The 2.2-Å Resolution X-ray Crystal Structure of the Complex of Trypsin Inhibited by 4-Chloro-3-ethoxy-7-guanidinoisocoumarin: A Proposed Model of the Thrombin–Inhibitor Complex†

Margaret M. Chow,‡ Edgar F. Meyer, Jr.,*,‡ Wolfram Bode,§ Chih-Min Kam,∥ R. Radhakrishnan,‡ J. Vijayalakshmi,‡ and James C. Powers∥

Contribution from the Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843-2128, Max-Planck-Institut für Biochemie, D-8033 Martinsried, West Germany, and School of Chemistry, Georgia Institute of Technology, Atlanta, Georgia 30332. *Received February 6, 1990*

Abstract: The crystal structure of trypsin complexed to 4-chloro-3-ethoxy-7-guanidinoisocoumarin, a potent, irreversible inhibitor of both trypsin and thrombin, has been determined to 2.20-Å resolution and refined to an R factor of 0.15 by a geometric restraint least-squares procedure. The results provide positive structural evidence for a mechanism of irreversible inactivation of the enzyme via a quinone imine methide intermediate to form a covalent adduct with the active-site His-57 residue. The chloro acyl enzyme intermediate at Ser-195 is also present. H-bonding interactions formed in the active-site complex by both inhibitor models are summarized. The structures of the complexes of 4-chloro-3-ethoxy-7-guanidinoisocoumarin with porcine pancreatic elastase and bovine trypsin have been compared, and clearly their structures are quite different. Binding conformations of this isocoumarin inhibitor are modeled by active-site homology into the structure of human α-thrombin, which is an enzyme of pharmacological interest and whose structure has recently been solved.

The design of specific, nonpeptidic, mechanism-based inhibitors of serine proteases is currently of great interest because of the causative roles of these proteolytic enzymes in development of numerous degradative diseases such as pulmonary emphysema, arthritis, pancreatitis, and respiratory distress syndrome. In addition, the formation of clots in various thrombotic diseases has been attributed to the action of thrombin and other serine proteases in the coagulation cascade. Most blood coagulation enzymes exhibit trypsin-like selectivity, but with much greater specificity and more complex structures than the digestive enzyme trypsin. Because crystal structures for most of the coagulation enzymes have yet to be elucidated, descriptions of the mechanism and mode of binding of inhibitors to a well-characterized enzyme with comparable substrate specificity such as trypsin can provide reasonable preliminary models on which to base the design of anticoagulant drugs and inhibitors for related proteases.

Various heterocyclic inhibitors (reviewed by Powers and Harper[1]) have previously been investigated by kinetic studies to deduce their modes of binding and mechanism of inhibition of serine proteases. More recently, heterocyclic compounds of the isocoumarin class substituted with basic functional groups were evaluated by solution kinetic studies for their reactivity and specificity toward trypsin and several blood coagulation serine proteases.[2] One inhibitor, 4-chloro-3-ethoxy-7-guanidinoisocoumarin (Figure 1), was found to be one of the most potent mechanism-based irreversible inactivators of trypsin and thrombin. This isocoumarin was active as an anticoagulant in rabbit plasma as well as human plasma.[3] Recently, the high-resolution X-ray crystal structure of this inhibitor bound to another serine protease, PPE,[4] was solved, providing further structural information on the enzyme–inhibitor interactions.[5] In this paper, we report a 2.2-Å resolution crystal structure of 4-chloro3-ethoxy-7-guanidinoisocoumarin complexed to bovine trypsin and compare the results with the evidence for a postulated mechanism of inhibition based on kinetic studies in solution.[2,6] The binding conformation of the inhibitor of trypsin is compared and contrasted with the PPE complex mentioned above. This is one of the few examples where the same low molecular weight inhibitor has been studied structurally with two different but homologous enzymes. As is true for several PPE + isocoumarin complexes,[6] the mode of binding is dramatically different, raising serious doubts about the ability of molecular modeling to predict and/or elaborate on the binding mode of novel inhibitors. However, the electrostatic interactions afforded by "trypsin-like" enzymes provide an additional, perhaps crucial, point of attachment, giving us confidence to design for and build on this specific binding mode. We therefore can propose a possible binding conformation of this inhibitor to human α-thrombin, a key blood coagulation protease whose structure has recently been solved and found to have an active-site structure similar to that of trypsin.[7]

Experimental Methods

4-Chloro-3-ethoxy-7-guanidinoisocoumarin was synthesized according to the method of Kam et al.[2] By use of the method of Bode and Schwager,[8] crystals of β-trypsin were grown in the presence of benzamidine, which was then removed by overnight dialysis against 0.1 M phosphate buffer, first at pH 7.5 and later at pH 5.0, in order to obtain optimum reaction conditions. The crystals were converted via dialysis to 0.1 M phosphate buffer with 10% acetonitrile in order to increase solubility of the inhibitor. A small clump of 4-chloro-3-ethoxy-7-guanidinoisocoumarin was added to the thin-walled capillary that contained the crystal and buffer and placed in physical contact with the crystal. The capillary was placed in a cold room (4 °C) and remained there for 18 days before data collection on the FAST area detector was begun at 5 °C. A Rigaku RU-200 generator with a Cu anode was operated at 120 mA and 45 kV, using a 0.5-mm collimator and Ni filtered radiation. Cell dimensions were estimated by the ALIGN feature

† This work was supported by grants from the Texas Agricultural Experiment Station and the Robert A. Welch Foundation (A328) to E.F.M., from the Deutsche Forschungsgemeinschaft to W.B. (SFB 207/H-1), and from the National Institutes of Health to J.C.P. (HL 29307 and 39035).

‡ Texas A&M University.
§ Max-Planck-Institut für Biochemie.
∥ Georgia Institute of Technology.

(1) Powers, J. C.; Harper, J. W. In *Proteinase Inhibitors*; Barrett, A. J., Salvensen, G., Eds.; Elsevier: Amsterdam, 1986; pp 55-152.
(2) Kam, C.-M.; Fujikawa, K.; Powers, J. C. *Biochemistry* **1988**, *27*, 2547-2557.
(3) Oweada, S.; Ku, D.; Kam, C.-M.; Powers, J. C. *Thromb. Haemostasis*, in press.
(4) Abbreviations used: PPE, porcine pancreatic elastase; SD, standard deviation; rms, root-mean-square; BPTI, bovine pancreatic trypsin inhibitor.
(5) Powers, J. C.; Oleksyszyn, J.; Narasimhan, S. L.; Kam, C.-M.; Radhakrishnan, R.; Meyer, E. F. *Biochemistry* **1990**, *29*, 3108-3118.
(6) Bode, W.; Meyer, E.; Powers, J. C. *Biochemistry* **1989**, *28*, 1951-1963.
(7) Bode, W.; Mayr, I.; Baumann, U.; Huber, R.; Stone, S. R.; Hofsteenge, J. *EMBO J.* **1989**, *8*, 3467-3475.
(8) Bode, W.; Schwager, P. *J. Mol. Biol.* **1975**, *98*, 693-717.

0002-7863/90/1512-7783$02.50/0

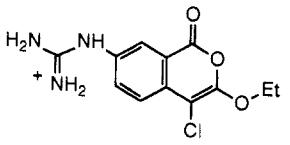

Figure 1. Structure of 4-chloro-3-ethoxy-7-guanidinoisocoumarin.

Table I. Optimum and Refined Values for the Geometries of Covalent Linkages between the Inhibitor and the Active-Site Ser-195 and His-57 Residues

geometric parameter	optimum	SD	refined model 1	refined model 2
Cl-Oγ	1.35 Å	0.01 Å	1.35 Å	1.33 Å
Cl-O1	1.246 Å	0.01 Å	1.23 Å	1.26 Å
O1-C1-Oγ	123°	1°	124.2°	123.2°
Oγ-C1-C2	116°	1°	112.6°	115.3°
C1-Oγ-Cβ	117°	1°	118.2°	120.0°
Nε-C4	1.45Å	0.015Å		1.44Å
Cε-Nε-C4	125.7°	1.0°		126.3°
Cδ-Nε-C4	125.7°	1.0°		125.6°
Nε-C4-C10	109.3°	1.0°		112.4°
Nε-C4-C3	109.3°	1.0°		109.9°

of program MADNES[9] as 55.0, 58.58, and 67.5 Å; the space group was $P2_12_12_1$. Data collection utilized a 100-s time step per 0.1° scan and required ca. 2 days. A total of 17630 reflections out to 1.99-Å resolution was measured from two data sets collected from the crystal. Data were analyzed with a program FASTSA by Huber and Schneider (unpublished), which also corrects for absorption and for multiple measurement of symmetry-related reflections. A total of 8675 unique reflections was obtained, which merged with an agreement of 3.24% to 2.2-Å resolution (completeness of data set to 2.2 Å, 75.7%; last shell, 2.25–2.20 Å, 42.6% complete).

Refinement Procedure. The isomorphous native structure of trypsin, together with associated ions and waters of hydration,[8] was used for initial phasing of the reflections. With the TNT system of least-squares refinement programs,[10] two cycles of refinement were calculated to adjust for small deviations in cell dimensions between the native and inhibited enzymes. The unbiased difference Fourier map generated at this stage clearly indicated continuous residual electron density in the active-site region. By use of the program FRODO[11] on an Evans & Sutherland PS330 graphics terminal, the inhibitor was then modeled into the residual density and included in the refinement. Associated ions and water molecules were added at consecutive steps as they could be assigned from residual density, using the TNTMAX program (S. Swanson, unpublished) to assign electron density peaks to water molecules and program WATER (R. Radhakrishnan, unpublished) to check for bad contacts with the enzyme. In the early stages, atomic temperature factors were refined periodically and averaged over entire residues; in the later stages of refinement, temperature factors for each atom were allowed to vary individually. A total of nearly 400 cycles was calculated, in which 3 cycles of geometric refinement were alternated with a cycle of temperature factor refinement. Full occupancy was assumed on the basis of unambiguous residual density in the initial unbiased difference map.

Geometric constraints were introduced during positional parameter refinements by combining the *R*-factor module (X-ray module) and the geometry module, with the weighting for *R* factor as 0.001, for bonds as 1.0, and for angles as 3.0. The default values of weights in TNT for torsion, plane, and contact distance were retained. The form factor for chlorine[12] was added to the TNT library. Standard parameters for inhibitor geometry[13] were used; geometries for key covalent linkages between enzyme and inhibitors are listed in Table I. Optimum geometric parameters for the ester linkage at the newly formed covalent bond between the inhibitor and Ser-195 were taken from Schweizer and Dunitz.[14] This ester bond was included in the refinement by specifying it as a long-range bond, analogous to a disulfide bond. The TNT library for the geometry module was modified in order to define the above geometric descriptors. The structure revealed the presence of Ca^{2+} in the conserved calcium binding site[8] and a total of 207 water molecules was located.

Table II. Refinement and Crystallographic Parameters for the Final Structure of the Complex of Trypsin and 4-Chloro-3-ethoxy-7-guanidinoisocoumarin

cell constants, Å	55.0 58.58 67.5
space group	$P2_12_12_1$
no. of unique reflections	8675
no. of atoms (both models included)	1889
resolution limits, Å	6.0–2.2
effective resolution,[18] Å	2.34
final *R* factor	0.146
overall temp factor for complex, Å^2	17.1
overall temp factor for inhibitor, Å^2 (model 1)	15.2
overall temp factor for inhibitor, Å^2 (model 2)	23.0
SD for bond length, Å	0.009 (unweighted)
SD for bond angles, deg	1.60
min and max cutoff for temp. factor, Å^2	4.0, 50.0
min and max electron density in final ΔF map, e/Å^3	−0.44, 0.52 ($F_o - F_c$)
mean positional parameter,[19] Å	0.18

Residual electron density was continuous for the first model throughout the entire inhibitor, with well-defined density for the guanidinium and ethoxy substituents and for the attached remnants of the heterocyclic ring, as shown by the $F_o - F_c$ map (Figure 2). The Cl atom was assigned to a prominent electron density peak and appears to remain covalently bound to the inhibitor. Density for the aromatic ring was also present, although some of the ring atoms were not exactly centered in density. The electron density required the placement of His-57 in the "out" position—i.e., not H bonding to either Ser-195 or Asp-102. His-57 has been found in the "out" conformation in most other solved structures of serine proteases complexed with heterocyclic inhibitors.[5,6,15,16]

Model refinement and evaluation were complicated at this stage by the presence of substantial regions of unassigned density, especially around His-57 and the aromatic ring of the inhibitor. In fact, continuous density existed between His-57 and the inhibitor, suggesting another model with covalent attachment. The most likely explanation for the unassigned density was partial occupancy of the inhibitor in another conformation/binding mode. The residual density suggested that this second model involved a doubly covalently bonded enzyme form of the enzyme–inhibitor complex where His-57, acting as a nucleophile, has been alkylated by the benzylic carbon atom of the inhibitor with displacement of Cl (Figure 3).[2,5,6,17] A second model was then constructed to lie within the residual electron density and to form covalent bonds to both Ser-195 and His-57. The two models were assigned partial occupancy values of x and y ($x + y = 1$) and refined simultaneously. The prevalence of either structure could then be compared with data from inactivation/reactivation kinetics of this inhibitor with trypsin in solution.

Ideal geometries for the covalent adduct formed between the inhibitor and His-57 were used[13] and defined in the TNT library for the geometry module (see Table I). Target standard deviations for the geometric parameters at both covalent bonds formed by the inhibitor were restricted during refinement, while for peptide backbone angles, standard deviations were set higher to allow for greater flexibility during refinement, and to ease any possible geometric distortions that might occur as a result of the forced constraints. Standard deviations of backbone angles adjacent to the inhibitor were later reduced gradually if distortions at those points became pronounced—a gradual "tightening of the screws".

The refinement procedure for the second model consisted of three cycles of geometrical refinement, alternated with one cycle each of temperature factor and occupancy refinement. Initially the second model was refined alone, like the first, for about 50 cycles, the first 15 cycles with TNT solvent contribution removed between refinements of the two models due to inflated temperature factor values. In general, the scattering contribution of bulk solvent is not necessarily included in refinement except for calculation of the final *R* factor, and also when making

(9) Messerschmidt, A.; Pflugrath, J. W. *J. Appl. Crystallogr.* **1987**, *20*, 306–315.
(10) Tronrud, D. E.; Ten Eyck, L. F.; Matthews, B. W. *Acta Crystallogr.* **1987**, *A43*, 489–501.
(11) Jones, T. A. *J. Appl. Crystallogr.* **1978**, *11*, 268–272.
(12) Lee, J. D.; Pakes, H. W. *Acta Crystallogr.* **1969**, *A25*, 712.
(13) Kennard, O.; Watson, D. G.; Allen, F. H.; Isaacs, N. W.; Motherwell, W. D. S.; Pettersen, R. C.; Town, W. G. *Molecular Structures and Dimensions*; N. V. a. Oosthoek's Uitgevers Mij: Utrecht, 1972; Vol. A1.
(14) Schweizer, W. B.; Dunitz, J. D. *Helv. Chim. Acta* **1982**, *65*, 1547–1554.
(15) Meyer, E. F.; Presta, L. G.; Radhakrishnan, R. *J. Am. Chem. Soc.* **1985**, *107*, 4091–4093.
(16) Radhakrishnan, R.; Presta, L. G.; Meyer, E.; Wildonger, R. *J. Mol. Biol.* **1987**, *198*, 417–424.
(17) Harper, J. W.; Powers, J. C. *J. Am. Chem. Soc.* **1984**, *106*, 7618–7619.
(18) Swanson, S. M. *Acta Crystallogr.* **1988**, *A44*, 437–442.
(19) Luzatti, V. *Acta Crystallogr.* **1952**, *5*, 802–810.
(20) Abbreviations used in Table III: 11, model 1 (chloro acyl enzyme); 12, model 2 (histidine covalent adduct); HOH, water molecule.

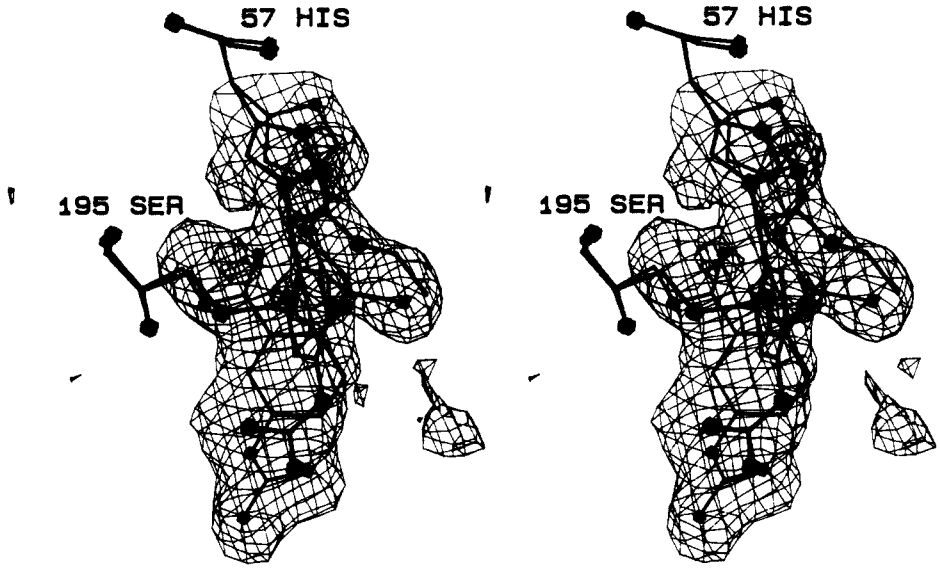

Figure 2. Final difference Fourier map with models of the inhibitor singly covalently bound (model 1, thin lines) and doubly covalently bound (model 2, thick lines) superimposed. The contributions from Ser-195 and His-57 side groups are removed to demonstrate continuous density at the covalent bonds. The contouring level is 0.22 e/Å^3.

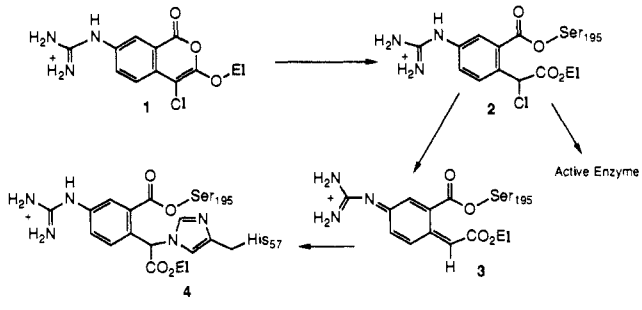

Figure 3. Postulated mechanism of inactivation of trypsin by 3-ethoxy-4-chloro-7-guanidinoisocoumarin.

electron density maps in order to dampen background noise. Then, to account for partial occupancies, the two models were combined into one atomic coordinate file and refined together for another 30 cycles. Initial partial occupancies were assigned to each inhibitor model as 0.5, and each inhibitor was allowed to refine as a whole. Ser-195 and His-57 residues, in assuming different geometries for each inhibitor model, were set at constant occupancies of 0.5, totaling the expected 100% occupancy of these residues in the enzyme. Final refined occupancies for each inhibitor model were around 0.38, suggesting nearly equal ratios of the two major binding forms of inhibitor present in about 75% of the total available binding sites. Temperature factors were arbitrarily set to 20.0 Å^2 and allowed to vary individually between different atoms, with average values for the overall complex and for each inhibitor model given in Table II along with other final refinement parameters. Problems arose with angles at the tetrahedral carbon atom covalently bound to His-57 and at the acyl carbonyl group bound to Ser-195, requiring more modeling using FRODO. At first His-57 was modeled in the "in" position, the standard position assumed in native trypsin allowing a H-bonding network/charge-relay system through the catalytic triad. This conferred *s* chirality on the tetrahedral carbon atom bound to His-57 and was an attempt to account for some excess density around the imidazole ring. Additional modeling revealed that the His-57 would be better centered in density if it were placed in the "out" position and thus was consistent with the first model for this complex. The positioning in density of the carbon atom to which His-57 was bound was improved somewhat and the configuration was converted to *R*. Each time after changing the model, temperature factors were reset to 20.0 Å^2 and occupancies for the models reset to 0.5 before beginning another series of refinement cycles.

For analysis of hydrogen bond lengths and angles, the coordinates of hydrogen atoms were generated from the final set of atomic coordinates with the program HYDGEN (R. Radhakrishnan, unpublished results). H atoms were added onto donor atoms in geometrically feasible positions, taking into account both covalent and H-bonding stereochemistry. Donor–acceptor distances and angles at hydrogen atoms for H bonds formed in the active-site complex are summarized in Table III and illustrated in Figure 4.

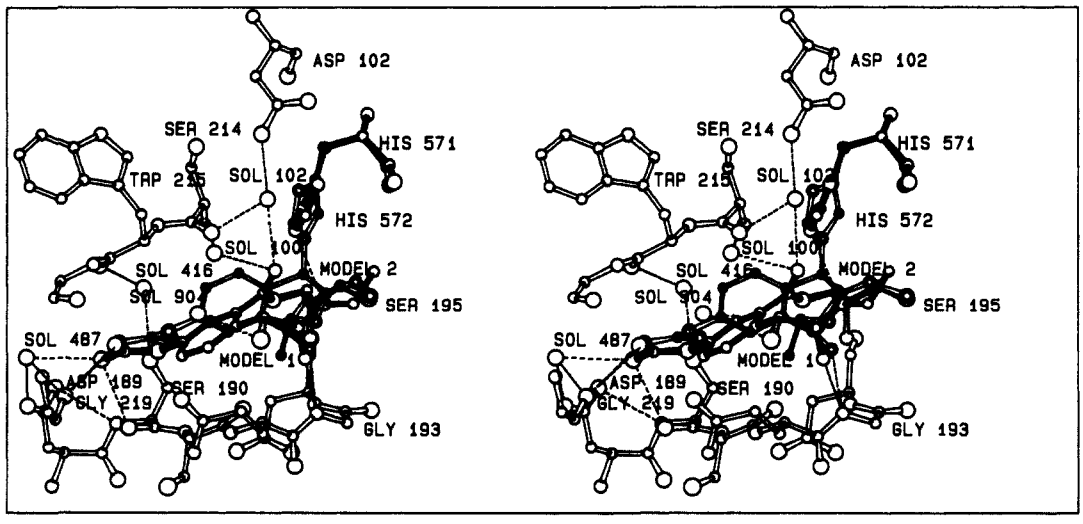

Figure 4. Stereoview of both conformations formed by the inhibitor in the active site of trypsin. Hydrogen bonds are represented by dashed lines. Model 1 (single attachment) is drawn with filled bonds and open atoms and is shown making electrostatic contact with Asp-189 via water (HOH 487). Model 2 is drawn with open bonds and partially filled atoms and is doubly attached to the enzyme.

Table III. Hydrogen Bonds in the Active Site of the Complex[20]

donor	acceptor	distance, Å	angle at H, deg
	Model 1		
I1>N2H$^+$	HOH 487	3.04	180
I1>N2H$^+$	O=C<Ser-190	3.19	117 (weak)
HOH 100	O1=C<I1	3.03	180
HOH 102	O1=C<I1	3.16	180
Gly-193>NH	O2=C<I1	3.17	140
HOH 904	Cl<I1	3.05	180
I1>N3H$^+$	O=C<Gly-219	2.92	162
Conceivable—Nε of His-57 if Protonated, with O3 of I1			
His-57>NεH	O3<I1	3.21	125 (weak)
	Model 2		
Gly-193>NH	O1=C<I2	2.82	165
I2>N3H$^+$	O=C<Gly-219	3.17	152
	Hydrogen Bonds Other Than with Inhibitor		
HOH 487	O=C<Gly-219	3.02	180
HOH 487	Oδ=C<Asp-189	2.92	180
Ser-190>NH	Oδ=<Asp-189	2.88	135
HOH 100	O=C<Ser-214	3.28	180 (weak)
HOH 102	O=C<Ser-214	2.57	180
HOH 102	Oδ=C<Asp-102	2.60	180
Ser-190>OγH	HOH 416	3.02	180
HOH 416	O=C<Trp-215	3.03	180

Results and Discussion

Description of Key Features of the Models. A proposed scheme for the reaction of 4-chloro-3-ethoxy-7-guanidinoisocoumarin and trypsin along with the possible reaction products is shown in Figure 3. The electron density showed that a covalent ester bond was formed between both inhibitor models and Ser-195 in the active site (Figure 2, bond length 1.33 Å). In the first model (Figure 3 (**2**), chloro acyl enzyme), the positively charged guanidinium group, which was expected to provide electrostatic interaction with Asp-189 in the S_1 primary specificity site of trypsin (notation of Schechter and Berger[21]), appears indeed to extend into the pocket. However, the two groups come no closer than 3.6–3.7 Å, which suggests that any possible electrostatic interaction is weak at best unless a water molecule comes between them as a H-bond donor/acceptor. HOH 487 apparently fulfills this purpose as it is located between Asp-189 and the guanidinium nitrogen atoms at the distances and angles shown in Table III.

The excess density neighboring His-57 was modeled as a water molecule (HOH 102) that is able to slip in when the imidazole ring is directed in the "out" position, forming a H-bond bridge between the carboxylate group of Asp-102 and carbonyl group of the benzoyl ester in the first model (**2**). This result is consistent with similar structural analyses of isocoumarins with PPE, which show His-57 to be in the "out" position.[5,6,15] In the first model, HOH 102 was initially in short contact with the peptide backbone carbonyl oxygen atom of Ser-214 (2.28 Å vs a 2.7-Å ideal distance), which increased the rms value to 0.065 Å. However, 50 additional cycles of refinement reduced the rms value to 0.041 Å and essentially relieved the contact distance, uncovering a third possible H bond that could be formed by this water molecule with Ser-214. This third hydrogen bond may exist only for the second inhibitor model, since the benzoyl ester carbonyl group is not in an orientation to form the H bond as in model 1 (see Figure 4).

The second model fit into continuous residual density, accounting for unassigned electron density left in the first model except for areas of excess density left near His-57 and along the aromatic ring of the second inhibitor model that were eventually modeled as water molecules HOH 102 and HOH 100 (see Figure 4). Continuous density was found for the carbonyl carbon atom that forms the acyl–enzyme covalent bond only at the 1σ level, at 0.1 e/Å^3. As discussed in the refinement procedure, centering His-57 in density results in the *R* configuration at the carbon atom covalently bound to His-57 (as shown in Figures 2 and 4). However, the density surrounding the bond itself and the tetrahedral carbon atom cannot distinguish between *R* and *S* configurations, and therefore, the assignment of *R* chirality to **4** in Figure 3 is a tentative one. The configuration of the carbon atom with Cl still attached can unambiguously be assigned as *R* in the first model (chloro acyl enzyme).

Evidence for Inhibition Mechanism. The stereochemistry of the covalent adduct (**4**) and the chloro acyl enzyme (**2**) may reveal additional details concerning the inactivation mechanism. If the configuration inverts upon going from the chloro acyl enzyme to the covalent adduct, the mechanism could involve either direct S_N2 nucleophilic displacement by His-57 or a quinone imine methide intermediate (Figure 3 (**3**)), whereas if retention of configuration occurs from the chloro acyl enzyme to the histidine adduct (**4**), then the quinone imine methide intermediate would be required and direct displacement could effectively be rule out. The evidence from our X-ray crystal analysis supports retention of configuration, providing structural support for the formation of a quinone imine methide intermediate during the process of irreversible inactivation for this particular enzyme–inhibitor complex. In contrast, 7-amino-3-(2-bromoethoxy)-4-chloroisocoumarin forms a doubly covalent adduct with PPE having the *S* chirality and His-57 in the "in" position.[22] The factors influencing favored conformation of His-57 in the active site require closer scrutiny.

The occupancies of both inhibitor models retain an almost equal 1:1 ratio, suggesting that only half of the inhibitor molecules found in the active site have irreversibly inactivated the enzyme by alkylating the active-site nucleophilic His-57 residue, while the rest could still be found as the potentially reversible chloro acyl enzyme intermediate. Clearly, the sample preparation procedures used have established a steady-state condition; after 18 days of soaking (pH 5.0, 4 °C) and 2 days for data collection (5 °C), a reaction equilibrium may be presumed to exist. In comparison, solution kinetic studies[2] found virtually total irreversible inactivation, with only 3–8% of the enzyme activity regained upon incubation in the presence of a pH 7.5 buffer or by the addition of buffered hydroxylamine. However, at pH 5.0, 70% of the enzyme activity was regained in the presence of 0.45 M hydroxylamine with a half-life of 2 h 25 min. Inhibited PPE did not regain any activity in the pH 5.0 buffer after staying at room temperature for 2 days. There are several possible explanations for the differences between the solution- and solid-state results. First, it is possible that after Ser-195 Oγ attacks to form the acyl enzyme (**2**), there is a time lag before subsequent formation of the quinone imine methide intermediate (**3**) and attack by His-57 to form the alkylated enzyme (**4**). In order to reach the proper geometry for nucleophilic attack by His-57, the plane of the inhibitor, as defined by its aromatic ring, must "flip up" or swivel so that the carbon atom targeted for attack is within striking range of the nucleophile. In the process, the guanidinium group shifts away from its favorable Coulombic interaction with Asp-189 by 2.5–3.0 Å (see Figure 4 and 7) to distances of 5.5–6.5 Å, separating the carboxylate oxygen and guanidinium nitrogen atoms. In addition, a comparison of the H bonds formed within the active site by the first and second models shows that mode 1 is stabilized by more H bonds than model 2 (Table III). These factors may in turn help to explain the postulated delay in formation of the covalent adduct and suggest why the chloro acyl enzyme, with the guanidinium group deeper in the specificity pocket and more H bonds to hold it in place, persists. Not to be overlooked is whether the *potentially* reversible chloro acyl enzyme is all that easily reversible or not. Deacylation of the ester bond with Ser-195 requires regeneration of the transient tetrahedral intermediate and stabilization of the oxy anion in the oxy anion hole. However, the benzoyl ester carbonyl oxygen atom is directed elsewhere, and nucleophilic attack would have to occur from the relatively inaccessible side of the plane of the ester bond facing the interior of the primary specificity pocket in order for the oxy anion that

(21) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.

(22) Vijayalakshmi, J.; Meyer, E.; Kam, C.-M.; Powers, J. C. *Biochemistry*, submitted.

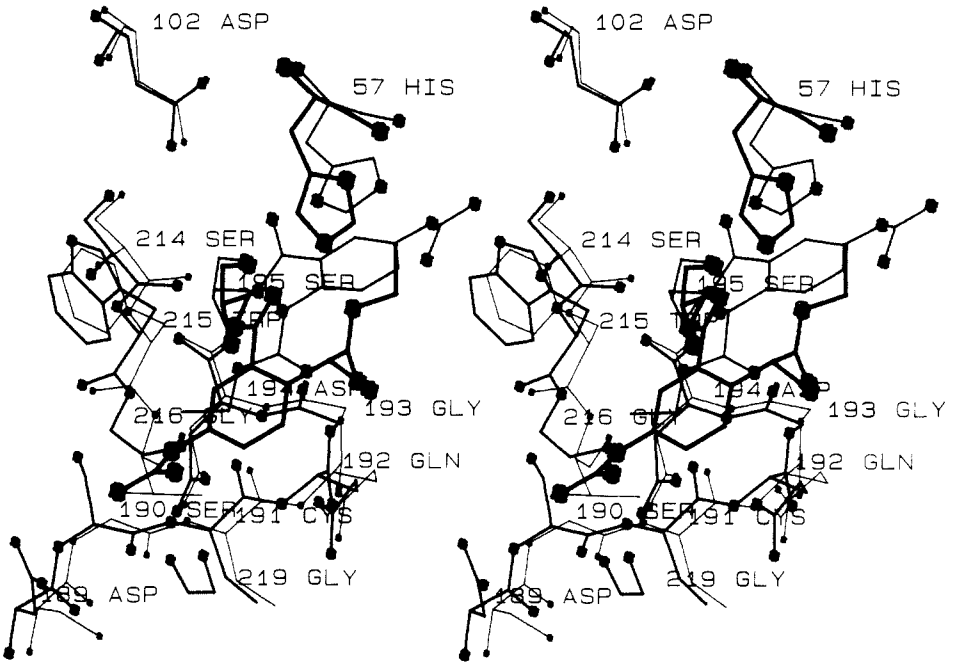

Figure 5. Stereoview of the active sites of PPE (thin lines) and of trypsin (medium lines) superimposed, containing the chloro acyl enzyme conformation of the inhibitor in each enzyme. N and O atoms are indicated by "spots" created by routine BLOB (A. Karrer, unpublished), which has been submitted for inclusion in program FRODO.[9]. The trypsin model is drawn with thick lines and large "blobs", the PPE model with thinner lines and smaller "blobs". Note that the two structures are oriented nearly 180° opposite to each other.

forms to be stabilized by the oxy anion hole. The apparently mobile ethoxy group could also sterically hinder access of a solvent nucleophile to ester carbonyl group. So, given the inhibitor geometry shown in Figure 4, a deacylation event seems less likely, whether by hydrolysis of the ester or regeneration of the intact isocoumarin. Thus, acyl enzymes can be stable entities that resist reverting to active enzyme and contribute to what appear to be irreversible inactivation kinetics, in addition to the presence of the histidine covalent adduct. Similar reasoning has been used for the complex of 4-chloro-3-ethoxy-7-guanidinoisocoumarin with PPE[5] and two benzoxazinone complexes with PPE.[16] Another possible factor concerns the pH during the reaction and how that affects the protonation state of His-57, which has a presumed pK_a of 6.0. At pH 7.5, the pH of the solution kinetic studies, His-57 is mostly unprotonated and can act as a good nucleophile, thus favoring formation of the irreversible histidine covalent adduct. However, the crystals were inhibited at pH 5.0; His-57 would therefore be predominantly (~90%) protonated and unable to react with the inhibitor to form the double covalent adduct, thus favoring the continued presence of the chloro acyl enzyme intermediate. The reactivation studies conducted by Powers et al.[4] of the guanidinoisocoumarin inhibitor with or without a 4-chloro group in PPE definitively demonstrated that, at pH 7.5, alkylation of His-57 seemed to be the pathway of choice, while at pH 5.0, formation of a stable acyl enzyme derivative was favored. The solution kinetic studies with trypsin at the lower pH, as mentioned previously, reaffirm this hypothesis. Variable partitioning between a reactivatable and nonreactivatable acyl enzyme has been observed during solution kinetic studies of other serine protease complexes with isocoumarin.[23] With PPE inhibited by 7-amino-4-chloro-3-methoxyisocoumarin at pH 5.0, 32–57% of the activity can be recovered by hydroxylamine treatment, depending on the buffer used for the inactivation. As in the case of the trypsin complex crystal structure reported here, the two elastase inhibitor complexes are nearly equally abundant in solution.

Comparison of the Trypsin and PPE Complexes. The structures of the complexes of 4-chloro-3-ethoxy-7-guanidinoisocoumarin with both PPE[5] and bovine trypsin are shown in Figure 5. This is one of very few examples where crystal structures of the same inhibitor bound to two related enzymes have been determined. Clearly the structures are quite different, with orientation of the opened isocoumarin rings twisted almost 180° opposite to each other. In the PPE crystal structure the 3-ethoxy group occupies the S_1 pocket, while the 7-guanidino group takes a similar position in the trypsin structure. It is apparent, upon examining Figure 5, that Val-216 in the S_1 pocket of PPE would have collided with the guanidinium group of the inhibitor in the chloro acyl enzyme model for trypsin, thus ruling out that orientation; also, the characteristics of the side groups extending into their respective primary specificity pockets are in agreement with the preference of PPE for small, hydrophobic residues and trypsin for positively charged, basic residues. With trypsin, both a chloro acyl enzyme (**2**) and a histidine adduct (**4**) are formed, while only the chloro acyl enzyme structure is formed with PPE. In the PPE complex the carbonyl group is twisted away from the oxy anion hole. The same is true of the chloro acyl enzyme structure in trypsin. However, the carbonyl group in trypsin does form H bonds with two water molecules located within the active site, HOH 102 and HOH 100 (Figure 4). These waters, in turn, form H-bonded bridges with the backbone carbonyl oxygen atom of Ser-214 and the carboxylate group of Asp-102. What one finds directed into the oxy anion hole instead is the carbonyl group formed as a result of cleavage of the isocoumarin ring (re: Table III, Figure 4). Interestingly enough, with the histidine adduct, the benzoyl carbonyl oxygen atom is again pointing at least partially toward the oxy anion hole, forming a H bond with Gly-193 NH. Clearly, the initial Michaelis complexes, which eventually lead to the two product crystal structures, may be presumed to be quite different and there must be at least two productive binding modes for isocoumarins with serine proteases, one with the 3-alkoxy substituent in the S_1 pocket and one with the 7-substituent in S_1. Most likely the S_1 primary specificity site would be the determining factor in distinguishing and selecting for productive vs nonproductive binding, especially for these small-molecule inhibitors.

Thrombin Model. The crystal structure of human thrombin, a trypsin-like serine protease, has recently been solved to 1.9-Å resolution.[7] Active-site residues of trypsin and thrombin can be superimposed (rms = 0.31 Å) and the putative binding geometry of the guanidinoisocoumarin complexed to thrombin can be predicted by the similarity of the enzymes' active-site regions, as shown in Figures 6 and 7. While any such predicted model is only a working hypothesis, we feel confident, on the basis of the

(23) Harper, J. W.; Powers, J. C. *Biochemistry* **1985**, *24*, 7200–7213.

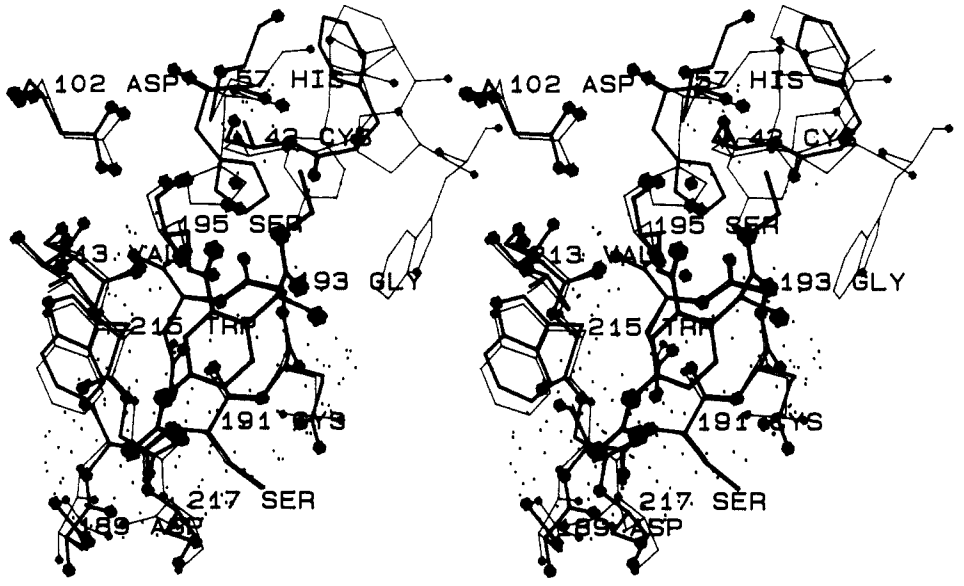

Figure 6. Stereoview of the active sites of thrombin (thin lines) and of trypsin (medium lines) superimposed, containing the chloro acyl enzyme form (model 1) of the inhibitor (heavy lines). The extent of enzyme–inhibitor interaction is illustrated by the double van der Waals surface contact surface around the inhibitor, generated by the INTER function of program FRODO.

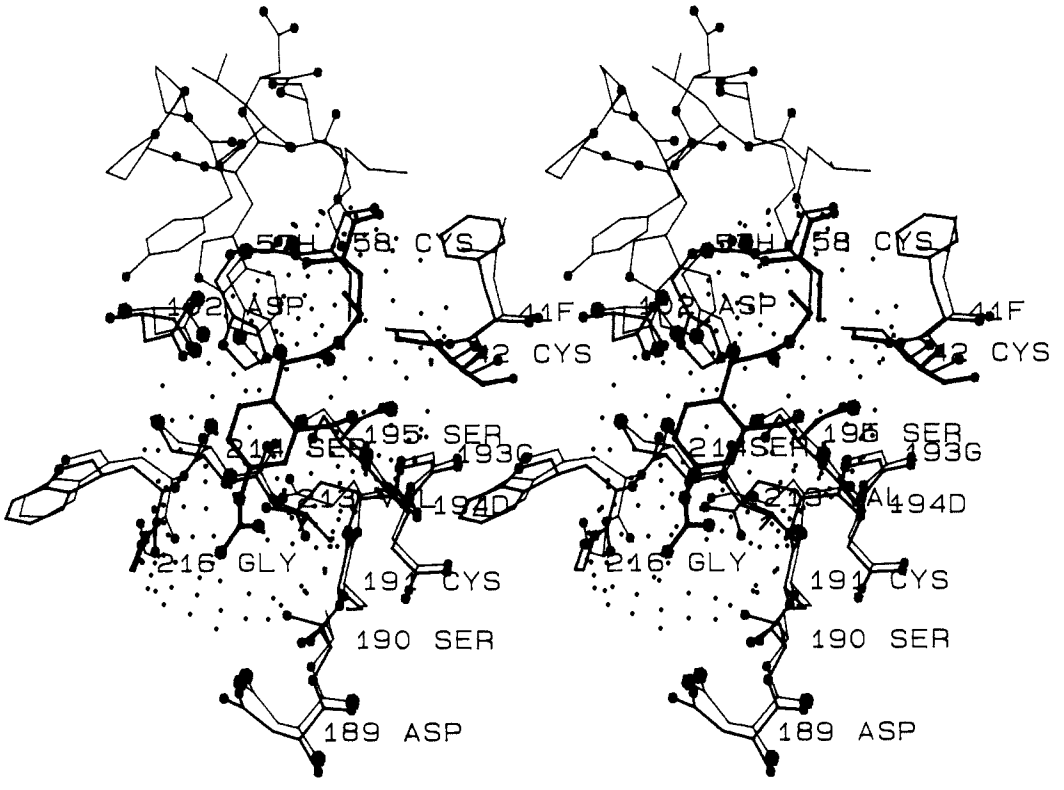

Figure 7. Stereoview of the active sites of thrombin and trypsin superimposed, containing the doubly covalently linked form (model 2) of the inhibitor. Assignments are as in Figure 6. No close contact of the guanidinium group to Asp-189 is allowed, suggesting an extension of this arm would increase binding specificity and reactivity. As in Figure 6, the ethoxy group points toward the 60A–60H insertion loop of thrombin (top center).

structures of other trypsin complexes[8,24,25] as well as the thrombin crystal structure,[7] that the negatively charged Asp-189 residue at the bottom of the S_1 pocket is an additional determinant that holds even for the highly geometrically constrained model 2 reported here. On the basis of this homology, we feel confident about studying the fit of isocoumarins to thrombin.

Of the two models reported here, the guanidinium group of model 1 (chloro acyl enzyme) could make a single H bond (3.4 Å) with Asp-189 of thrombin (vs the experimentally determined interaction distances of 3.7 Å in the trypsin structure and 2.6–2.8 Å for the Arg guanidino group of D-Phe-Pro-Arg-CH_2Cl complexed with thrombin;[7] see Figure 6); however, direct electrostatic interaction is not likely, as is the case for the reported trypsin complex. A water molecule may serve as a H-bond donor/acceptor bridge, as has been demonstrated in the trypsin complex with HOH 487 and which is reminiscent of the binding geometry of the Lys-151 side chain in the trypsin + BPTI complex.[24] However, the second inhibitor model (Figure 7), being covalently linked to His-57, does not permit deep penetration of the guanidinium group into the S_1 primary specificity site (4.8 Å in the thrombin model, 5.1 Å in the trypsin complex). Most likely another water molecule mediates this interaction as a H-bond

(24) Huber, R.; Bode, W. *Acc. Chem. Res.* **1978**, *11*, 114–122.
(25) Matsuzaki, T.; Sasaki, C.; Okumura, C.; Umeyama, H. *J. Biochem.* **1989**, *105*, 949–952.

bridge as in the trypsin + BPTI complex or a hydroxyl (OH^-) ion serves as a counterion, since H-bond formation cannot occur over such a distance and the presence of naked positive charge on the guanidinium group would be energetically highly unfavorable. Here, competition with the model I side chain in the S_1 pocket does not permit crystallographic resolution of such a water molecule or OH^- ion other than HOH 487, which is associated with the chloro acyl enzyme and not the histidine covalent adduct.

Bovine thrombin is effectively inhibited by 3-alkoxy-4-chloro-~7-guanidinoisocoumarins.[2] The most effective inhibitor is the 3-methoxy derivative with a second-order inhibition rate constant ($k_{obsd}/[I]$) of 290000 M^{-1} s^{-1}. The 3-ethoxy (55000 M^{-1} s^{-1}) and 3-phenylethoxy (30000 M^{-1} s^{-1}) derivatives are less potent. The kinetics of these three derivatives in trypsin parallel their behavior in thrombin, with second-order inactivation rates of 310000 and >110000 M^{-1} s^{-1} for 3-methoxy, and 3-ethoxy and 3-phenylethoxy derivatives, respectively. Reactivation kinetics for these inhibitors in trypsin and thrombin are also comparable, with the formation of very stable inactivated enzyme and inhibited trypsin regaining only 1–4% activity upon standing or 8–16% with the addition of hydroxylamine, while inhibitied thrombin regained <9% or 15% activity under the same conditions. Thus, it seems reasonable to conclude that these guanidinoisocoumarins behave similarly in both enzymes in terms of binding and inhibition mechanism. Although kinetic data were collected for bovine thrombin and not human thrombin, it should be noted that both enzymes have exactly the same sequence at the insertion loop 60A–60H. Therefore, there should be no difference between bovine and human thrombin with respect to interactions of isocoumarins with this portion of the structure, and the basis for drawing parallels between structural modeling in human thrombin and kinetic data with bovine thrombin is reasonable. At the moment, no kinetic data with human thrombin or a ready source of the enzyme are available (J. C. Powers, personal communication).

In the two models of the thrombin–isocoumarin complex, the 3-alkoxy group is directed toward thrombin's insertion loop 60A–60H (consisting of residues Tyr-Pro-Pro-Trp-Asp-Lys-Asn-Phe). This loop affords the opportunity for enhanced active-site interactions; the inhibitor alkoxy group is pointed toward Lys-60F, which may be an explanation for the observation that the longer, more hydrophobic 3-alkoxy derivatives are less potent inhibitors than the 3-methoxy derivative. In fact, Figure 7 clearly shows the van der Waals surface of the inhibitor 3-ethoxy side chain overlapping the Lys-60F residue in the loop. It is clear that selective substitution on the 3-alkoxy group of the isocoumarin would provide greater interaction with this region of thrombin's active site and should lead to more potent and specific isocoumarin inhibitors for thrombin. It has been previously demonstrated that a serine protease–isocoumarin crystal structure is an ideal starting point for the design of additional, more potent inhibitors;[5] the report of the title structure now permits us to extend our modeling efforts to analogous members of the trypsin-like proteases with greater confidence, especially because supporting solution kinetics data are also available. To the extent that the homology of binding modes of isocoumarins to trypsin and thrombin is constrained by the electrostatic attractions afforded by the S_1 pocket, we feel confident in pursuing such modeling efforts. However, such models are clearly based upon initial, experimental evidence (i.e., the title complex), which again shows the importance of high-resolution crystallographic analyses of key enzyme–inhibitor complexes in forming a basis for rational molecular modeling.

Conclusion

We have determined the crystal structure of bovine trypsin inhibited by 4-chloro-3-ethoxy-7-guanidinoisocoumarin and constructed models of human α-thrombin bound by the same inhibitor. One form of the inhibited enzyme contains covalent bonds to both Ser-195 and His-57. This structure, and the 7-amino-3-(2-bromoethoxy)-4-chloroisocoumarin complex with PPE,[22] are the first two examples of a doubly covalently bound isocoumarin derivative with a serine protease. The binding geometry of the 4-chloro-3-ethoxy-7-guanidinoisocoumarin to the active site of trypsin is substantially different from that observed in previously reported structures of elastase inhibited by isocoumarins and other heterocyclic inhibitors.[5,6,15] Small-molecule heterocyclic inhibitors have been found to occupy a number of different geometries in the active sites of serine proteases—or in this case, even within the active site of the same enzyme—and can react by several competing reaction pathways. The unexpected observations reported here for the trypsin complex necessitate, again, a word of caution to molecular modelers who generate what seem to be plausible inhibitors in the absence of physical and/or chemical evidence; nature again proves to be too wily for simplistic models. Crystal structures of these inhibitor complexes are therefore invaluable for defining the rules—or trends or tendencies as it may be—that determine the binding of small molecules to related enzymes and for designing therapeutically useful inhibitors of physiologically important enzymes such as thrombin and other trypsin-like serine proteases.