonding pattern of His57 in the peptidylboronate−α-lytic proteases is similar, but in those complexes a shorter hydrogen bond is made to O2 than to $O\gamma 195$, whereas in trypsin−tripeptide boronate these two hydrogen bonds are identical (Table 3).

Substrate-like Interactions in Trypsin−Tripeptide Boronate and Ester Complexes. At the P2 site the Val side chain of the inhibitor makes hydrophobic contact with Leu199 and His57 of the protein (Table 4). Its $\chi 1$ angle of 133° is rare (Ponder & Richards, 1987) and of high dihedral energy (Weiner et al., 1984). The Ala side chain and the hydrophobic part of the BOC group are not in close contact with any group of the protein (closest carbon−carbon contact involving BOC = 4.3 Å to Cζ3 of Trp215). The carbonyl oxygen of the BOC may be bound to a loose water, Wat417 (B = 53 Å^2). The lack of favorable contacts is consistent with the relatively high temperature factors for the BOC group compared with the other non-hydrogen atoms in the inhibitor or with all heavy atoms in the protein (Table 4). The relatively enhanced mobility of the BOC group probably explains the inability to locate it in the initial maps.

The backbone of the tripeptide boronate and tripeptide boronate esters makes two good antiparallel β-sheet hydrogen bonds to the backbone of the protein, N_{Ala}−O216 and O_{Ala}−N216 and a weaker one from N_{Lys}−O214 (Table 3). The backbone dihedrals of the inhibitor (Table 4) are characteristic of twisted β-sheet geometry (Salemme, 1981). The backbone of the inhibitor also hydrogen bonds to the $N_\epsilon 2$ atom of the Gln192 side chain. $N_{\epsilon 2}192$ is also hydrogen bonded to Wat409 which in turn is within hydrogen-bonding distance to the Ala (P3) carbonyl oxygen of the inhibitor, as well as to two other main-chain carbonyl oxygens of the protein, O216 and O219 (Table 3).

In the P1 pocket, the amine group of the well-ordered Lys side chain of the inhibitor forms a salt bridge with Asp189 (Table 4). The dihedral angles of the lysine side chain (Table 4) correspond to the most common conformation (for $\chi 1$ and $\chi 2$) observed for lysines in protein crystal structures (Ponder & Richards, 1987).

"Episelection": Methyl Ester at O1. For methanol-treated trypsin−peptidylboronate, the ($2|F_o| - |F_c|$), α_c electron density map shows good density for a methyl ester linkage extending from the boronate oxygen O1 [labeled according to Bone et al. (1989)] that occupies the oxyanion hole (Figure 2). The refined occupancy of this methyl group is unity; the boronate is 100% esterified with methanol at this oxygen alone. Mass spectrometry of single crystals of the complex confirmed the presence of an additional single methyl group (Figure 3). The methyl group is 3.73 Å from the carbonyl oxygen of residue 41 and 4.18 Å from the $S\gamma 42$ atom of the Cys42−Cys58 disulfide. A well-ordered water molecule, coordinated by both the inhibitor and the protein (Wat409), and an ordering of the Gln192 side chain are also observed as a result of inhibitor binding. There is also a water molecule hydrogen bonded to the boronate O2.

The O1−methyl and $O\gamma 195$−$C\beta 195$ bond lengths are 1.46 and 1.43 Å, respectively, similar to O−C bond lengths where the oxygen is bound to boron in small molecule structures. The B−O1−CH_3 angle is similar to B−O−C angles in small

Table 2: Comparison of Geometry Involving Boron in Small Molecule Crystal Structures, in Peptidylboronate−α-Lytic Proteases, in Trypsin−Tripeptide Boronate, and in Trypsin−Tripeptide Boronate Methyl Ester Structures

	small molecules			peptidylboronate−enzymes	
parameter	trigonal planer	(pseudo)tetrahedral	four-coordinate boron in β-HBO_2	trypsin−tripeptide boronate[a]	α-lytic proteases−peptidylboronate[b]
bond length (Å)					
B−O	1.37(1)[c]	1.47(3)[d]	1.45(1)		
B−OH_2			1.55		
B−Oγ195				1.56(1)	1.55(2)
B−O1				1.49(1)	1.51(2)
B−O2				1.47(1)	1.49(1)
B−C	1.56(1)[e]				
B−Cα				1.63(1)	1.63(6)
O−C	1.45(1)[f]	1.41(3)[g]			
O1−CH_3				1.46[h]	
Oγ195−Cβ195				1.43(1)	1.43(2)
bond angle (deg)					
O−B−O	115(1)[e]	108(6)[i]			
O1−B−O2				113(2)	117(4)
Oγ195−B−O1				104(2)	103(4)
Oγ195−B−O2				100(1)	99(7)
O−B−C	123(2)[j]				
O1−B−Cα				119(1)	118(2)
O2−B−Cα				112(1)	113(4)
Oγ195−B−Cα				108(1)	103(2)
B−O−C		117(1)[k]			
B−Oγ195−Cβ195				127(1)	132(4)
B−O1−CH_3				120[h]	
B−trigonal plane (Å)[l]			0.45	0.39(5)	0.37(10)

[a] Bond length and angle averages and standard deviations are for the two data sets for trypsin-tripeptide boronic acid. [b] Averages are for seven structures (Bone et al., 1989, and Brookhaven Data Bank Entry Codes 1P08 and 1P10). [c] Average of the 16 bond lengths in β-HBO_2 (Zachariasen, 1963), in D-mannitol tris(benzeneboronic ester) (Gupta et al., 1977), and in phenylboronic acid (Rettig et al., 1977). [d] Average for β-HBO_2, γ-HBO_2, H_3BO_3, $KB_5O_8{\cdot}4H_2O$, and $K_2B_4O_7{\cdot}4H_2O$ (Table 6 of Zachariasen, 1963), $(MeOH)_2Li(u\text{-}OMe)_2B(OMe)_2$ (Al-Juaid et al., 1989), and piperidinium tetramethoxyborate (Alcock et al., 1982). [e] Average of five parameters (Gupta et al., 1977; Rettig &Trotter, 1977). [f] Average of six bond lengths (Gupta et al., 1977). [g] Average of six bond lengths (Al-Juaidet al., 1989; Alcock et al., 1982). [h] Trypsin−tripeptide boronate methyl ester. [i] Average of eight angles (Al-Juaid et al., 1989; Alcock et al., 1982). [j] Average of ten angles (Gupta et al., 1977; Rettig & Trotter, 1977). [k] Average of five angles (Al-Juaidet al., 1989; Alcock et al., 1982). [l] Distance of boron to the plane defined by the non-water oxygens in the four-coordinate boron unit in β-HBO_2 or to the plane defined by O1, O2, and Cα of boroLys in the peptidylboronate complexes of trypsin or the corresponding distance in peptidylboronate−α-lytic proteases.

molecule boron structures, whereas the B−Oγ195−Cβ195 angles are somewhat larger in the peptidylboronate enzyme structures (Table 2).

Ethyl Esters at O1 and O2. Figure 4 shows an ($|F_o| - |F_c|$), α_c map in which no ester groups are included in the refined model. The density indicates an ethyl group disordered between two orientations on the boronate oxygen that is not in the oxyanion hole (O2), as well as an ethyl group on the boronate oxygen that is in the oxyanion hole (O1). The density for the latter ethyl group indicates that it is also disordered or of partial occupancy (refined occupancy = 0.32). The latter ethyl group makes van der Waals contact with the Cys42−Cys58 disulfide (CH_3−Sγ42 = 3.41 Å).

In ($|F_o| - |F_c|$), α_c difference maps obtained using models where the orientations and occupancies and temperature factors of the esterifying ethyl groups were refined, a residual positive peak was observed between boronate O1 and the carbonyl oxygen of residue 41. This peak is located at the position where a water molecule resides in the trypsin−tripeptide boronic acid complex hydrogen bonded to carbonyl 41 and the boronate O1. A model with partially occupied water at this site (as well as a partially occupied ethyl group) yielded an occupancy of 0.42 (B = 26 Å^2).

1,3-Propanediol Esterifies O1 and Extends into the P2′ Site. The initial ($|F_o| - |F_c|$), α_c and ($2|F_o| - |F_c|$), α_c maps

Table 3: Comparison of Hydrogen Bond Lengths in Trypsin−Tripeptide Boronate, in α-Lytic Protease−Peptidylboronate Complexes, and in Trypsin(ogen)−BPTI

hydrogen bond (Å)	trypsin−tripeptide boronate[a]	peptidylboronate−α-lytic proteases[b]	trypsin(ogen)−BPTI[c]
O1−N193	2.78(2)	2.53(8)	
O1−N195	3.00(1)	3.04(7)	
$N_{\epsilon 2}$57−Oγ195	2.86(4)	3.01(6)	2.62(8)
$N_{\epsilon 2}$57−O2	2.87(4)	2.69(9)	
N_{P3}−O216[d]	3.01(1)	2.99(8)	
O_{P3}−N216	2.96(2)	3.03(7)	3.15(12)
N_{P1}−O214	3.00(3)	3.12(9)	3.45(16)
N_{P1}−Oγ195			3.12(3)
O_{P2}−$N_{\epsilon 2}$192	2.80(1)		2.82(7)
Wat301−$N_{\epsilon 2}$192	2.92(1)		
Wat301−O_{P3}	3.25(10)		
Wat301−O216	2.94(2)		
Wat301−O219	2.84(1)		
Wat332−O1	2.83(1)		
Wat332−O2	3.20(2)		
Wat332−O41	2.70(1)		

[a] Hydrogen bond averages and standard deviations are for the two trypsin−tripeptide boronic acid data sets. [b] Averages and standard deviations are for 15 structures (Bone et al., 1991a,b). [c] Brookhaven Data Bank Entry Codes 2TPI (Kossiakoff et al., 1976; Kossiakoff, 1984) and 2PTC and 2TGP (Marquart et al., 1983). [d] P1 refers to the inhibitor residue before the scissile bond in the corresponding substrate; residues P2 and P3 extend toward the N-terminus (Schecter & Berger, 1987).

Table 4: Comparison of Geometry in Trypsin−Tripeptide Boronate, in Peptidylboronate Complexes of α-Lytic Protease, and in Protein Inhibitor Complexes of Trypsin-like Proteases

inhibitor parameter type	trypsin−tripeptide boronate	peptidylboronate−α-lytic proteases[a]	trypsin-like protease−protein complexes[b]
dihedral (deg)			
φ_{P3}	−155	140(3)	−78(12), −118[c]
ψ_{P3}	143	153(2)	−28(7), 149[c]
φ_{P2}	−67	−68(3)	−73(11)
ψ_{P2}	142	146(3)	156(8)
φ_{P1}	−107	−113(3)	−109(9)
$\chi 1_{P1}$	−69	−68(7)	−68(6)
$\chi 2_{P1}$	172		−175(6)
$\chi 3_{P1}$	−172		180(8)
$\chi 4_{P1}$	−179		176(11)
salt bridge (Å)			
$N\epsilon-O_{\delta 1}189$	3.40		3.54(15)
$N\epsilon-O_{\delta 2}189$	3.82		3.78(15)
B (Å^2)			
boroLys[d]	13(1)		
BOC[e]	40(2)		
Ala-Val-boroLys[e]	17(5)		
protein[e]	15(8)		
carbon−carbon contact (Å)			
$C_{\gamma 2}Val-C_{\delta 2}199$	3.49		
$C_{\gamma 1}Val-C_{\delta 2}57$	3.42		

[a] Averages and standard deviations for the seven α-lytic protease structures referenced in Table 3. [b] Averages for 2TGP, 2PTC, 2TPI, 1TGS, and 2KAI (Bernstein et al., 1979). [c] Values for 1TGS are not included in the averages and are listed separately. [d] Average for Cα, Cβ, Cγ, Cδ, Cϵ, and Nζ. [e] Average for non-hydrogen atoms.

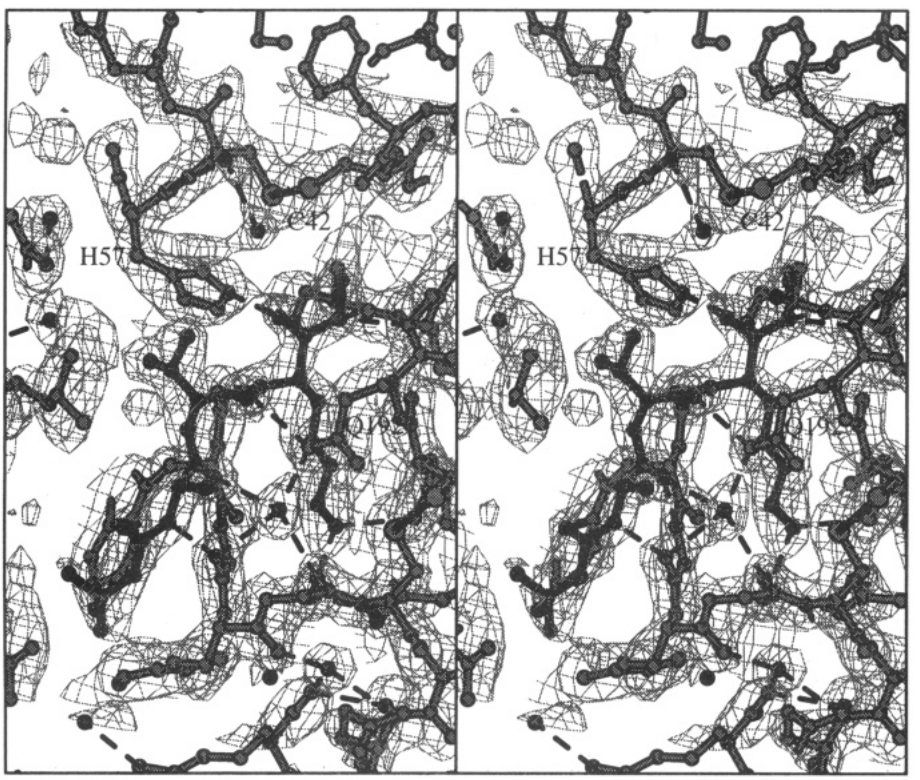


FIGURE 2: Refined structure of trypsin−tripeptide boronate methyl ester superimposed on the final $2(|F_o| - |F_c|)$, α_c map in the region of the boronate. His57 is the donor to O2, and the methyl substituent attached to O1 lies upward toward Cys41.

obtained when a model derived from the structure of trypsin−tripeptide boronic acid was refined to convergence were not as interpretable, as for other trypsin−tripeptide boronate complexes in this study. In the $(2|F_o| - |F_c|)$, α_c map there was continuous density extending from Oγ195 through a position near the boron to the oxyanion hole but much weaker density from Cα of boroLys to the boron. The temperature factor of Oγ195 was much larger [27 Å^2 ($u = \pm 0.58$ Å)] than its value in the trypsin−tripeptide boronic acid structures (12 Å^2) or in trypsin−benzamidine [18 Å^2 ($u = \pm 0.39$ Å)], indicating a somewhat less fixed position. Gln192 was also disordered or mobile with a temperature of 42 Å^2 at Oϵ1 compared with 17 Å^2 in the trypsin−tripeptide boronic acid structure, where it is highly ordered. The density for the entire inhibitor was weaker in the $(2|F_o| - |F_c|)$, α_c map than for the other trypsin−tripeptide boronate and boronate ester complexes. These observations suggested disorder involving the tripeptide boronate inhibitor, which nevertheless samples the environment in the P2′ site.

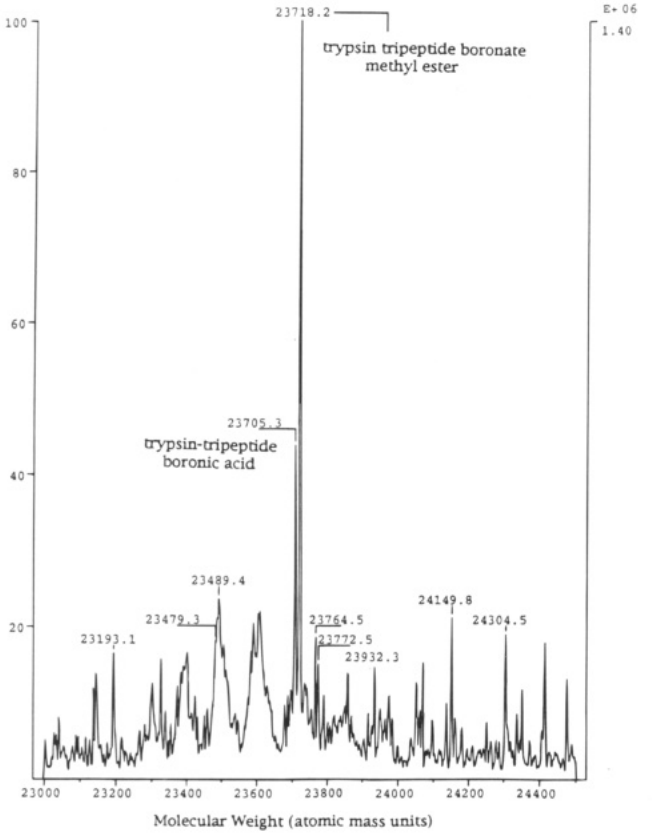


FIGURE 3: Mass spectrograph of freshly dissolved single crystals of trypsin−tripeptide boronate methyl ester uniquely identifying the boronate methyl ester as the major peak.

An ($|F_o| - |F_c|$), α_c map and a ($2|F_o| - |F_c|$), α_c map were calculated after refinement to convergence of a model lacking boron, boronate oxygens, and ester substituents. On the basis of these maps and on the results of the refinement and analysis of other models (and resulting maps) discussed in Experimental Procedures, a major component in the structure of trypsin−tripeptide boronate 1,3-propanediol ester is a boronate bound to Cα of boroLys, Oγ of Ser195, a hydroxyl group, and one oxygen of a 1,3-propanediol monoester whose oxygen (O1) occupies the oxyanion hole. In this component the alcohol is disordered at its free hydroxyl terminus from where it hydrogen bonds in one orientation to O41 and to a water molecule (Wat337) and in the alternate orientation to another water molecule (Wat579). As in the other tetrahedral boronate complexes, His57 makes good hydrogen bonds with Ser195 ($N_{\epsilon 2}57-O\gamma 195 = 2.85$ Å) and with a boronate oxygen ($N_{\epsilon 2}57-O2 = 2.73$ Å). The oxygen in the oxyanion hole O1 makes a hydrogen bond with NH193 (2.9 Å) but is further from NH195 (3.5 Å). The final structural model incorporating these features is shown in Figure 5 superimposed on the resulting ($2|F_o| - |F_c|$), α_c map.

Plasticity and Conformational Change on Binding. Binding of all tripeptide boronate and tripeptide boronate esters induces subtle conformational changes, as seen before in binding the transition-state analog MIP versus P1 site occupation by benzamidine (Krieger et al., 1974). Changes incurred by binding of the tripeptide boronate methyl ester were the most fully analyzed and are described here.

The Gln192 is disordered in the reference structures (Figure 6a, 7a) and also in MIP−trypsin (4PTP, Chambers & Stroud, 1979, and 1NTP, Kossiakoff, 1984), and trigonal trypsin (Kossiakoff et al., 1976; Walter et al., 1982). However, in the methyl ester complex, Gln192 becomes ordered [average side chain $B = 8.9$ Å^2 ($u = 0.34$ Å)], adopting a conformation ($\chi 1 = -84°$, $\chi 2 = 70°$, $\chi 3 = 48°$) rarely observed for Gln residues in proteins (Ponder & Richards, 1987). This conformation is the same as that

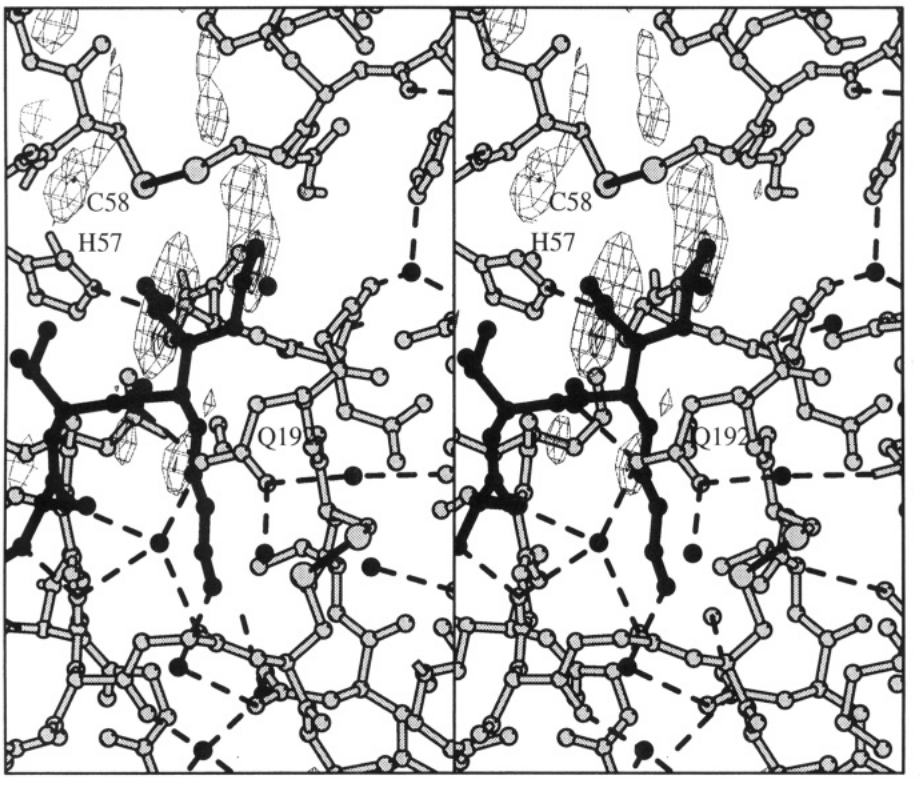


FIGURE 4: Refined structure of trypsin−tripeptide boronate ethyl ester in the region of the boronate superimposed on the ($|F_o| - |F_c|$), α_c map that is obtained when ester groups are not included in the model. The map shows partial occupancy at each oxygen, O1 and O2.

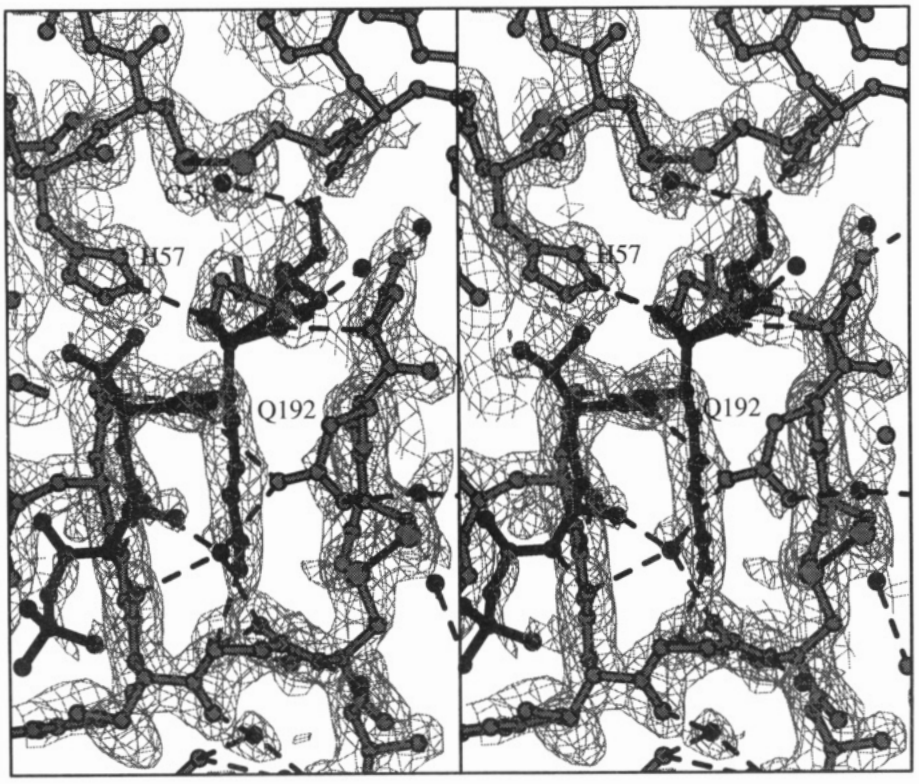


FIGURE 5: Refined structure of tripeptide boronate 1,3-propanediol ester showing the boronate components superimposed on the resulting $(2|F_o| - |F_c|)$, α_c map.

Table 5: Deviation of Core Atoms (rms in Å) between Crystal Structures of Inhibitor-Free Trypsin, Benzamidine-Inhibited Trypsin, and Trypsin−Tripeptide Boronate Methyl Ester[a]

	1TPO[b]	2PTN[b]	3PTB[b]	1BNZ[c]	2BNZ[c]	MeOH[c]
1TPO		0.090[d]	0.102	0.129	0.139	0.171
2PTN			0.095	0.107	0.116	0.166
3PTB				0.089	0.102	0.160
4PTB					0.063	0.141
5PTB						0.140
CONS[e]						0.141

[a] Pairs of structures were superimposed on the basis of the corresponding "core" main-chain atoms identified using difference distance matrices. [b] Protein Data Bank Entry Codes are explained in Experimental Procedures. [c] 1BNZ and 2BNZ are two trypsin−benzamidine crystal structures determined in our laboratory. MeOH refers to the structure of trypsin−tripeptide methyl ester. [d] rms deviation of core atoms in angstroms. [e] Consensus structure is derived from two inhibitor-free and three benzamidine-inhibited trypsin crystal structures (see Experimental Procedures).

observed for Gln192 in trypsin−BPTI (2PTC; Marquart et al., 1983). In both trypsin−tripeptide boronate methyl ester and trypsin−BPTI the Nϵ2 atom of the amide side chain of Gln192 is hydrogen bonded to a main-chain carbonyl oxygen of the inhibitor.

In Table 5 are the rms deviations for "core" backbone atoms involving trypsin−tripeptide methyl ester and each of five reference structures (see Experimental Procedures), which do not involve inhibitor binding at the oxyanion hole. Also listed are the rms deviations for the core backbone atoms involving each pair of the reference structures alone, as well as that involving the trypsin−tripeptide boronate methyl ester and the consensus structure (derived from the five reference structures). The rms deviations between each of the reference structures and trypsin−tripeptide boronate methyl ester are somewhat greater (~0.14 Å) than that for each pair of reference structures (~0.08 Å).

The structure of the trypsin−tripeptide boronate methyl ester, superimposed on the five reference structures, and the consensus structure, along with the average *main-chain* displacement vectors, calculated in a moving window between trypsin−tripeptide boronate methyl ester and the consensus structure (see Experimental Procedures), are shown in Figure 6. In Figure 7 the average *residue* displacement vectors, complementary to the running average main-chain displacement vectors of Figure 6, are represented for non-disordered residues.

Local alterations of the enzyme structure (up to 0.35 Å, 6.9σ) are produced from the contact of the P2 Val with the side chain of Leu99. The χ1 dihedral of this Val is 47° from an optimal value, and the the χ2 dihedral of Leu99 is 26° (7.6σ) from its optimal, low-energy value in the consensus structure (Table 6). These changes typically correspond to calculated energies of ~0.4−1.3 kcal/mol each. The χ1 angle of Ser96 (Table 6), ψ98, and φ99 (Table 7) are also perturbed. Large concerted main-chain shifts (~0.3 Å, 6σ) of the loop residues 96−98, away from the Val of the inhibitor, are represented by white arrows (lengths >5.0σ) in Figure 6. Perturbations propagate through residues 101−103. The magnitudes and significance of both the concerted main-chain and residue displacement vectors for residues 94−103 around Asp102 (segment A) are summarized in Table 8. Subtle but significant differences in the side-chain dihedrals (~10°, 4σ) of Ile103, Leu105, Ile106, Lys107, and Leu108 (Table 6) and in the main-chain dihedrals of Asn101, Ile103, and Met104 (Table 7) are observed.

Perturbations in residues 103−108 are transmitted to an adjacent antiparallel β-strand, segment B, residues 84−91

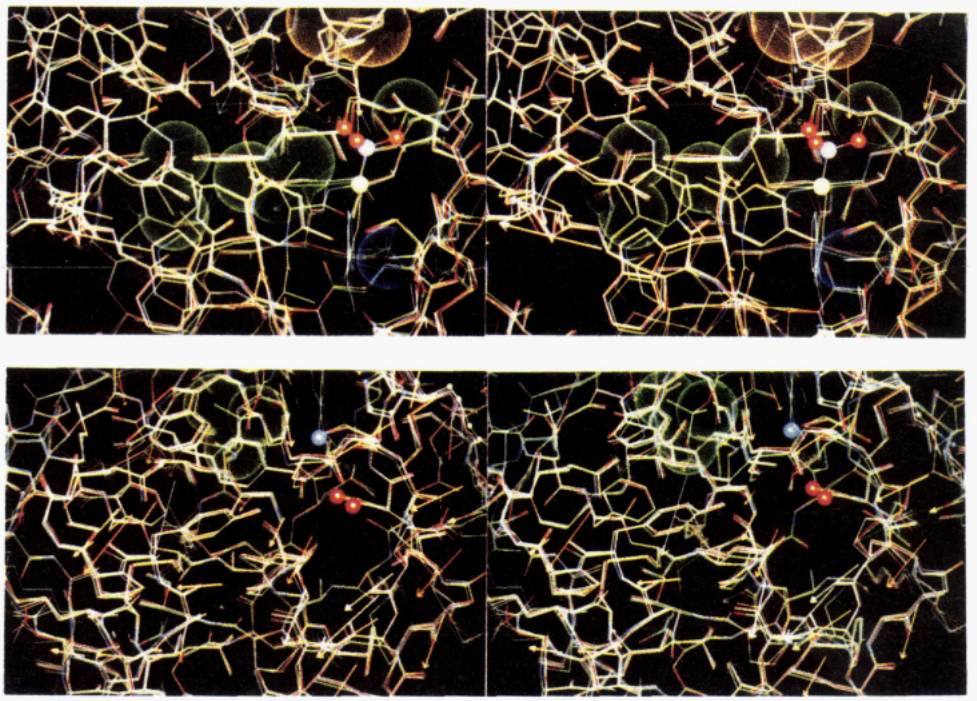


FIGURE 6: (a, top) Structure of trypsin−tripeptide boronate methyl ester superimposed on the five reference structures in the vicinity of the active site. The boron (blue), oxygen (red), and carbon (white) directly attached to it are shown as balls. Van der Waals surfaces around the side chains of P2 Val and Leu99 in contact and around the methyl group and the contacting Cys42−Cys58 disulfide are shown as dotted surfaces. Superimposed on the structures are average main-chain displacement vectors in angstroms scaled by a factor of 10 (between atoms in trypsin−tripeptide boronate methyl ester and the corresponding ones in the consensus structure) calculated in a moving window. White signifies lengths with significance greater than 5.0σ, orange 4.0−5.0σ, yellow 3.0−4.0σ, and thin light blue 2.75−3.0σ. (b, bottom) Viewed from behind the P1 pocket, the structure of trypsin−tripeptide boronate methyl ester superimposed on the five reference structures. The positively charged Nϵ atom of the boroLys at P1 (blue) and contacting negatively charged Oδ1 and Oδ2 (red) of Asp189 are highlighted with balls. Van der Waals surface (dotted) is around the BOC group of the inhibitor. Average main-chain displacement vectors show marked closure of the P1 pocket around the lysine side chain, color coded as in (a).

(Figure 7 and Table 8). There are also perturbations of −18° in χ1 of Ser84 and of 6−16° in the main-chain dihedrals of Ser84, Ala85, Lys87, and Val90. The P1 specificity residue, Asp189, shifted by 0.19 Å (3.3σ), is another location where perturbations propagate through the β-sheet; Val17, which makes antiparallel β-sheet hydrogen bonds with Asp189, is also shifted (0.22 Å, 3.3 σ) in the same direction (Figure 7).

The most dramatic shifts for the entire structure occur within the loop comprising residues 221−225 at the base of the binding poccket (segment D) which both follow sequentially as well as adjoin spatially the residues forming the lip to the P1 pocket (segment C). Perturbations in the P1 pocket are transmitted to and amplified within this loop (Table 8). The directions of the shifts are similar to those observed for the lip of the P1 pocket. There are also large changes in φ222 (−12°, 8.1σ) and Ω223 (−5°, 16σ) (Table 7).

Structural perturbations are transmitted to additional regions of the protein which contact segment D (residues 221−225). Segment E, below segment D (Figures 6 and 7), comprising residues 184−187 moves ~0.16−0.30 Å (Table 8). The directionalities of the concerted main-chain shifts in segment E (Figure 7b) are similar to those observed in segments C and D (residues 214−224). Segment F comprising residues 164−173, also contacting the segment D loop (residues 221−225) into which structural perturbations are propagated, moves 0.14−0.34 Å in directions similar to those in segments C, D, and E (Figure 7, Table 7). In segments E and F there are also large perturbations in the main-chain dihedrals of Ser164, Ser166, Tyr172, Pro173, Ala183, Gly184, Leu185, Glu186, and Lys188A (Table 7).[3]

DISCUSSION

The tripeptide boronates described here are analogs of the tetrahedral intermediate, covalently attached to Ser195, and are potent inhibitors of trypsin. Incubation in the presence of alcohols produced ester complexes wherein substituents derived from alcohols were incorporated through ester linkages to boronate oxygen atoms. Differential esterification of either boronate oxygen on the enzyme surface offers the opportunity for alkyl groups derived from alcohols to make favorable hydrophobic interactions with the P2′ cleft abutting the Cys42−Cys58 disulfide (Figure 2). The region surrounding this cleft differs significantly among serine proteases and thus suggests a basis for engineering selectivity toward different proteases of therapeutic significance.

Esterifying Groups on the Trypsin-Bound Tripeptide Boronates Bind near P2′. The peptidylboronate esters define a general class of novel serine protease inhibitors. In addition to interactions with the oxyanion site, and with the P1 and P2 binding subsites, they utilize other regions, most specif-

[3] Some of the perturbations induced by the binding of the tripeptide boronate methyl ester are propagated to symmetry-related molecules via intermolecular contact regions especially for crystals corresponding to the smaller orthorhombic unit cell. Significant concerted main-chain shifts are observed in loop G (residues 201−204) opposite the active site of trypsin. Cys201 is shifted by 0.28 Å (4.6σ). This loop lies close to the side chain of P3 Ala from a symmetry-related molecule in the crystallographic unit cell (Figure 7 and Table 8). Another symmetry-related region proximal to perturbed region segment E (residues 184−187) includes the calcium binding loop residues (76−79); Asn79 is shifted by 0.32 Å (3.5σ) such that its Nδ2 atom forms a hydrogen bond with the Oδ1 atom of Asp25 (2.96 Å) versus 3.7(2) Å in the reference structures.

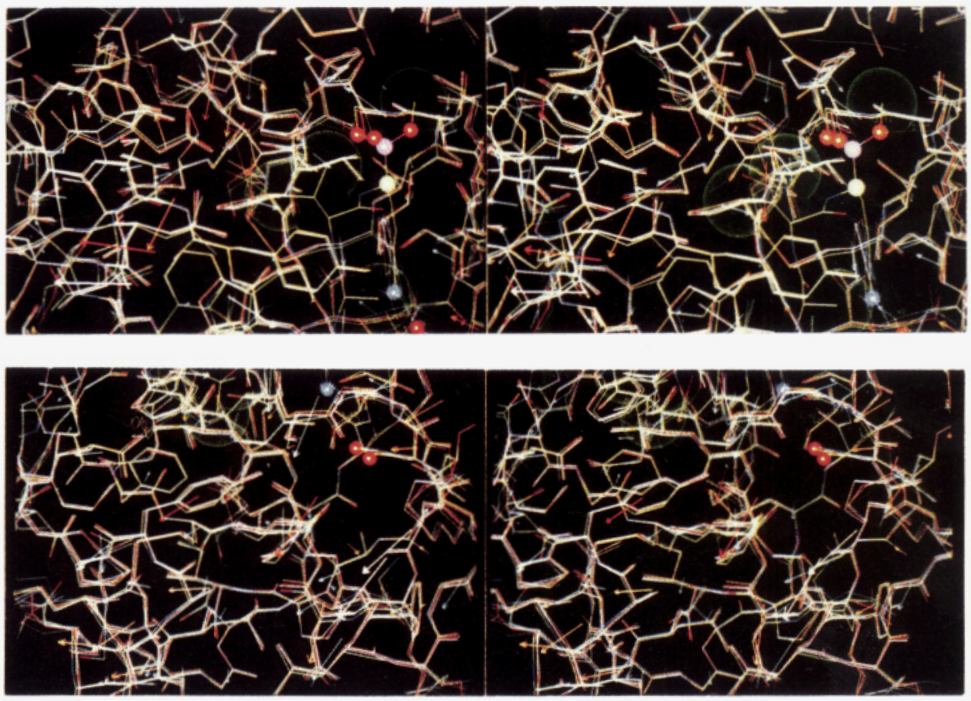


FIGURE 7: (a, top) As in Figure 6a, except that van der Waals surfaces of the Val and Ala side chains of the inhibitor and the methyl group bound to the boronate oxygen are shown as dotted surfaces and "average residue" displacement vectors are shown, in angstroms amplified 10 times, between atoms in trypsin−tripeptide boronate methyl ester and the corresponding ones in the consensus structure. White signifies lengths greater than 5.0σ, red 4.0−5.0σ, orange 3.0−4.0σ, thin yellow 2.5−3.0σ, and thin light blue 0.0−2.5σ. Segment C, the main-chain segment of trypsin with which the tripeptide boronate methyl ester makes antiparallel β-sheet hydrogen bonds, and which forms the entry lip to the P1 pocket, is moved by ~0.25 Å (3.5σ); Trp215 shows a residue shift of 0.29 Å (6.3σ), and perturbations occur in the χ1 dihedrals of Trp215 (−6°, 3.5σ) and of Cys220 (−10°, 4.2σ) (Table 6) and in Ω216 (−5°, 5σ) corresponding to ~0.6 kcal/mol. (b, bottom) As in Figure 6b, except showing average main-chain (as opposed to residue) displacement vectors times 10, color coded as in Figure 6. Though less dramatic than for side chains, there is significant closure around the P1 side chain.

Table 6: Side-Chain Conformational Differences between Trypsin−AP18 Methyl Ester and the Consensus[a] Trypsin Structure

residue	dihedral angle (deg)	consensus	trypsin−AP18 Me ester	$\Delta\chi$[b]	σ[c]	ΔE[d] (kcal)
Lys188A	χ1	−60(2)	−35	25	13.4	1.2
	χ2	−176(5)	−144	32	6.6	2.0
Ser110[e]	χ1	−62(2)	−88	-26	11.3	1.4
Leu99	χ1	−68(2)	−76	-8	4.2	0.4
	χ2	180(3)	154	−26	7.6	1.3
Ile121	χ2	64(2)	55	−9	6.4	0.0
Ile106	χ1	−61(2)	−75	−14	6.3	0.5
	χ2	166(3)	176	10	3.3	−0.4
Ser84	χ1	−58(3)	−76	-18	5.5	0.5
Leu108	χ2	178(2)	−171	11	5.2	0.2
Cys136	χ1	−55(2)	−66	−11	4.8	0.0
Lys107	χ1	−172(3)	−163	9	2.9	0.4
	χ3	−173(4)	171	−16	4.2	0.1
Cys220	χ1	−55(2)	−45	10	4.2	0.4
Trp215	χ1	63(2)	69	−6	3.5	0.1
Thr134	χ1	−55(2)	−48	7	3.1	0.2
Ile103	χ2	176(6)	158	−11	3.0	0.9
Ser139	χ1	47(4)	37	−10	2.9	0.7
Leu105	χ1	−73(2)	−67	6	2.9	−0.2
Ser96	χ1	59(5)	44	−15	2.9	0.5
Asn223	χ1	−73(2)	−67	6	2.6	−0.2

[a] The consensus structure is derived from two inhibitor-free and three benzamadine-inhibited trypsin structures (see text). [b] $\Delta\chi$ = AP18 methyl ester dihedral − consensus dihedral. [c] σ = Δ/rms deviation. [d] ΔE = dihedral energy (trypsin−AP18 methyl ester) − dihedral energy (consensus). Dihedral energies were calculated with X-PLOR. For these dihedrals, $E_{\text{dihedral}} = 1.6[1 + \cos(3\chi)]$. [e] Ser110 may be discretely disordered; the values for 4PTB (46°) and 5PTB (40°) were not used for averaging.

ically near the P2′ site, through both hydrophobic or hydrophilic interactions. The methyl and ethyl esters of trypsin−tripeptide boronate terminal ester methyl groups approach or attain van der Waals distance from Sγ42 (4.18 and 3.75 Å, respectively). For the 1,3-propanediol monoester of propanediol-treated trypsin−tripeptide boronate, the terminal hydroxyl group of the alcohol makes hydrogen bonds to carbonyl oxygen 41 in one of its orientations or to a water molecule in another observed orientation.

The competition and interplay between hydrophobic and hydrophilic forces in the region bounded by the O1 oxygen of the inhibitor boronate and the Sγ42−Sγ58 disulfide and carbonyl oxygen 41 of the protein can be seen in the range of interactions, orientations, and occupancies of alcohol groups esterified to O1 in the series of trypsin−tripeptide boronate ester crystal structures. In trypsin−tripeptide boronic acid a water molecule is hydrogen bonded to O1, O2, and carbonyl oxygen 41. In trypsin−tripeptide methyl ester, a fully occupied methyl group, bound to O1, is ~0.3 Å away from this water binding site. In trypsin−tripeptide ethyl ester, disorder allows a water molecule to partially occupy this site while an ethyl group, at partial occupancy, also extends through it to make contact with Sγ42. In trypsin−tripeptide 1,3-propanediol ester, a hydroxyl group of propanediol, hydrogen bonded to O41, occupies the site with partial occupancy.

In these trypsin−tripeptide boronate ester systems, the location, orientation, and occupancy in and around the P2′ site for methyl, methylene, and hydroxyl groups from the boronate ester inhibitors versus that for a competing water molecule yield insights regarding the differences in solvation energy changes accompanying the binding to trypsin of the peptidylboronate esters versus the corresponding peptidylboronic acids. Since all systems are crystallized from, and

Table 7: Listing, in Order of Significance, of Main-Chain Conformational Angular Differences between Trypsin−Tripeptide Boronate Methyl Ester and the Consensus[a] Trypsin Structure, in Columns $\Delta\varphi$, $\Delta\psi$, and $\Delta\Omega$

residue	$\Delta\varphi$ (deg)[b]	n_σ[c]	ΔE (kcal)[d]	residue	$\Delta\psi$ (deg)	n_σ	ΔE (kcal)	residue	$\Delta\Omega$ (deg)	n_σ	ΔE (kcal)[e]
Cys232	2	10.4	−0.2	Leu185	−12	8.7	0.4	Asn223	−5	16.0	0.4
Gly78	−17	8.1	0.0	Thr98	−6	7.3	0.0	Ser166	−8	13.0	1.4
Lys222	−12	8.1	1.0	Lys87	16	4.8	0.5	Ala85	−4	6.8	0.8
Gly184	21	5.7	0.0	Ala183	−13	4.7	0.0	Val53	−3	6.2	0.8
Ser130	−20	5.4	0.1	Asp194	−5	4.2	−0.1	Trp141	2	5.2	−0.1
Glu186	9	5.4	−0.1	Pro198	−6	3.7	−0.1	Gly216	−5	5.1	0.6
Asn25	6	4.9	−0.4	Ser164	10	3.5	0.0	Ala243	−4	4.6	0.4
Lys87	6	4.2	0.5	Ile121	10	3.4	0.0	Lys145	−4	4.4	0.3
Gln64	7	3.9	0.0	Gly78	−16	3.3	0.0	Pro173	3	4.4	0.0
Asp194	10	3.9	−0.1	Tyr29	−4	3.3	0.0	Lys188A	−8	4.3	−0.1
Val52	−6	3.8	0.0	Gly140	−6	3.2	0.0	Ile138	−1	4.0	0.1
Lys145	−6	3.7	0.0	Gly193	−9	3.1	0.1	Cys201	−3	3.6	0.4
Glu80	17	3.7	0.1	Val90	7	3.1	0.0	Leu155	5	3.4	−0.8
Gly193	8	3.4	0.1	Thr229	3	3.0	0.1	Ser84	−3	3.3	0.4
Gln175	5	3.4	0.0	Val227	10	3.0	0.0	Ser202	−3	3.2	0.3
Gly148	13	3.2	−0.2	Gln135	15	3.0	0.1	Thr144	1	3.2	−0.1
Pro173	8	3.2	−0.4	Ala243	9	3.0	−0.1	Gly196	1	3.1	0.0
Ile103	−11	3.2	−0.2	Ala85	−7	3.0	0.0	Val213	−2	3.0	0.2
Gln210	9	3.1	0.0	Pro28	−6	3.0	−0.1	Asn245	−6	2.7	1.0
Leu209	−8	3.0	0.2	Trp237	−5	3.0	−0.1	Gly133	−1	2.7	0.1
Val75	−8	3.0	0.0	Ala24	−5	3.0	−0.2	Met104	4	2.7	0.0
Cys136	10	3.0	−0.1	Leu163	−6	2.9	−0.1	Pro198	7	2.7	−0.2
Val199	−4	2.9	0.0	Asn233	5	2.8	0.1				
Ala85	−5	2.9	0.0	Tyr172	9	2.7	0.0				
Ser244	−7	2.9	0.0	Pro161	5	2.7	0.0				
Val227	−8	2.9	0.0	Gln210	−9	2.7	0.0				
Asn233	−5	2.8	0.1								
Gly142	5	2.8	0.1								
Asn101	5	2.7	−0.1								

[a] Defined in Table 5. [b] $\Delta\varphi = \varphi$(trypsin−tripeptide boronate methyl ester) − φ(consensus). [c] $n_\sigma = \Delta\varphi$/rms deviation. [d] $\Delta E = (\varphi, \psi$ (van der Waals) energy in trypsin−tripeptide boronate methyl ester) − $(\varphi, \psi$ energy in consensus). $E_{\varphi,\psi}$ was calculated with X-PLOR by evaluating and summing the following van der Waals interactions: {(H_i, N_i, $C\beta_i$) and (H_{i+1}, N_{i+1})}; {O_{i-1} and (H_{i+1}, N_{i+1})}; {O_{i-1} and (C_i, O_i, $C\beta_i$)}. [e] ΔE = (dihedral energy in trypsin−tripeptide boronate methyl ester) − (dihedral energy in consensus). Dihedral energies were calculated with X-PLOR; for Ω, $E_{\text{dihedral}} = 100.0[1 + \cos(\Omega)]$.

maintained in, largely aqueous systems, peptidylboronic acid complexes are relatively stable. In alcohol−water solutions an equilibrium exists between the free acid and mono- and diesters of the tripeptide boronate, and the enzyme specifically selects the component(s) with the highest affinities out of that equilibrium mixture. This "epitaxial selection" process on the enzyme surface leads to novel, selective inhibitors whose chemical structures and three-dimensional interactions with the enzyme, determined crystallographically, yield insight into subsequent design of related inhibitors with increased selectivity.

For the trypsin−tripeptide methyl ester, space filling by the methyl group alone may impart significantly greater binding energy to the methyl ester versus the boronic acid even though the methyl group does not make direct contact with Sγ42. However, it prevents a water molecule from hydrogen bonding to both the O1 boronate oxygen and O41, resulting in decreased entropy cost upon binding. For the trypsin−tripeptide ethyl ester, the site is shared by both a water molecule, when O1 is underivatized, and an ethyl group, corresponding to the esterified O1 oxygen. Although the terminal methyl of the ethyl group directly contacts Sγ42, the increased entropic cost of ordering the ethyl group (compared with that for the methyl group in the trypsin−tripeptide methyl ester system) probably plays a role in the observed disorder.

Peptidylboronates as Transition-State Mimics. The distances $N_{\epsilon 2}57$−Oγ195 = 2.86(4) Å and $N_{\epsilon 2}57$−O2 = 2.87(4) Å indicate strong, bifurcated hydrogen bonds (Figure 8). The orientation of the Nϵ2−H bond of His57 is directed midway between O2 and Oγ195. Thus, in the peptidylboronate transition-state analog, His57 lies in the ideal position required in a true transition state for its action as both a base to deprotonate Ser195 and an acid to protonate the substrate leaving group. These bond lengths are compared with the single, relatively strong hydrogen bonds between $N_{\epsilon 2}57$ and Oγ195 that average 2.60(10) Å in seven inhibited complexes of trypsin-like proteases,[2] versus the weaker hydrogen bonds seen in uninhibited trypsin-like serine proteases which average 3.06(23) Å for 38 structures in the Protein Data Bank (Bernstein et al., 1977). A charged (doubly protonated) His57 is also expected in all complexes analyzed here. The tetrahedral intermediate is also mimicked by the covalent bond between the boron and Oγ195, the hydrogen bonds between the boronate oxygen and the main-chain amides forming the oxyanion hole, the antiparallel β-sheet hydrogen bonds between the backbone of the inhibitor and the backbone of the enzyme at the active site, the interactions between the boroLys side chain and Asp189 in the P1 pocket, and the interactions of the peptidylboronate Val residue at the P2 subsite. The boronate inhibitors strictly differ from catalytic transition states for serine proteases, since the geometry about the boron is not strictly tetrahedral, and the negative charge resides formally in an orbital centered on the boron.

Boronate Geometry. The geometry of the boron moiety in tetrahedral trypsin−tripeptide boronate and esters is distorted part way between terahedral and trigonal planar with O1, O2, and C representing the putative plane. The Ser195γO−B bond is long (1.56 Å) compared with B−O1

Table 8: Structural Shifts between Trypsin–Tripeptide Boronate Methyl Ester and the Consensus Structure

	segment A									
shifts[a]	Tyr 94	Asn 95	Ser 96	Asn 97	Thr 98	Leu 99	Asn 100	Asn 101	Asp 102	Ile 103
main chain (Å)	0.18	0.13	0.20	0.35	0.34	0.21	0.11	0.17	0.18	0.18
σ[b]	2.4	2.8	4.7	6.9	6.6	4.3	2.8	4.1	3.3	3.2
residue (Å)	0.15	0.17	0.15	0.34	0.34	0.25	0.10	0.10	0.16	0.22
σ	4.3	2.3	2.3	4.9	6.4	3.8	1.7	1.8	2.8	3.0

	segment B							
shifts[a]	Ser 84	Ala 85	Ser 86	Lys 87	Ser 88	Ile 89	Val 90	His 91
main chain (Å)	0.16	0.16	0.18	0.07	0.14	0.16	0.21	0.21
σ	4.2	4.2	3.9	1.6	2.7	3.0	3.9	3.6

	segment C					
shifts[a]	Ser 214	Trp 215	Gly 216	Ser 217	Gly 219	Cys 220
main chain (Å)	0.21	0.19	0.28	0.25	0.31	0.23
σ	3.8	3.4	3.4	2.7	3.2	3.8

	segment D					
shifts[a]	Ala 221	Gln 221A	Lys 222	Asn 223	Lys 224	Pro 225
main chain (Å)	0.30	0.38	0.47	0.45	0.35	0.23
σ	4.6	5.9	7.5	6.3	4.7	3.2
residue (Å)	0.24	dis[c]	dis	0.38	0.37	0.22
σ	3.3			3.2	4.7	2.1

	segment E			
shifts[a]	Gly 184	Tyr 184A	Leu 185	Gly 187
main chain (Å)	0.17	0.28	0.30	0.16
σ	3.4	3.7	3.8	3.0

	segment F									
shifts[a]	Ser 164	Asp 165	Ser 166	Ser 167	Cys 168	Lys 169	Ser 170	Ala 171	Tyr 172	Pro 173
main chain (Å)	0.21	0.20	0.16	0.14	0.16	0.20	0.18	0.20	0.22	0.20
σ	3.7	3.2	2.3	2.0	2.7	2.8	2.5	3.6	4.2	3.5
residue (Å)	0.34	dis	dis	dis	0.12	dis	dis	0.22	0.23	0.19
σ	3.9				1.7			2.9	4.8	2.2

	segment G						
shifts[a]	Pro 198	Val 199	Val 200	Cys 201	Ser 202	Gly 203	Lys 204
main chain (Å)	0.17	0.16	0.19	0.30	0.30	0.25	0.24
σ	3.6	3.5	3.7	6.3	5.7	4.1	3.6

[a] Main-chain shift is a vectorial average in a running window which includes the target atom as well as three atoms before and after it. For each residue the most signficant shift is listed. Residue shift is a vectorial average of all main chain and side chain atoms in the residue. [b] σ = (absolute magnitude of change in shift)/(rms deviation based on the consensus structures). [c] Disordered.

and B–O2 (1.49 and 1.47 Å). However, this geometry resembles that in the uncharged four-coordinate boron unit in so-called "monoclinic β-HBO_2" (Zachariasen, 1963). In that case the three short B–O distances (1.43, 1.45, and 1.45 Å) are identical with the two B–O1 and B–O2 distances in our trypsin–peptidylboronate complexes; the long B–OH_2 distance (~1.55 Å) in β-HBO_2 is similar to the B–Oγ195 distance in the enzyme complexes, and in both systems the distortion from tetrahedral corresponds to shift of the boron along the direction of the long B–OH_2 (or B–Oγ195) bond direction. Trigonal boron complexes provide an unfilled p orbital perpendicular to the trigonal plane, and the bond from a ligand on one side can be dative, formed by donating a lone pair of electrons, as from oxygen to the empty p orbital.

Interactions and Specificity at P3, P2, and P1. The optimal side chains at the P2 position of substrate or inhibitors that bind to trypsin, thrombin, or factor Xa are Pro and Gly. Indeed Val (as used here) is not a good substituent at P2 with regard to inhibition of trypsin, at least as seen in modifications of antithrombin–trypsin (Sheffield & Blajchman, 1994), suggesting that more efficacious inhibitors can be generated using different P2 substituents. The side chain at the P3 site (Ala in our case) interacts with the side chain of Gln192 in the trypsin-like enzymes; thus neutral polar side chains might be preferred to Ala at P3 for inhibiors of trypsin. This correlation has been seen for example through mutational analysis of Glu192 to Gln in thrombin, which results in altered specificity for substrate

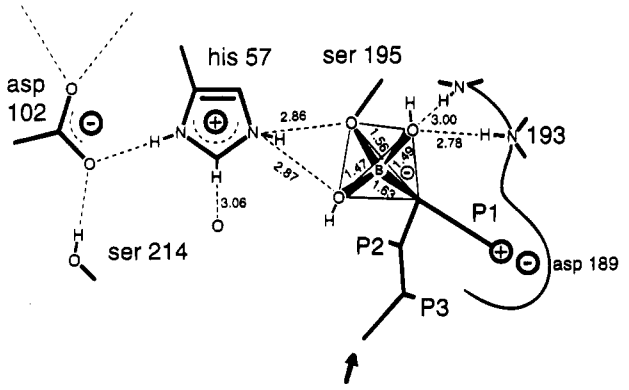


FIGURE 8: Schematic showing the bond lengths and hydrogen bond lengths obtained in the boronate−trypsin complex.

variants at P3 (Rezaie & Esmon, 1993). Overall, therefore, other sequences of the tripeptide should offer higher affinity for trypsin, and a systematic approach to optimization of the side chains presented at P3, P2, and even P1 should give rise to the diversity necessary to tune selectivity toward a particular serine protease. This is especially likely at P3, which may be useful in identifying inhibitors specifically oriented toward particular trypsin-like serine proteases and away from others.

Conformational Change upon Binding. The use of several independently determined crystal structures to construct a consensus structure for comparison with that of the enzyme−inhibitor complexes enabled detection of structural shifts as small as ~0.15 Å and dihedral angle perturbations as subtle as a few degrees. Estimates of expected error in specific coordinate positions, as a function of thermal factors, are calculated for each atom position according to $\sigma_{err}(B) = (3/4)Rr(0.00107B^2 - 0.01493B + 0.264)$ [eq 1 of Finer Moore et al. (1992) and see Perry et al. (1990)]. Overall, these average ~0.05 Å, and individual error estimates can be compared to individual shifts listed in Tables 6−8. Calculation of average main-chain shifts determined from vectors using a moving window provided a sensitive measure of concerted structural rearrangements, and calculation of average individual residue displacement vectors gave complementary measurements of structural shifts. The effective methodology for comparison and visualization (Katz & Kossiakoff, 1990) provided a detailed description of structural perturbations induced by the binding of peptidylboronate transition-state analogs to trypsin.

The long-range propagation of structural changes observed in trypsin−tripeptide boronate methyl ester shows that trypsin exhibits a large degree of structural plasticity, of the kind that must be taken account of in structure-based drug design. Accommodation of the peptidylboronate inhibitors in the trypsin structure is accomplished by conformational changes involving a combination of concerted structural shifts and perturbations of main-chain and side-chain dihedral angles at the P1 and P2 pockets. A surprising finding is the observation of propagation of perturbations to segments not only adjacent to the inhibitor binding sites but also adjacent to more distant regions. In two regions, perturbations introduced to a strand of β-sheet were transmitted to the associated strand. Perturbations at the P1 binding residues were both transmitted and amplified to an adjacent loop. A large fraction of the protein was perturbed from a conformational change more global than anticipated.

Crystallographic assessment of the nature, degree, and extent of conformational changes incurred in and around the specificity subsites of an enzyme by inhibitor binding yields information useful for further design of inhibitors more specific to the target enzyme. For example, observation of the dihedral strain in the P2 Val residue of the tripeptide boronate methyl ester and in the contacted Leu99 protein residue, along with the large shift in protein residues 96−98, suggests a possible redesign of the inhibitor at the P2 side chain into one with a smaller, more compatible, and not necessarily biological amino acid side chain.

It has long been recognized that there are well-defined changes induced in trypsin by the binding of substrate analogs or inhibitors, such as in the transition-state analog monoisopropylphosphoryl- (MIP-) inhibited trypsin, versus benzamidine-inhibited trypsin which occupies the P1 site (Krieger et al., 1974). But the extent of these well-defined changes in response to binding the boronates, using highly refined structures where small changes can be reliably described, remains remarkable, extending to ~50, out of the total 224 amino acid residues. Estimating the energy cost of inducing such changes is crude at best; nevertheless, when the changes in energies associated with changes in side-chain dihedral angles and main-chain φ and ψ angles are estimated using the force field of X-PLOR, the aggregate sum of changes corresponds to $\sim +10$ kcal/mol in enthalpy, as tabulated in Tables 6 and 7. This enthalpic term, an unfavorable cost of binding, counteracts the free energy gained through the favorable interactions. The balance of other entropic terms associated with fixation and desolvation also contributes to give the resultant binding free energy observed, $\Delta G°_f = -11.4$ kcal/mol, corresponding to the observed $K_i \sim 7$ nM. Thus if it were possible to accurately sum enthalpic terms due to interaction energies in an observed complex, the intrinsic binding energy calculated for the complex described here might correspond to ~−21.4 kcal/mol, irrespective of and before accounting for other electrostatic and entropic contributions. Thus although ~10 kcal/mol is necessarily a crude estimate of the torsional enthalpic cost paid to realize enthalpic gain upon binding, it is a reminder of the critical need to better evaluate such energy costs in a program of inhibitor or drug design. Having accurate structures of the free protein and bound complex is clearly essential to these developments.

ACKNOWLEDGMENT

We thank Michael Ross, Tom Jenkins, David Agard, and Tony Kossiakoff for discussion and insights and Michael Venuti and Heinz Gschwend for reviewing the manuscript.

REFERENCES

Alcock, N. W., Hagger, R. M., Harrison, W. D., & Wallbridge, M. G. H. (1982) *Acta Crystallogr. B38*, 676−677.

Al-Juaid, S. S., Eaborn, C., El-Kheli, M. N. A., Hitchcock, P. B., Lickiss, P. D., Molla, M. E., Smith, J. D., & Zora, J. A. (1989) *J. Chem. Soc., Dalton Trans.*, 447−452.

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

Bachovchin, W. W., Plaut, A. G., Flentke, G. R., Lynch, M., & Kettner, C. A. (1990) *J. Biol. Chem. 265*, 3738−3743.

Bajzer, Z., & Prendergast, F. G. (1992) *Methods Enzymol. 210*, 200−237.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Blow, D. M., Birktoft, J. J., & Hartley, B. S. (1969) *Nature (London) 221*, 337−340.
Bode, W., & Schwager, P. (1975) *J. Mol. Biol. 98*, 693−717.
Bolognesi, M., Gatti, G., Menegatti, E., Guarneri, M., Marquart, M., Papamokos, E., & Huber, R. (1982) *J. Mol. Biol. 162*, 839−868.
Bone, R., Frank, D., Kettner, C. A., & Agard, D. A. (1989) *Biochemistry 28*, 7600−7609.
Bone, R., Fujishige, A., Kettner, C. A., & Agard, D. A. (1991b) *Biochemistry 30*, 10388−10398.
Bone, R., Sampson, N. S., Bartlett, P. A., & Agard, D. A (1991a) *Biochemistry 30*, 2263−2272.
Brunger, A. T. (1990) *X-PLOR Manual, Version 2.1*, Yale University, New Haven, CT.
Cha, S. (1975) *Biochem. Pharmacol. 24*, 2177−2185.
Chambers, J. L., & Stroud, R. M. (1977) *Acta Crystallogr. B33*, 1824-1837.
Chambers, J. L., & Stroud, R. M. (1979) *Acta Crystallogr. B33*, 1861−1874.
Chen, Z., & Bode, W. (1983) *J. Mol. Biol. 164*, 283−311.
Collen, D., & Lijnen, H. R. (1991) in *Hematology Basic Principles & Practice* (Hoffman, R., Benz, E. J., Jr., Shattil, S. J., Furie, B., & Cohen, H. J., Eds.) pp 1232−1242, Churchill Livingstone, New York.
Davie E. W., Fujikawa K., & Kisiel W. (1991) *Biochemistry 30*, 10363−10370.
Delbaere, L. T. J., & Brayer, G. D. (1985) *J. Mol. Biol. 183*, 89−103.
Finer-Moore, J. S., Kossiakoff, A. A., Hurley, J. H., Earnest, T. N. E., & Stroud, R. M. (1992) *Proteins 3*, 203−222.
Flentke, G. R., Munoz, E., Huber, B. T., Plaut, A. G., Kettner, C. A., & Bachovchin, W. W. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1556−1559.
Franconi, G. M., Graf, P. D., Lazarus, S. C., Nadel, J. A., & Caughey, G. H. (1989) *J. Pharmacol. Exp. Ther. 248*, 947−951.
Frank, M. M., & Fries, L. F. (1989) in *Fundamental Immunology* (Paul, W. E., Ed.) pp 679−701, Raven Press Ltd., New York.
Franks, A. (1955) *Proc. Phys. Soc., London B62*, 1054.
Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M., Jr., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 397−418.
Greenblatt, H. M., Ryan, C. A., & James, M. N. G. (1989) *J. Mol. Biol. 205*, 201−228.
Gupta, A., Kirfel, A., & Will, G. (1977) *Acta Crystallogr. B33*, 637−641.
Hance, A. J., & Crystal, R. G. (1975) *Am. Rev. Respir. Dis. 112*, 657−711.
Hartley, B. S., & Kauffman, D. L. (1966) *Biochem. J. 101*, 229−231.
Hedstrom, L., Szilagyi, L., & Rutter, W. J. (1992) *Science 255*, 1249−1253.
Hedstrom, L., Farr-jones, S., Kettner, C. A., & Rutter, W. J. (1994a) *Biochemistry 33*, 8764−8769.
Hedstrom, L., Perona, J. J., & Rutter, W. J. (1994b) *Biochemistry 33*, 8757−63.
James, M. N. G., Brayer, G. D., Delbaere, L. T. J., Sielecki, A. R., & Gertler, A. (1980) *J. Mol. Biol. 139*, 423−438.
Jones, T. A. (1982) in *Computational Crystallography* (Sayre, D., Ed.) pp 303−317, Oxford University Press, Oxford, England.
Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916−924.
Katz, B. A., & Kossiakoff, A. A. (1990) *Proteins: Struct., Funct., Genet. 7*, 343−357.
Kettner, C., & Shenvi, A. B. (1984) *J. Biol. Chem. 259*, 15106−15114.
Kettner, C., & Knabb, R. M. (1993) *Adv. Exp. Med. Biol. 340*, 109−118.
Kettner, C., Bone, R., Agard, D. A., & Bachovchin, W. (1988) *Biochemistry 27*, 7682−7688.
Kettner, C., Mersinger, L., & Knabb, R. (1990) *J. Biol. Chem. 265*, 18289−18297.
Kossiakoff, A. A. (1984) *Basic Life Sci. 27*, 281−304.
Kossiakoff, A. A., & Spencer, S. A. (1981) *Biochemistry 20*, 6462−6474.
Kossiakoff, A. A., Chambers, J. L., Kay, L. M., & Stroud, R. M. (1976) *Biochemistry 15*, 654−664.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331−358.
Krieger, M., Kay, L. M., & Stroud, R. M. (1974) *J. Mol. Biol. 83*, 209−230.
Mangel, W. F., Singer, P. T., Cyr, D. M., Umland, T. C., Toledo, D. L., Stroud, R. M., Pflugrath, J. W., & Sweet, R. M. (1990) *Biochemistry 29*, 8351−8357.
Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math. 11*, 431−441.
Marquart, M., Walter, J., Deisenhofer, J., Bode, W., & Huber, R. (1983) *Acta Crystallogr. B39*, 480−490.
Matteson, D. S., & Majumber, D. (1983) *Organometallics 2*, 1529−1535.
Pauling, L. (1946) *Chem. Eng. News 24*, 1375.
Perona, J. J., Hedstrom, L., Wagner, R. L., Rutter, W. J., Craik, C. S., & Fletterick, R. J. (1994) *Biochemistry 33*, 3252−3259.
Perry, K. M., Fauman, E. B., Finer-Moore, J. S., Montfort, W. R., Maley, G. F., Maley, F., & Stroud, R. M. (1990) *Proteins 8*, 315−333.
Ponder, J. W., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775−791.
Powers, J. C. (1986) *Adv. Inflam. Res. 11*, 145−157.
Read, R. J., Fujinaga, M., Sielecki, A. R., & James, M. N. G. (1983) *Biochemistry 22*, 4420−4433.
Rettig, S. J., & Trotter, J. (1977) *Can. J. Chem. 55*, 3071−3075.
Rezaie, A. R., & Esmon, C. T. (1993) *J. Biol. Chem. 268*, 19943−19948.
Salemme, F. R., & Weatherford, D. W. (1981) *J. Mol. Biol. 146*, 119−141.
Schecter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
Sekizawa, K., Caughey, G. H., Lazarus, S. C., Gold, W. M., & Nadel, J. A. (1989) *J. Clin. Invest. 83*, 175−179.
Sheffield, W. P., & Blajchman, M. A. (1994) *FEBS Lett. 339*, 147−150.
Sommerhoff, C. P., Caughey, G. H., Finkbeiner, W. E., Lazarus, S. C., Basbaum, C. B., & Nadel, J. A. (1989) *J. Immunol. 142*, 2450−2456.
Stroud, R. M. (1974) *Sci. Am. 231*, 74−88.
Stroud, R. M., Kay, L. M., & Dickerson, R. E. (1971) *Cold Spring Harbor Symp. Quant. Biol. 36*, 125−140.
Stroud, R. M., Kay, L. M., & Dickerson, R. E. (1974) *J. Mol. Biol. 83*, 185−208.
Stroud, R. M., Kossiakoff, A. A., & Chambers, J. L. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 177−193.
Tam, E. K., & Caughey, G. H. (1990) *J. Respir. Cell Mol. Biol. 3*, 27−32.
Tsilikounas, E., Kettner, C. A., & Bachovchin, W. W. (1992) *Biochemistry 31*, 12839−12846.
Vartio, T., Seppa, H., & Vaheri, A. (1981) *J. Biol. Chem. 256*, 471−477.
Walter, J., Steigemann, W., Singh, T. P., Bartunik, H., Bode, W., & Huber, R. (1982) *Acta Crystallogr. B38*, 1462−1472.
Weber, I. T. (1987) *J. Appl. Crystallogr. 20*, 388−393.
Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S. J. R., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765−784.
Zachariasen, W. H. (1963) *Acta Crystallogr. 16*, 385−389.
Zhou, G. W., Guo, J., Huang, W., Fletterick, R. J., & Scanlan, T. S. (1994) *Science 265*, 1059−1064.

BI9501186