Inhibition of Trypsin and Thrombin by Amino(4-amidinophenyl)methanephosphonate Diphenyl Ester Derivatives: X-ray Structures and Molecular Models[†,‡]

Jay A. Bertrand, Josef Oleksyszyn, Chih-Min Kam, Bogdan Boduszek, Steven Presnell, R. Richard Plaskon, F. L. Suddath,[§] James C. Powers, and Loren Dean Williams*

School of Chemistry & Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400

Received September 5, 1995; Revised Manuscript Received December 20, 1995[⊗]

ABSTRACT: X-ray structures of trypsin from bovine pancreas inactivated by diphenyl [*N*-(benzyloxycarbonyl)amino](4-amidinophenyl)methanephosphonate [Z-(4-AmPhGly)^{P}(OPh)$_2$] were determined at 113 and 293 K to 1.8 Å resolution and refined to *R* factors of 0.211 (113 K) and 0.178 (293 K). The structures reveal a tetrahedral phosphorus covalently bonded to the Oγ of the active site serine. Covalent bond formation is accompanied by the loss of both phenoxy groups. The D-stereoisomer of Z-(4-AmPhGly)^{P}-(OPh)$_2$ is not observed in the complex. The L-stereoisomer of the inhibitor forms contacts with several residues in the trypsin active site. One of the phosphonate oxygens is inserted into the oxyanion hole and forms hydrogen bonds to the amides of Gly193, Asp194, and Ser195. The second phosphonate oxygen forms hydrogen bonds to Nϵ2 of His 57. The *p*-amidinophenylglycine moiety binds into the trypsin primary specificity pocket, interacting with Asp189. The amide forms a hydrogen bond to the carbonyl oxygen atom of Ser214. The inhibitor moiety, from the 113 K structure of trypsin inactivated by the reaction product of Z-(4-AmPhGly)^{P}(OPh)$_2$, was docked into human thrombin [Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. R., & Hofsteenge, J. (1989) *EMBO J. 8*, 3467−3475] and energy minimized. The inhibitor fits well into the thrombin active site, forming favorable contacts similar to those in the trypsin complex with no bad contacts.

Trypsin-like enzymes, a group of serine proteases, are causative or adjunctive in many disease states. The hyperproteolytic activities of this homologous family of enzymes are attractive chemotherapeutic targets in pathways of blood coagulation, fibrinolysis, kinin formation, complement activation, digestion, reproduction, and phagocytosis (Neurath, 1984). Tryptase, a pathogenic mediator of asthma, increases the contractility of airway smooth muscle (Sekizawa et al., 1989). Plasma kallikrein plays an important role in fibrinolysis, kinin formation, and blood coagulation (Lorand & Konishi, 1964). Thrombin is pivotal in blood coagulation, converting fibrinogen to fibrin. During thrombosis and hemostasis, thrombin activates factors V, VIII, and XIII and protein C. Thrombin cleavage of the thrombin receptor induces platelet aggregation (Vu et al., 1991).

Inhibitors of specific serine proteases could, in principle, remedy many disease states. The arginine-based inhibitor MQPA[1] (Kikumoto et al., 1984) and the amidinophenylalanyl-based inhibitor NAPAP (Stürzbecher et al., 1983) both reversibly inhibit trypsin and thrombin at submicromolar concentrations. The bis(benzamidino) compound 2,7-bis-(4-amidinophenyl)methenylcycloheptanone is a potent inhibitor of trypsin (K_i = 90 nM; Stürzebecher et al., 1992), and α,α′-bis(4-amidino-2-iodophenoxy)-*m*-xylene is a potent inhibitor of plasma kallikrein (K_i = 31 nM; Geratz et al., 1976).

Several classes of transition state analogues, including aldehydes (Umezawa, 1976; Umezawa & Aoyagi, 1977) and boronic acids (Kettner et al., 1990; Tapparelli et al., 1993a), have been designed to bind to serine proteases. Specificity for trypsin and thrombin is conferred by conjugation of appropriate transition state analogues with peptides such as D-Phe-Pro-Arg (Bajusz et al., 1990; Kettner et al., 1990; Kettner & Shaw, 1981; Shuman et al., 1993; Tapparelli et al., 1993b).

Another class of transition state inhibitors of serine proteases is composed of electrophilic phosphorus moieties

[†] This work was supported in part by a grant from the National Institutes of Health (HL29307 to J.C.P.).

[‡] Atomic coordinates have been deposited with the Brookhaven Protein Data Bank [ID codes 1MAX (LT) and 1MAY (RT)].

* Address correspondence to this author.

[§] Deceased.

[⊗] Abstract published in *Advance ACS Abstracts,* February 1, 1996.

[1] Abbreviations: ALP, α-lytic protease; Z, benzyloxycarbonyl; (4-AmPhGly)^{P}(OPh)$_2$, diphenyl amino(4-amidinophenyl)methanephosphonate [or diphenyl (4-amidinophenylglycyl)phosphonate]; (4-AmPhe)^{P}-(OPh)$_2$, diphenyl 1-amino-2-(4-amidinophenyl)ethanephosphonate [or diphenyl (4-amidinophenylalanyl)phosphonate]; APPA, (4-amidinophenyl)pyruvic acid; Boc, *tert*-butyloxycarbonyl; Boc-Ala-Ala-Pro-Val^{P}-(OPh)-Lac-Ala-OMe, *N*-[(2*S*)-2-[[[(1*R*)-1-[*N*-[(*tert*-butyloxycarbonyl)-L-alanyl-L-alanyl-L-prolyl]amino]-2-methylpropyl]phenoxyphosphinyl]oxy]propanoyl]-L-alanine methyl ester; MQPA, (2*R*,4*R*)-4-methyl-1-[N^{α}-[(3-methyl-1,2,3,4-tetrahydro-8-quinolinyl)sulfonyl]-L-arginyl]-2-piperidinecarboxylic acid; NAPAP, N^{α}-[(2-naphthylsulfonyl)-glycyl]-*p*-amidinophenylalanylpiperidine; 2-NpSO_2-Gly, (2-naphthylsulfonyl)glycyl; PPACK, D-Phe-Pro-Arg chloromethyl ketone; DMSO, dimethyl sulfoxide; DPP-IV, dipeptidyl peptidase IV; RT, room-temperature; LT, low-temperature; MES, 2-(*N*-morpholino)ethanesulfonic acid; SBzl, thiobenzyl ester.

[2] The nomenclature of Schechter and Berger (1967) is used to designate the amino acid residues (P_2, P_1, P_1', P_2', etc.) of a peptide substrate and the corresponding subsites (S_2, S_1, S_1', S_2', etc.) of the enzyme. The scissile peptide bond is P_1-P_1'.

[3] The (α-aminoalkyl)phosphonic acids are analogues of natural α-amino acids and are designated by the generally accepted three-letter abbreviations for the amino acid followed by a superscript P. For example, diphenyl [α-[*N*-(benzyloxycarbonyl)amino]ethyl]phosphonate, which is related to alanine, is abbreviated as Z-Ala^{P}(OPh)$_2$.

0006-2960/96/0435-3147$12.00/0

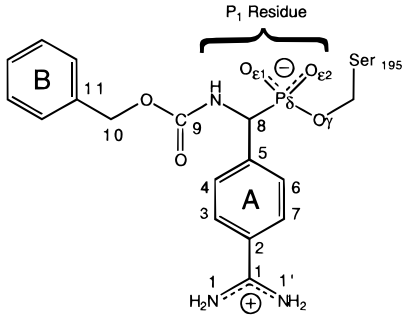

FIGURE 1: Covalent structure of trypsin inhibited by Z-(4-AmPhGlyP).

linked to peptides (Bartlett & Lambden, 1986; Lambden & Bartlett, 1983; Oleksyszyn et al., 1994; Oleksyszyn & Powers, 1991; Sampson & Bartlett, 1991). In particular, arginine and ornithine-linked phosphonates [(α-amino-δ-methoxybutyl)phosphonyl or (α-amino-*n*-hexyl)phosphonyl] are specific and irreversible inhibitors of trypsin and thrombin (Cheng et al., 1991; Fastrez et al., 1989; Hamilton et al., 1993; Wang et al., 1992). Phosphonates are chemically stable and synthetically accessible. Complexes of the phosphonate Boc-Ala-Ala-Pro-Val^{P}(OPh)-Lac-Ala-OMe bound to the serine protease ALP have been determined crystallographically (Bone et al., 1991). The structures show that the inhibitor forms a covalent bond with the active site Ser195, mimicking the high-energy tetrahedral intermediate.

Recently Oleksyszyn et al. (1994) described a series of phosphonate inhibitors that form stable complexes with trypsin-like serine proteases. These inhibitors are peptidyl-(α-aminoalkyl)phosphonate diphenyl esters containing 4-amidinophenyl groups. The most potent thrombin inhibitor in this series is Boc-D-Phe-Pro-(4-AmPhGly)^{P}(OPh)$_2$ with $k_{obsd}/[I] = 11\,000$ M^{-1} s^{-1} (thrombin) and $k_{obsd}/[I] = 2200$ M^{-1} s^{-1} (trypsin). Similar potency for trypsin has been achieved with Z-(4-AmPhGly)^{P}(OPh)$_2$, with $k_{obsd}/[I] = 2000$ M^{-1} s^{-1}. This phosphonate is also a potent inhibitor of plasma kallikrein ($k_{obsd}/[I] = 18\,000$ M^{-1} s^{-1}).

We describe room (RT)- and low-temperature (LT) X-ray structures of irreversible complexes formed by the reaction of trypsin with Z-(4-AmPhGly)^{P}(OPh)$_2$ (Figure 1). This structure is the first of a trypsin-like serine protease inhibited by a diphenyl ester phosphonate derivative. We show that only the L-stereoisomer binds and that the O-phenyl groups dissociate in an aging process. Using this inhibited trypsin structure, we have constructed a model of Z-(4-AmPhGlyP) covalently bound in the active site of human thrombin. It appears that structural differences in both the peptide moiety and the side chain group of the P$_1$ residue are related to the inhibitory potency toward trypsin and thrombin.

EXPERIMENTAL PROCEDURES

Diphenyl [*N*-(benzoxylcarbonyl)amino](4-amidinophenyl)-methanephosphonate, Z-(4-AmPhGly)^{P}(OPh)$_2$, was synthesized by the method of Oleksyszyn et al. (1994). Trypsin was purified by the method of Schroeder and Shaw (1968) as modified by Bode and Schwager (1975a). Fractions of β-trypsin were pooled, dialyzed against 5.0 mM MES pH 6.0 buffer, 1.0 mM CaCl$_2$, and 5.0 mM benzamidine, and concentrated by ultrafiltration. Orthorhombic crystals were grown by vapor diffusion from 20 μL hanging drops containing 10−15 mg/mL β-trypsin, 5.0 mM MES pH 6.0 buffer, 1.0 mM CaCl$_2$, 2.5 mM benzamidine, and 0.9 M ammonium sulfate equilibrated against a reservoir of 5.0 mM MES pH 6.0 buffer, 1.0 mM CaCl$_2$, 2.5 mM benzamidine, and 1.8 M ammonium sulfate.

To replace the trypsin-bound benzamidine with the reaction product of Z-(4-AmPhGly)^{P}(OPh)$_2$, crystals of approximate size 0.4 × 0.4 × 0.7 mm were dialyzed against 2.8 M ammonium sulfate, 2.0% DMSO, and 10 mM phosphate (pH 7.0). After 24 h, the dialysate was replaced with a solution of 0.1 mg/mL Z-(4-AmPhGly)^{P}(OPh)$_2$, 2.8 M ammonium sulfate, 2.0% DMSO, and 10 mM phosphate (pH 7.0). After 7 days of soaking, a crystal was mounted in a glass capillary and a room-temperature data set was collected. For the low-temperature data set, a different crystal was subjected to shock cooling by immersion of the capillary in liquid nitrogen before transfer to a stream of 113 K nitrogen gas bathing the goniostat. It appears that cryoprotection is provided by the relatively high ammonium sulfate concentration and the DMSO. The crystals were of space group $P2_12_12_1$ and were isomorphous to those of Bode and Schwager (1975a) with unit cell dimensions $a = 54.95$, $b = 58.55$, and $c = 67.70$ Å (RT) and $a = 54.66$, $b = 58.48$, and $c = 66.93$ Å (LT).

X-ray intensity data were collected at 293 K (RT) and 113 K (LT) with a San Diego Multiwire Systems (SDMS) area detector (Xuong et al., 1985) mounted with an Enraf-Nonius cryostat (Model FR558-S). X-rays were generated with a Rigaku RU-200 generator with a copper anode (λ = 1.54 Å) operated at 100 mA/46 kV, using a 0.7 mm collimator and a graphite monochromator. Data were collected by ω-scan.

The initial refinement model was benzamidine-inhibited trypsin (Bode & Schwager, 1975b) minus all associated ions, water molecules, and the inhibitor. Using XPLOR (Brunger et al., 1987), the model was subjected to a cycle of simulated annealing followed by positional and displacement factor refinement. A calcium ion was added to the conserved calcium binding site (Bode & Schwager, 1975b), and the model was subjected to additional cycles of positional and displacement factor refinement. $2F_o - F_c$ Fourier maps of the initial refined model show clear, clean, and continuous electron density for the protein with the exception of Ser146 to Ser147 in the autolysis loop (Schroeder & Shaw, 1968), the "soft" main chain residues Asn115-Ser116 (Bode & Schwager, 1975a), and a few side chain atoms. $F_o - F_c$ maps show difference electron density for approximately half of the atoms of the bound phosphonate inhibitor. The electron density of the initial refined model is continuous between Pδ of the inhibitor and Oγ of Ser195 (Figure 1), indicating formation of a covalent adduct.

With the program CHAIN (Sack, 1990), the inhibitor was docked piecewise, beginning with the 4-amidinophenyl moiety, into difference electron density. Restraints and initial coordinates for the inhibitor were generated with standard bond lengths and bond and dihedral angles. After the phosphate group was fit to correct geometry and electron density, the inhibitor was treated as a continuous chain bound to Oγ of Ser195. The positions of the phosphonate oxygen atoms (Figure 1; Oε1 and Oε2), C8, N8, and the benzyloxycarbonyl C9, O9, C10, and O10 atoms are well-defined in initial LT but not in RT electron density maps. Initial positions of these atoms in the RT structure were derived from the LT structure. All histidine residues were treated

Table 1: RT and LT Refinement Parameters for the Complexes of Trypsin with Z-(4-AmPhGlyP)

	RT	LT
resolution (Å)	1.80	1.80
R_{merge}	0.12	0.19
number of reflections (%)	96	91
unit cell constants (Å)		
a	54.95	54.66
b	58.55	58.48
c	67.70	66.93
final *R* factor	0.178	0.211
rms deviation of bond lengths from ideal (Å)	0.012	0.015
rms deviation of bond angles from ideal (deg)	2.68	3.15

as singly protonated, with His57 protonated at Nϵ and all others at Nδ. Solvent molecules, located from $2F_o - F_c$ and $F_o - F_c$ maps, were added to the model in groups of 12 or less per cycle of refinement. Although the majority of the inhibitor could be fit to electron density during the course of refinement, ring B could not, and the occupancy of this portion of the inhibitor was maintained at zero. The final *R* factor ($\sum||F_o| - |F_c||/\sum|F_o|$) is 0.178 for the RT structure and 0.211 for the LT structure for all data between 5.0 and 1.8 Å (Table 1). To obtain a relatively unbiased omit map, at the end of the refinement, the inhibitor and side chain atoms of Ser195 (Cβ and Oγ) were deleted. The refinement was continued for an additional 40 cycles, and an $F_o - F_c$ map was calculated.

Models of a Covalent Complex with Thrombin. Z-(4-amidinophenylglycine)phosphonate was docked onto the active site of human α-thrombin. A total of 193 α-carbon atoms of the LT trypsin structure were superimposed onto topologically equivalent atoms of PPACK-inhibited human α-thrombin (Bode et al., 1989) with PPACK removed [root mean square (rms) deviation of 0.81 Å]. The model was modified as expected for the reaction of Z-(4-AmPhGly)^P-$(OPh)_2$ with Ser195 Oγ (Figure 1) and was energy minimized using the program CHARMm (version 21.3) (Brooks et al., 1983; Momany & Rone, 1992) with movement allowed for the modified Ser195 and waters within 10 Å of the Ser195 Oγ. Hydrogen atoms were included in the minimizations. Charges for the modified Ser195 and Z-(4-AmPhGlyP) were obtained from quantum mechanical geometry optimizations using MOPAC 6.0 (Stewart, 1991). All other parameters were those of Kam et al. (1994). The rms deviation of positions of initial vs minimized modified Ser195 is 0.61 Å with the largest movement being ring B (average rms deviation with ring B removed was 0.36 Å).

Models of Noncovalent Complexes with Thrombin. Modeling of noncovalent complexes of Z-(4-AmPhGly)$^P(OPh)_2$ with human α-thrombin was performed to determine structures of reaction intermediates, which are not amenable to X-ray crystallographic analysis. Modeling was conducted in a fashion similar to that performed on isocoumarin−thrombin complexes (Kam et al., 1994). In the initial model, the phosphonyl group was oriented with a Oϵ2−Pδ−Ser195 Oγ angle of 100°. The distance from the phosphorus atom to Ser195 Oγ was 2.2 Å. The distance from Oϵ2 to the amide hydrogen of Ser195 was 1.6 Å and to the amide hydrogen of Gly193 was 1.8 Å. The smallest distances between amidino hydrogen and Asp189 carbonyl oxygen were 2.0 and 3.0 Å. The initial model was translated to 64 different positions within the thrombin active site. The resulting structures were energy minimized using CHARMm in a manner similar to that described previously (Plaskon et al., 1993). Complexes with minimized energies less than or equal to −1250 kcal/mol and P to Ser195 Oγ distances of less than or equal to 3.5 Å were selected as potentially productive noncovalent complexes.

RESULTS

In maps synthesized from trypsin data collected at low temperatures, most of the inhibitor [Z-(4-AmPhGlyP)] is well-defined by the electron density. It appears that at low temperatures much of the inhibitor is well-ordered. In maps synthesized from data collected at room temperature, the electron density of most of the inhibitor is diffuse, indicating significant thermal disorder. Figure 2 shows a LT $F_o - F_c$ map of the active site regions of trypsin inhibited by Z-(4-AmPhGlyP). At both temperatures, the electron density connecting Ser195 Oγ to Pδ of the inhibitor is clear and continuous, indicating formation of a covalent bond between the enzyme and the inhibitor. The LT maps indicate unambiguously that the phosphorus center is tetrahedral, consistent with results of ^{31}P NMR experiments (Oleksyszyn et al., 1994) and with the structure of Boc-Ala-Ala-Pro-Val^P-(OPh)-Lac-Ala-OMe bound to ALP (Bone et al., 1991). The atoms bonded to C8 (Pδ, N8, and C5) are well-defined at LT, and their positions indicate preferential binding of the L-stereoisomer of Z-(4-AmPhGlyP). The 4-amidinophenyl group is well-defined at both temperatures. Ring B is not observed in the electron density at either temperature and

FIGURE 2: Stereodrawing of 1.4 σ level, LT omit $F_o - F_c$ map of trypsin inhibited by Z-(4-AmPhGlyP). The coordinates of Ser195 and the inhibitor moiety from the fully refined trypsin phosphonate are overlaid.

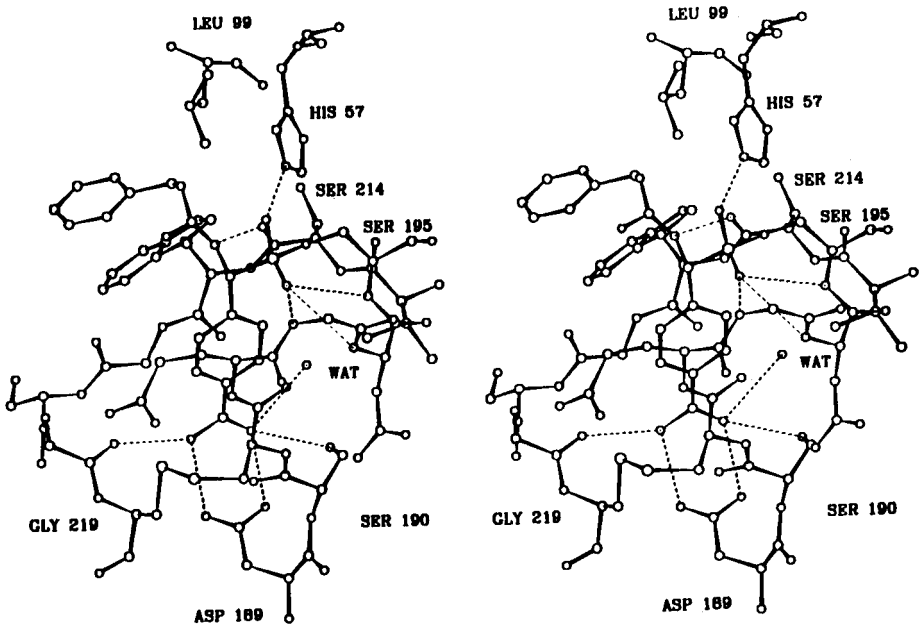

FIGURE 3: Stereodrawing of the active site of trypsin inhibited by Z-(4-AmPhGlyP) (LT structure). Hydrogen bonds are shown by dashed lines. WAT indicates the characteristic "buried" water molecule.

appears to be statically disordered at LT. A smaller degree of disorder, extending to other regions of the inhibitor, is suggested by a break in LT electron density between C8 and C5 (Figure 2). The phenoxy groups are not observable in the electron density maps at either temperature as expected if the initial reaction product, a serine phosphono diester, is ultimately hydrolyzed to a serine phosphono monoester (Oleksyszyn et al., 1994).

A lobe of electron density is observed adjacent to atom C6 of ring A (Figure 2). This lobe of electron density is not noise. Nearly identical features are observed in maps of a series of trypsin complexes with amidinophenyl-based inhibitors (J. A. Bertrand and L. D. Williams, unpublished). It is conceivable that these peaks of electron density result from occupancy of the S_1 pocket by solvent molecules in some fraction of the trypsin molecules that are not occupied by the inhibitor.

Enzyme/Inhibitor Interactions. Many interactions stabilizing the complex of trypsin with Z-(4-AmPhGlyP) are similar to those observed previously in inhibited trypsin complexes. The amidinophenyl group of Z-(4-AmPhGlyP) is inserted between main chain segments Ser190-Gln192 and Trp215-Gly216. The position and contacts of the amidinophenyl group are similar to other amidino- and arginyl-based trypsin/thrombin inhibitors (Bode & Schwager, 1975a; Bode et al., 1990; Lee et al., 1993; Sweet et al., 1974; Walter & Bode, 1983). The amidino nitrogen atoms form hydrogen bonds with the carboxylate oxygens of Asp189 and with Gly219 O, Ser190 Oγ, and "buried" water molecule 557 (Figure 3 and Table 2).

The phosphonate moiety of the inhibitor forms four hydrogen bonds with the protein (Figure 3 and Table 2). Oϵ2 is located within the oxyanion hole, forming hydrogen bonds to the amides of Gly193, Asp194, and Ser195. Phosphonate Oϵ1 forms a hydrogen bond with Nϵ2 of His57. The inhibitor amide (N8) forms a hydrogen bond with the carbonyl oxygen of Ser214.

The benzyloxycarbonyl moiety does not form hydrogen bonds with the protein. The carbonyl group is directed (C9 to O9) into the solvent channel with an orientation and a position similar to those of a normal P_2 carbonyl group. Similarly, O10 is in a position analogous to that of a P_2 α-carbon. Leu99 Cδ1 is in van der Waals contact with C10 (3.7 Å). The position of ring B is uncertain, not defined by electron density, and probably variable. This ring appears to be located in the solvent region, out of contact with the protein surface. A range of positions for ring B can be estimated from the known position of C10 and stereochemical considerations.

Table 2: Hydrogen Bond Distances

inhibitor atom	protease/cofactor atom	distance (Å)	
		phosphonate[a]	Bone et al.[b]
195 N1	Asp189 Oδ1	3.0	
	Gly219 O	2.8	
195 N1′	Asp189 Oδ1	3.2	
	Asp189 Oδ2	2.9	
	Ser190 Oγ	3.0	
	Wat557 OH_2	3.3	
195 N8	Ser214 O	3.3	3.0
195 Oϵ1	His57 Nϵ2	2.8	(2.9)[b]
195 Oϵ2	Gly193 N	2.7	2.6
	Asp194 N	3.3	
	Ser195 N	3.0	2.8

[a] Complex of trypsin inhibited by Z-(4-AmPhGlyP), low-temperature structure. [b] Complex of α-lytic protease inhibited by Boc-Ala-Ala-Pro-Pval-Lac-Ala-OMe, where Pval is the phosphonic acid analogue of Val and Lac is lactate, from Bone et al. (1991). The binding site is occupied partially by the intact inhibitor and partially by a fragment of the inhibitor. In the cleaved form, the distance from His57 Nϵ2 to 195 Oϵ1 is 2.9 Å. In the intact form, His57 is displaced, increasing this distance to 5.5 Å. Other H bond distances are nearly identical in the two forms of the complex.

Model of Human Thrombin Inhibited by the Reaction Product of Z-(4-AmPhGly)^P(OPh)$_2$. Contacts of the amidinophenyl group of Z-(4-AmPhGlyP) in the covalent thrombin model are similar to those in the trypsin complex. An initial model of thrombin inhibited by the reaction product of Z-(4-AmPhGlyP) (Figure 4) was obtained by superimposition of the LT trypsin−inhibitor structure onto thrombin (Bode et

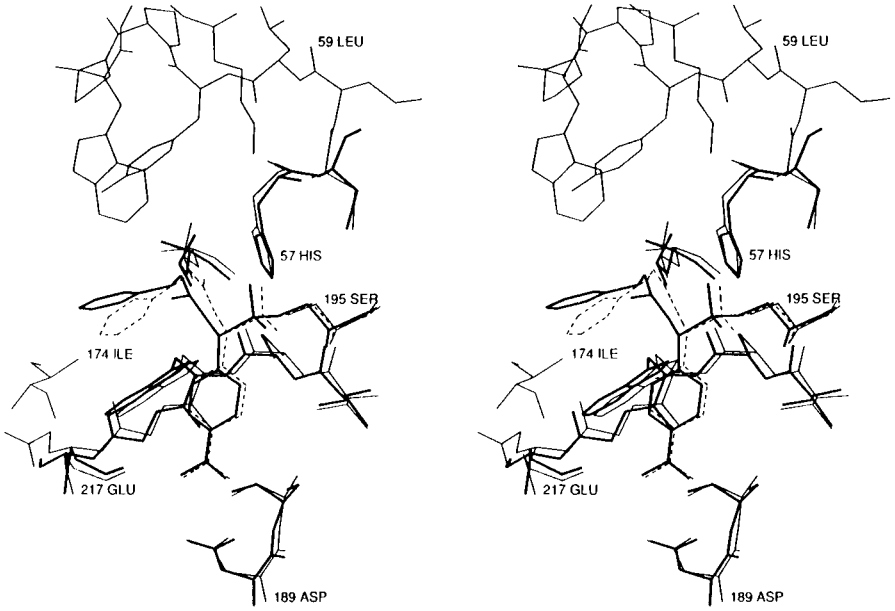

FIGURE 4: Stereodrawing of the active site of trypsin inhibited by Z-(4-AmPhGlyP) (thick lines) superimposed on human thrombin (thin lines; Bode et al., 1989). Z-(4-AmPhGlyP) is shown before (thick lines) and after (dashed lines) energy minimization.

al., 1989). Energy minimization performed of this thrombin model shows that it is energetically reasonable.

Subtle differences in hydrogen-bonding patterns distinguish the trypsin and thrombin complexes. Replacement of Ser190 in trypsin with Ala190 in thrombin distinguishes the amidinophenyl moiety binding regions This difference makes the thrombin pocket slightly less polar and proscribes the hydrogen bond between N1′ of the inhibitor and Oγ of Ser190 of trypsin (Bode et al., 1989). Another difference is an apparently weakened hydrogen bond of N8 with the carbonyl oxygen of Ser214 in thrombin [3.3 Å (LT, trypsin) to 3.4 Å (thrombin)], resulting from small differences in the positions of the carbonyl oxygen of Ser214 in trypsin and in thrombin.

The active site regions of trypsin and thrombin have different hydrophobic characters. The extent of hydrophobic interaction has been inferred here by a decrease in solvent accessible surface area upon binding. Figure 5 illustrates a method for comparing the extent of hydrophobic interactions of various complexes: (i) our X-ray structure of trypsin inhibited by Z-(4-AmPhGlyP), (ii) our model of thrombin inhibited by Z-(4-AmPhGlyP), and (iii) a previous X-ray structure of thrombin inhibited by PPACK (Bode et al., 1989). In a manner analogous to that of PPACK, Z-(4-AmPhGlyP) makes thrombin-specific interactions with hydrophobic groups of Tyr60A and Trp60D, part of a nine-residue cap (Leu59−Asn62) of the thrombin active site region. The benzyl group of the inhibitor makes thrombin-specific hydrophobic contacts with residue 174 [Ile174 (thrombin), Gly174 (trypsin)], part of an open, exposed turn (Bode et al., 1990).

Models of Noncovalent Thrombin Complexes with Z-(4-AmPhGly)^P(OPh)$_2$. Noncovalent structures of human α-thrombin complexed with Z-(4-AmPhGly)^P(OPh)$_2$ (Figure 6) are similar to the covalent structure derived from the trypsin X-ray structure. Modeling shows that the 4-amidinophenylglycine and phosphorus atoms are in nearly identical positions. The Z-phenyl groups are located in similar regions. There is ample room in the active site of the noncovalent complex for the phenoxy groups that have dissociated from the covalent complex. The angles formed by Oϵ_2-Pδ-Ser195 Oγ are 166−174° for eight of the ten productive structures and 155 and 156° for the other two.

Thermal Expansion. Increasing the temperature from LT to RT increases the volume of the unit cell by 1.83% (3919 Å^3). Unit cell parameters for crystals at LT and RT indicate expansion of the unit cell in all dimensions with increasing temperature, with the largest change in the *c* axis (Table 1). The increase in volume can be attributed to structural changes within individual protein molecules and the surrounding solvent and/or to changes in the orientation or packing of protein molecules in the unit cell (Tilton et al., 1992). The rms deviations between the LT and RT enzyme structures is 0.35 Å for α-carbon atoms and 0.61 Å for whole residues. The rms deviations for α-carbon atoms show small but noticeable effects of temperature on α-carbon atom positions, providing evidence for structural changes in the enzyme complex. Similar observations were made by Tilton and colleagues for ribonuclease A (Tilton et al., 1992).

Dynamics. The degree of local disorder can be inferred from isotropic displacement factors (*B*) indicating dispersion of electron density around average atomic position. The positional uncertainty of the atom is related to the displacement factor by the mean square vibration ($\langle x\rangle^2$) of an atom ($B = 8\pi\langle x\rangle^2$). Positional uncertainty is caused by multiple static conformations, dynamic conversion between conformers, model errors, lattice flaws, and errors in measuring structure factor amplitudes (Tilton et al., 1992). If other contributions to uncertainty are small, multiple static conformations can be distinguished from dynamic conversion between conformers by varying temperature. Static disorder is independent of temperature, but dynamic disorder will have a temperature dependence related to the shape of the potential well (Frauenfelder et al., 1979). Variation of displacement factors with temperature is predominantly related to the dynamic behavior of protein atoms (Petsko & Ringe, 1984). Figure 7 shows LT and RT backbone displacement factors averaged over each residue. Displacement factors generally increase with temperature. Uneven individual temperature dependence implies that the stabilizing forces that govern

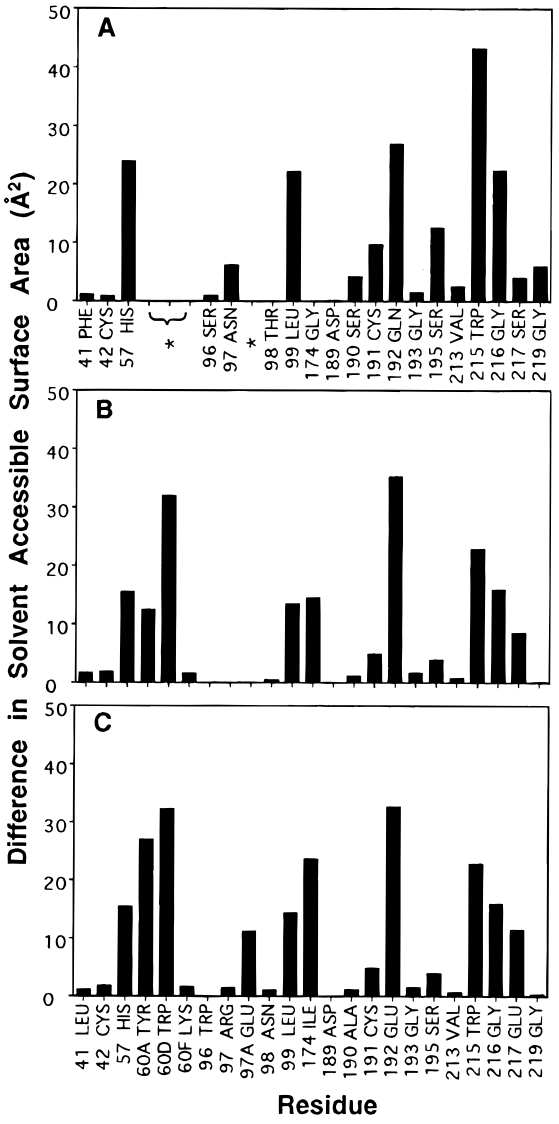

FIGURE 5: Absolute change in solvent accessible surface area upon binding to an uninhibited enzyme: (A) binding of Z-(4-AmPhGlyP) to trypsin (* indicates thrombin specific residues, not present in trypsin), (B) binding of Z-(4-AmPhGlyP) to thrombin, and (C) binding of PPACK to thrombin.

structure and dynamics are not uniform throughout the protein (Tilton et al., 1992).

DISCUSSION

We have determined room- and low-temperature X-ray structures of trypsin inhibited by the reaction product of the (4-amidinophenylglycine)phosphonate derivative, Z-(4-AmPhGly)^{P}(OPh)$_2$. This phosphonate is a member of a potent class of transition state inhibitors of trypsin-like proteases. These trypsin inhibitors are inactive toward acetylcholinesterase (Oleksyszyn et al., 1994). The second order rate constant ($k_{obsd}/[I]$) of Z-(4-AmPhGly)^{P}(OPh)$_2$ with trypsin is 2000 M^{-1} s^{-1}.

The reaction of trypsin with Z-(4-AmPhGly)^{P}(OPh)$_2$ to form a tetrahedral adduct is thought to proceed via an associative pathway. The transition state of this reaction contains a trigonal bipyramidal, pentacoordinated phosphorus atom with donor and acceptor ligands in apical positions (Oleksyszyn & Powers, 1994). Phosphonylation of the active serine results in immediate loss of one phenoxy group. The second phenoxy group appears to be hydrolyzed during aging of the phosphonylated derivative. The X-ray structure of trypsin and models of human α-thrombin inhibited by Z-(4-AmPhGlyP) are consistent with this proposed reaction mechanism. In noncovalent models, the Ser195 Oγ and phosphonyl O atoms are in apical positions in the phosphonylation transition state. The ground state tetrahedral phosphorus adduct is similar to the tetrahedral transition state of peptide bond hydrolysis with one phosphonate oxygen located within the oxyanion hole, forming hydrogen bonds to the amide NH atoms of Gly193, Asp194, and Ser195. The other phosphate oxygen forms hydrogen bonds to Nϵ2 of His57 (Figure 3 and Table 2). Experiments using ^{31}P NMR give results consistent with a tetrahedral ground state in serine protease inhibition by phosphonate diphenyl esters (Oleksyszyn et al., 1994; Oleksyszyn & Powers, 1991). Further, results for the inhibition of chymotrypsin by Suc-Val-Pro-Phe^{P}(OPh)$_2$ (Oleksyszyn & Powers, 1991) and by DPP-IV (Boduszek et al., 1994) suggest that one diastereomer reacts preferentially with trypsin. The present report establishes that the L-stereoisomer of the diphenyl ester phosphonate is the preferred reactant.

The carbonyl oxygen of Ser214 forms a hydrogen bond with the amide nitrogen (N8) of the inhibitor. This hydrogen bond is similar to that formed to the amide nitrogen of a normal peptide substrate. Such hydrogen bonds were observed by Bone et al. (1991) in the complexes formed between ALP and the stereoisomers of Boc-Ala-Ala-Pro-Val^{P}(OPh)-Lac-Ala-OMe (Table 2). Trypsin inhibited by Z-(4-AmPhGlyP) is most similar to the ALP−phosphonate complex in which the ester between the phosphonate and the lactate residues has been cleaved, leaving a phosphonic acid moiety [Boc-Ala-Ala-Pro-Val^{P}(OPh)] covalently linked to the active site serine.

Although there are a number of similarities in complexes of trypsin with Z-(4-AmPhGlyP) and ALP with Boc-Ala-Ala-Pro-Val^{P}(OPh)-Lac-Ala-OMe (Bone et al., 1991), there are also some notable differences. In the ALP complexes, the P_1 valine residue makes van der Waals contact with residues that form a shallow, hydrophobic S_1 pocket. The ALP S_1 pocket is lined by main chain residues 214−216 and 192−193. The base of the pocket is composed of the side chains of Met192, Met213, and Val217A. In trypsin, the S_1 pocket is lined by main chain residues 215−216 and 190−192 and the base of the pocket is formed by the carboxylate oxygens of Asp189. The amidino nitrogens of the P_1 4-amidinophenylglycine moiety form hydrogen bonds with the carboxylate oxygens of Asp189 (Figure 3 and Table 2). Consistent with other trypsin structures with amidino and arginyl inhibitors (Bode & Schwager, 1975a; Lee et al., 1993; Sweet et al., 1974; Walter & Bode, 1983), the amidino nitrogens also form hydrogen bonds with Gly219 O, Ser190 Oγ, and a buried water molecule. The Z-(4-AmPhGlyP) appears to fit ideally into the active site of trypsin (and thrombin). The covalent linkage to Ser195 does not induce strain or distortion in the hydrogen bonding interactions of P_1 with S_1.

The trypsin potencies of diphenyl (4-amidinophenylglycine)phosphonate derivatives are significantly decreased by insertion of a methylene group between C5 and C8. This modification converts (4-amidinophenylglycine)phosphonate to (4-amidinophenylalanine)phosphonate and decreases the $k_{obsd}/[I]$ from 100−2200 to 24−130 M^{-1} s^{-1} (Oleksyszyn et

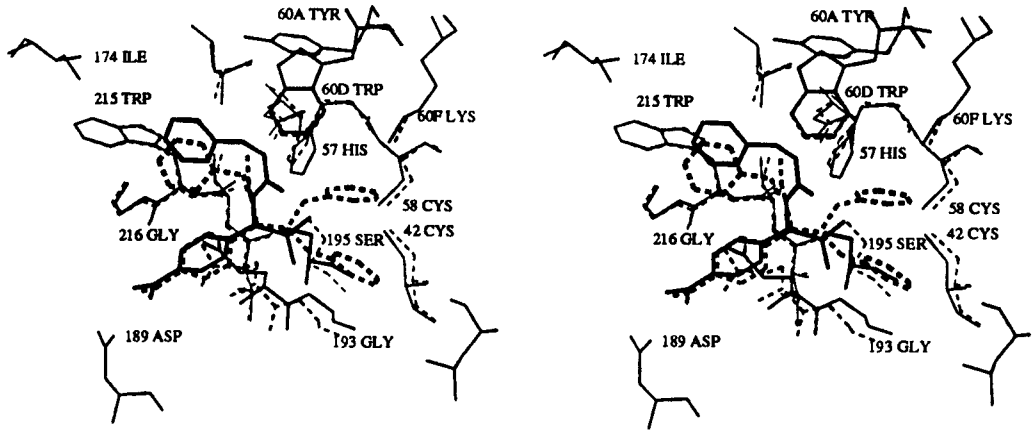

FIGURE 6: Stereodrawing of a model of the active site of human α-thrombin (dashed thin lines) noncovalently complexed with Z-(4-AmPhGly)^{P}(OPh)$_2$ (dashed thick lines) superimposed on the model of α-thrombin (solid thin lines) covalently inhibited by Z-(4-AmPhGlyP) (solid thick lines). The drawing was made with the program MOLSCRIPT (Kraulis, 1991).

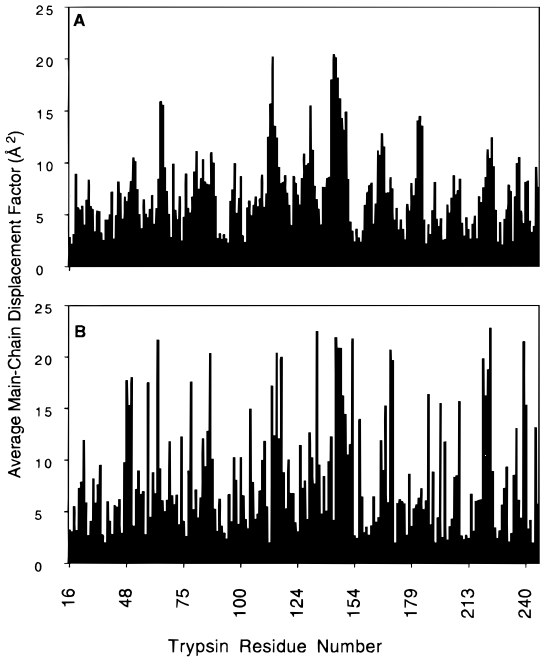

FIGURE 7: Average (A) RT and (B) LT displacement factors of the backbone atoms of each residue of trypsin.

al., 1994). Insertion of this methylene group into the three-dimensional structure of Z-(4-AmPhGlyP) with trypsin would increase the distance between the phosphorus atom and the amidinophenyl moiety. Formation of the covalent bond between Pδ and Oγ of Ser195 would then constrain hydrogen bonds in the P_1 pocket to suboptimal geometry.

Thrombin Inhibited by Z-(4-AmPhGlyP). The contacts in the model of Z-(4-AmPhGlyP) covalently bound to thrombin are generally similar to those observed in the complex with trypsin (Figure 4). A notable difference results from replacement of Ser190 in trypsin with Ala190 in thrombin. This switch makes the S_1 pocket of thrombin less polar than that of trypsin and results in the loss of the hydrogen bond between N1′ of the phosphate of Z-(4-AmPhGlyP) and Oγ of Ser190. The hydrophobic interactions of Z-(4-AmPhGlyP) with thrombin also differ from those with trypsin. Z-(4-AmPhGlyP) makes thrombin-specific interactions with hydrophobic residues Tyr60A, Trp60D, and Ile174 (Figure 5).

Comparison with Thrombin Inhibited by NAPAP or PPACK. In thrombin, an aryl binding site is formed by residues Ile174, Trp215, and Glu97A−Leu99 (Bode et al., 1990). Aryl groups generally bind there in a favorable edge-on arrangement with Trp215 as observed in complexes of NAPAP, MQPA, hirudin, and PPACK (Bode et al., 1989, 1990; Brandstetter et al., 1992; Rydel et al., 1990). Although the position of ring B is uncertain in our trypsin structure and thrombin models, the aryl binding site appears to be beyond its range of possible positions. Ring B is directed into the solvent region, leaving thrombin residues such as Ile174 exposed to solvent (Figure 5). It is possible that steric conflicts with Leu99 prevent occupation of the aryl binding site by ring B.

Like those of NAPAP and PPACK, the amidinophenyl ring of Z-(4-AmPhGlyP) binds within the S_1 pocket. However, the amidinophenyl ring of Z-(4-AmPhGlyP) is rotated by 24° relative to those of NAPAP and PPACK. The carbonyl group (C9 and O9) of Z-(4-AmPhGlyP) is nearby the piperidine amide carbonyl of NAPAP. The location of this carbonyl group may explain the surprisingly low thrombin potency of a related inhibitor, 2-NpSO_2-Gly-(4-AmPhGly)^{P}(OPh)$_2$ (k_{obsd}/[I] of 170 M^{-1} s^{-1}; Oleksyszyn et al., 1994). We initially expected greater inhibitory potency of this inhibitor since the structurally similar NAPAP inhibitor has a very low K_i value with bovine thrombin (6 nM, DL-racemate; Stürzbecher et al., 1983). However, if the carbonyl of the P_2 glycine of 2-NpSO_2-Gly-(4-amidinophenylglycine)phosphonate occupies a position similar to that of Z-(4-AmPhGlyP), the glycine residue would be directed up through the S_2 pocket, causing the naphthyl group to be improperly positioned for insertion into the aryl binding pocket.

The potency of D-Phe-Pro-(4-AmPhGly)^{P}(OPh)$_2$ is surprisingly low [k_{obsd}/[I] of 700 M^{-1} s^{-1} (human thrombin) and 730 M^{-1} s^{-1} (bovine thrombin; Oleksyszyn et al., 1994)] in comparison with the potency of PPACK (k_{obsd}/[I] of 9.60 × 10^6 M^{-1} s^{-1}; Lijnen et al., 1984). One factor that may relate to these low potencies is the positions of α-carbon atoms (C8) of the P_1 residues. The P_1 α-carbon of the phosphonate inhibitor is further from the imidazole group of His57 than the P_1 α-carbon atom of PPACK. This variation in P_1 α-carbon position results from differences in P_1 structure (4-amidinophenylglycine vs arginine), differences in bond torsion angles and lengths, and the presence/absence of a covalent bond with His57.

The position of the P_2 residue in other (4-amidinophenylglycine)phosphonate derivatives can be inferred from position of the rigidly bound amidinophenyl moiety and the covalent phosphonate linkage to Ser195. The only structural variations expected in di- or tripeptide (4-amidinophenylglycine)phosphonates would be in the substituent group. In comparison with the proline of PPACK, the proline of D-Phe-Pro-(4-AmPhGly)^P(OPh)$_2$ would be positioned further out into the solvent region. As a result, the hydrophobic contribution to the stability of thrombin complexes by the proline of D-Phe-Pro-(4-AmPhGly)^P(OPh)$_2$ would be less than that by the proline of PPACK. This shift in the P_2 proline position would also be expected to decrease hydrophobic interactions between the D-Phe phenyl group and the aryl binding site. Further, in the thrombin complex, PPACK forms critical hydrogen bonds with N and O of Gly216 (Bode et al., 1992). Similar hydrogen bonds between inhibitor atoms and Gly216 have been observed in bovine thrombin complexes with NAPAP and MQPA (Brandstetter et al., 1992). In the models of thrombin complexed with Z-(4-AmPhGlyP) and D-Phe-Pro-(4-AmPhGly)^P(OPh)$_2$, the inhibitors are not properly positioned to form hydrogen bonds with Gly216.

Inhibitor Dynamics. With data collected at room temperature, the initial electron density maps in the inhibitor binding region were diffuse, consistent with significant static/thermal disorder. Electron density in the active site region indicated the position of the amidinophenyl moiety and the covalent linkage formed between Oγ of Ser195 and the inhibitor. The positions of C8, N8, C9, O9, O10, C10, Oϵ1, and Oϵ2 were poorly defined. Electron density from LT data, absent thermal disorder, is defined much better. Ring B of Z-(4-AmPhGlyP) was not definable by electron density at either temperature, presumably as a result of static disorder. The inhibitor is anchored to the enzyme by the covalent linkage and by the amidino group interactions in the bottom of the specificity pocket. Both static and thermal disorder of the inhibitor increase with distance from these anchor points.

Conclusion. We have determined the three-dimensional structure of the irreversible complex of trypsin and the reaction product of Z-(4-AmPhGly)^P(OPh)$_2$ and show that only the L-stereoisomer is bound in the trypsin active site. In addition, we have docked Z-(4-AmPhGlyP) from the trypsin structure into the active site of human thrombin and have used the resulting model to relate structural differences in both the peptide moiety and the side chain group of the P_1 residue to the inhibition kinetic constants. In the design of inhibitors, a logical goal would be to increase the number of interactions between the enzyme and inhibitor. The number of hydrogen bonds in thrombin could be increased for the (4-AmPhGly)^P(OPh)$_2$ series inhibitors by utilization of the hydrogen bonding capabilities of the extended substrate binding site. The hydrophobic contribution to stability could be increased in thrombin complexes with (4-AmPhGly)^P-(OPh)$_2$ series inhibitors by better utilization of the hydrophobic aryl binding site.

ACKNOWLEDGMENT

The authors thank Dr. Wolfram Bode for helpful discussions and for providing seed crystals.

REFERENCES

Bajusz, S., Szell, E., Bagdy, D., Barabas, E., Horvath, G., Dioszegi, M., Fittler, Z., Szabo, G., Juhasz, A., Tomori, E., & Szilagyi, G. (1990) *J. Med. Chem. 33*, 1729−1735.

Bartlett, P. A., & Lambden, L. A. (1986) *Bioorg. Chem. 14*, 356−377.

Bode, W., & Schwager, P. (1975a) *J. Mol. Biol. 98*, 693−717.

Bode, W., & Schwager, P. (1975b) *FEBS Lett. 56*, 139−143.

Bode, W., Ulrich, I., Baumann, U., Huber, R., Stone, S. R., & Hofsteenge, J. (1989) *EMBO J. 8*, 3467−3475.

Bode, W., Turk, D., & Stürzebecher, J. (1990) *Eur. J. Biochem. 193*, 175−182.

Bode, W., Turk, D., & Karshikov, A. (1992) *Protein Sci. 1*, 426−471.

Boduszek, B., Oleksyszyn, J., Kam, C.-M., Selzler, J., Smith, R. E., & Powers, J. C. (1994) *J. Med. Chem. 37*, 3969−3976.

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

Brandstetter, H., Turk, S., Hoeffken, H. W., Grosse, D., Stürzebecher, J., Martin, P. D., Edwards, B. F. P., & Bode, W. (1992) *J. Mol. Biol. 226*, 1085−1099.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Brunger, A., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458−460.

Cheng, L., Goodwin, C. A., Scully, M. F., Kakkar, V. V., & Claeson, G. (1991) *Tetrahedron Lett. 32*, 7333−7336.

Fastrez, J., Jaspers, L., Lison, D., Renard, M., & Sonveaux, E. (1989) *Tetrahedron Lett. 30*, 6861−6864.

Frauenfelder, H., Petsko, G. A., & Tsernoglou, D. (1979) *Nature 280*, 558−563.

Geratz, J. D., Cheng, M. C.-F., & Tidiwell, R. R. (1976) *J. Med. Chem. 19*, 634−639.

Hamilton, R., Walker, B. J., & Walker, B. (1993) *Tetrahedron Lett. 34*, 2847−2850.

Kam, C.-M., Kerrigan, J. E., Plaskon, R. R., Duffy, E. J., Lollar, P., Suddath, F. L., & Powers, J. C. (1994) *J. Med. Chem. 37*, 1298−1306.

Kettner, C., & Shaw, E. (1981) *Methods Enzymol. 80*, 826−842.

Kettner, C., Mersinger, L., & Knabb, R. (1990) *J. Biol. Chem. 265*, 18289−18297.

Kikumoto, R., Tamao, Y., Tezuka, T., Tonomura, S., Hara, H., Ninomiya, K., Hijikata, A., & Okamoto, S. (1984) *Biochemistry 23*, 85−90.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Lambden, L. A., & Bartlett, P. A. (1983) *Biochem. Biophys. Res. Commun. 112*, 1085−1090.

Lee, A. Y., Hagihara, M., Karmacharya, R., Albers, M. W., Schreiber, S. L., & Clardy, J. (1993) *J. Am. Chem. Soc. 115*, 12619−12620.

Lijnen, H. R., Ugtterhoeven, M., & Collen, D. (1984) *Thromb. Res. 34*, 431−437.

Lorand, L., & Konishi, K. (1964) *Arch. Biochem. Biophys. 105*, 58−67.

Momany, F. A., & Rone, R. (1992) *J. Comput. Chem. 13*, 888−900.

Neurath, H. (1984) *Science 224*, 350−357.

Oleksyszyn, J., & Powers, J. C. (1991) *Biochemistry 30*, 485−493.

Oleksyszyn, J., & Powers, J. C. (1994) *Methods Enzymol. 244*, 423−441.

Oleksyszyn, J., Boduszek, B., Kam, C., & Powers, J. C. (1994) *J. Med. Chem. 37*, 226−231.

Petsko, G. A., & Ringe, D. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 331−372.

Plaskon, R. R., Kam, C.-M., Kerrigan, J. E., Burgess, E. M., Powers, J. C., & Suddath, F. L. (1993) *Arch. Biochem. Biophys. 300*, 588−597.

Rydel, T. J., Ravichandran, K. G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., & Fenton, J. W. I. (1990) *Science 249*, 277−280.

Sack, J. (1990) *Chain: A Crystallographic Modelling Program*, Baylor College of Medicine, Waco, TX.

Sampson, N. S., & Bartlett, P. A. (1991) *Biochemistry 30*, 2255−2263.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
Schroeder, D. D., & Shaw, E. (1968) *J. Biol. Chem. 243*, 2943−2949.
Sekizawa, K., Caughey, G. H., Lazarus, S. C., Gold, W. M., & Nadel, J. A. (1989) *J. Clin. Invest. 83*, 183−188.
Shuman, R. T., Rothenberger, R. B., Campbell, C. S., Smith, G. F., Gifford-Moore, D. S., & Gesellchen, P. D. (1993) *J. Med. Chem. 36*, 314−319.
Stewart, J. J. P. (1991) *QCPE Catalogue Volume XXIII*, Program number 455.
Stürzbecher, J., Markwardt, F., Voigt, B., Wagner, G., & Walsmann, P. (1983) *Thromb. Res. 29*, 635−642.
Stürzebecher, J., Prasa, D., & Sommerhoff, C. (1992) *Hoppe-Seyler's Z. Physiol. Chem. 373*, 1025−1030.
Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H., & Blow, D. M. (1974) *Biochemistry 13*, 4212−4228.
Tapparelli, C., Metternich, R., Ehrhardt, C., Zurini, M., Claeson, G., Scully, M., & Stone, S. (1993a) *J. Biol. Chem. 268*, 4734−4741.
Tilton, R. F., Dewan, J. C., & Petsko, G. A. (1992) *Biochemistry 31*, 2469−2481.
Umezawa, H. (1976) *Methods Enzymol. 45*, 678−695.
Umezawa, H., & Aoyagi, T. (1977) in *Proteinases in Mammalian Cells and Tissues* (Barrett, A., Ed.) pp 637−662, Elsevier Science Publishers, Amsterdam.
Vu, T.-K. H., Hung, D. T., Wheaton, V. I., & Coughlin, S. R. (1991) *Cell 64*, 1057−1068.
Walter, J., & Bode, W. (1983) *Hoppe-Seyler's Z. Physiol. Chem. 364*, 949−959.
Wang, C.-L., Taylor, T. L., Mical, A. J., Spitz, S., & Reilly, T. M. (1992) *Tetrahedron Lett. 33*, 7667−7670.
Xuong, N. H., Nielsen, C., Hamlin, R., & Anderson, D. (1985) *J. Appl. Crystallogr. 18*, 342−350.

BI9520996