Atomic structure of the trypsin–A90720A complex: a unified approach to structure and function

Angela Y Lee[1], Tim A Smitka[2], Rosanne Bonjouklian[2] and Jon Clardy[1]*

[1]Department of Chemistry, Baker Laboratory, Cornell University, Ithaca, NY 14853–1301, USA
and [2]Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN 46285, USA

Background: A90720A is a potent serine proteinase inhibitor produced by the terrestrial blue–green alga *Microchaete loktakensis*. Most of its structure has been defined by spectroscopic and degradative studies, but the configurations of several stereochemical centers are unknown, and its mode of inhibition of serine proteinases is not understood. We therefore examined the structure of the compound in a complex with trypsin.

Results: We have crystallized the bovine trypsin–A90720A complex and determined its three-dimensional structure at 1.90 Å resolution using single crystal X-ray diffraction. The structure of the bound inhibitor is clearly shown in the electron density. The structure defines the absolute stereostructure of A90720A, establishes its bound conformation and illuminates its mode of inhibition.

Conclusions: A90720A interacts with trypsin in a substrate-like manner through an extensive series of hydrogen bonds, hydrophobic interactions and steric complementarity. The compound uses a mixture of peptidal and nonpeptidal elements to imitate the canonical conformation of the exposed binding loop of 'small' proteinase inhibitors.

Chemistry & Biology October 1994, **1**:113–117

Key words: crystal structure, cyclic peptide, *Microchaete loktakensis*, proteinase inhibitor, serine proteinase

Introduction

Proteolysis constitutes a major theme in biology; understanding and controlling proteolysis thus forms an important goal for chemistry. The search for serine proteinase inhibitors, especially low molecular weight thrombin inhibitors that could become therapeutic agents for stroke and coronary artery occlusion [1], has been particularly intense [2–4]. Natural products often suggest new molecular motifs; in an effort to find novel biologically active materials, a collaborative program at Eli Lilly and the University of Hawaii has screened blue–green algal metabolites [5]. The terrestrial blue–green alga, *Microchaete loktakensis*, produces a cyclic depsipeptide, A90720A (Fig. 1), with potent inhibitory properties for serine proteinases (IC_{50} = 275 nM for bovine thrombin, 10 nM for bovine trypsin, and 30 nM for human plasmin). Although spectroscopic techniques were sufficient to reveal the planar structure of A90720A [6], it was not possible to determine the configuration of several stereochemical centers or the conformation of the molecule, and its mechanism of inhibition of serine proteinases remained unknown. In an effort to address all of these unresolved issues at once, we used X-ray crystallographic methods to determine a high resolution (1.9 Å) three-dimensional structure of A90720A bound to bovine trypsin.

A90720A has a 19-membered depsipeptide ring and a side chain terminating in a sulfated glyceric acid. The 3-amino-6-hydroxy-2-piperidone (Ahp) unit, which can be thought of as a glutamate in which the side chain carboxyl has been reduced to an aldehyde and the aldehyde cyclized with the n+1 amide to form a cyclic hemiaminol, is arguably its most interesting structural feature. Ahp was first described as a constituent of dolastatin-13 [7], a cytostatic agent from the sea hare *Dolabella auricularia*, and has recently been found in several metabolites of blue–green algae, namely the aeruginopeptins [8], cyanopeptolins [9], microcystilide A [10], and the micropeptins [11], all of which were isolated from *Microcystis* species. Although *Microchaete loktakensis* is taxonomically quite distinct from *Microcystis*, A90720A bears a close structural relationship to the *Microcystis*-derived metabolites. Most Ahp-containing metabolites have been discovered in antibacterial or anticancer-based assays, but the micropeptins [11] selectively inhibit

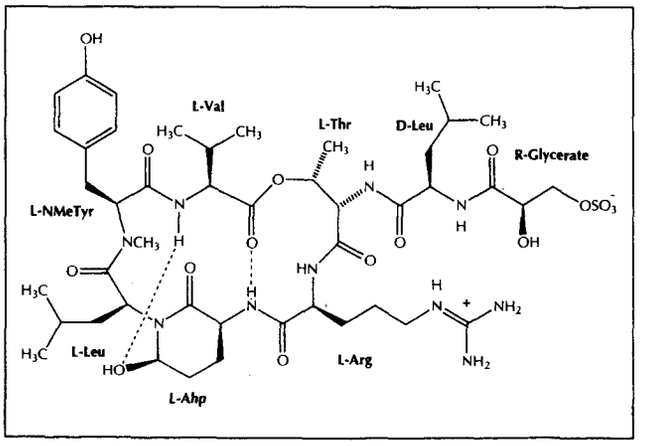

Fig. 1. Structure of A90720A.

*Corresponding author.

 ISSN 1074–5521

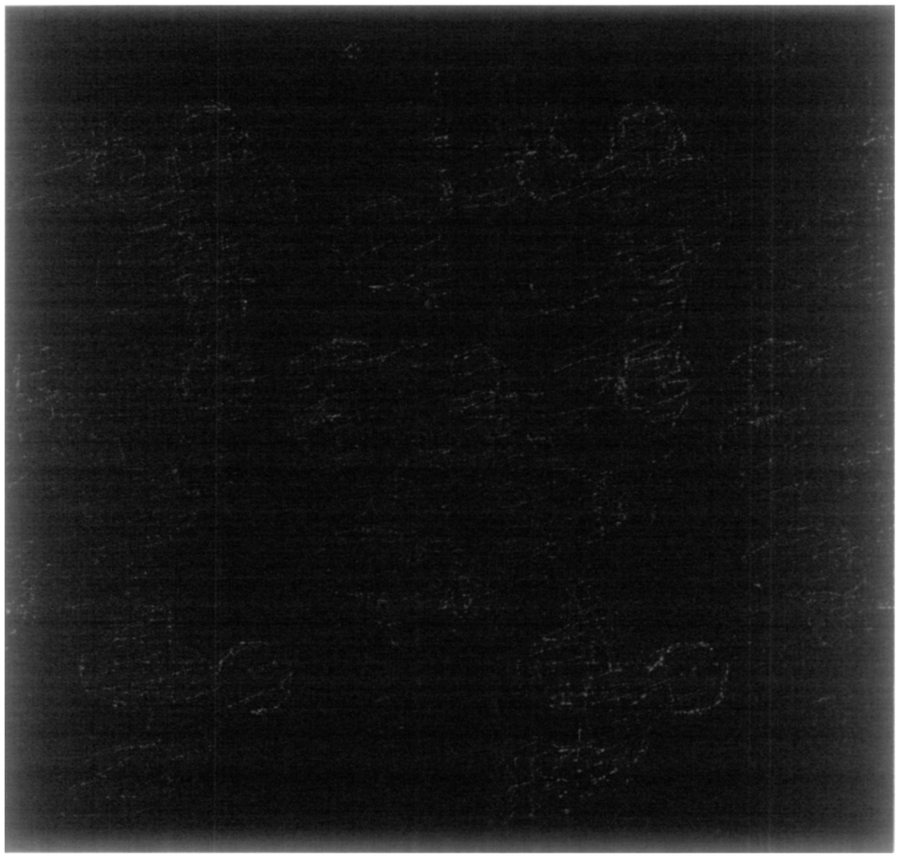

Fig. 2. Stereo view ... the electron density ($2|F_o| - |F_c|$ ...) in the active site of the complex. F ... the catalytic triad of trypsin, Ser195 ... d His57, and the Ahp and argini ... residues of A90270A are shown. ... e the absence of continuous electron ... nsity between protein and inhibitor ... The plot was made using CHAIN [23]

trypsin and plasmin but not other serine proteinases. Thus, an understanding of how A90720A functions could lead to more general insights for other Ahp-containing depsipeptides.

Results and discussion

Conformation and interactions of A90720A

The structure of A90720A is clearly revealed by the continuous electron density ($2|F_o|-|F_c|$) in the final X-ray model illustrated in Fig. 2. The bottom half of the molecule is deeply buried in the catalytic pocket, whereas the upper half, from tyrosine to the glyceric sulfate, is at least partially exposed to solvent. The 19-membered ring is elliptical, and this shape seems to be determined by two hydrogen bonds involving Ahp, one between the lactone carbonyl and the Ahp NH (2.81 Å), a ... d the other between the ring hydroxy of Ahp and t ... e valine NH (3.05 Å) (Fig. 3; see also Fig. 1).

The inhibitor employs several strategies t ... ensure tight binding in the catalytic pocket. The inhibit ...'s side chain, the portion of A90720A running from the ... lyceric sulfate to the amide NH of arginine, uses five hydr ... gen bonds to interact with the protein (Fig. 4a). The side chain is held so that its bottom half (Fig. 4b) is held again ... t the protein surface and its top half is exposed to solver ... The protein contacts are: Ser217 OγH to sulfate oxy ... en (2.86 Å), Gly219 NH to the glyceric free hydro ... yl (2.79 Å), Gly216 O to D-Leu NH (2.90 Å), Gly 216 NH to D-Leu O (3.04 Å), and Ser214 O to L-Arg ... NH (3.12 Å) (Fig. 4a). In the other Ahp-containing met ... bolites, all of

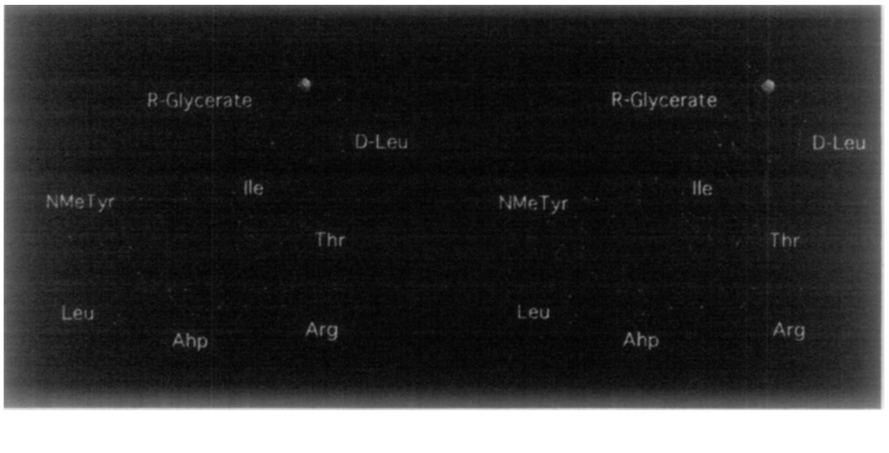

Fig. 3. Stereo view of th ... bound conformation of A90720A. Bo ... n the hydroxyl OH and amide NH of t ... e Ahp residue make transannular hy ... lrogen bonds. Imaging software describ ... d in [24].

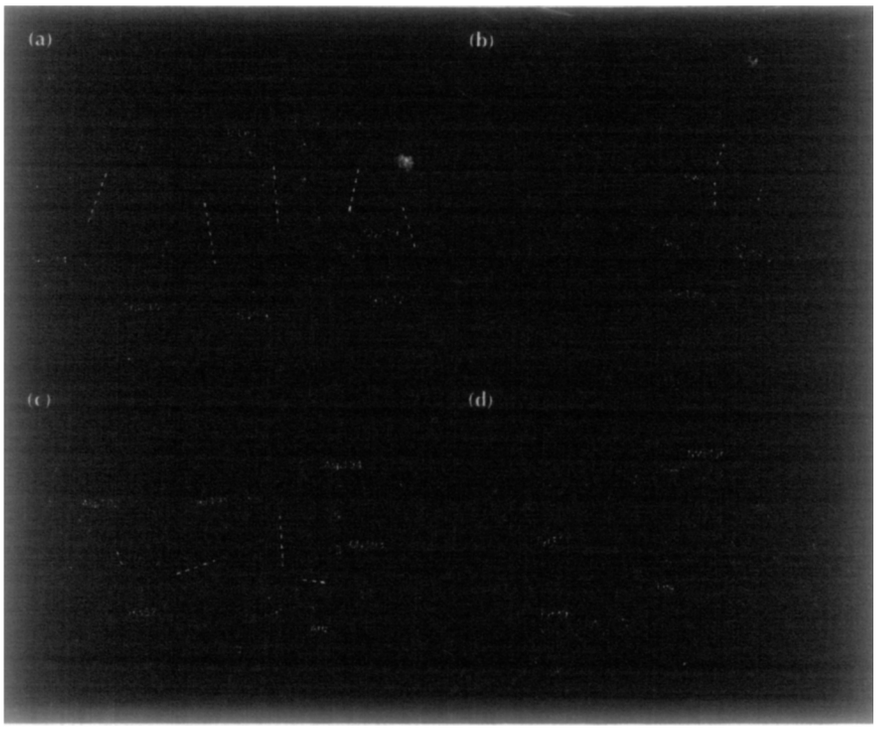

Fig. 4. Hydrogen-bonding interactions between A90720A and trypsin. **(a)** The hydrogen-bonding interactions between A90720A's side chain (top) and bovine trypsin (bottom). **(b)** The hydrogen-bonding interactions of A90720A's arginine side chain with residues in the specificity pocket of trypsin. **(c)** The catalytic site of trypsin occupied by A90720A. Note that the Ser195 hydroxyl is hydrogen bonded to His57 and that the carbonyl oxygen of the A90720A's arginine residue is hydrogen bonded to the main chain NHs of Gly193 and Asp194. **(d)** A hydrophobic pocket of bovine trypsin occupied by the leucine and *N*-methyl-tyrosine of A90720A.

the amino acids are L, the D-leucine in this region of A90720A is the sole exception. If this residue were L-leucine, and if the hydrogen-bonding scheme were preserved, the hydrophobic side chain would project into solvent (Fig. 4a). This series of hydrogen bonds resembles the hydrogen bonds formed between serine proteinases and their substrates or their 'small' protein inhibitors [12–14]. In the case of A90720A, however, the sulfate residue adds two hydrogen bonds to this interaction — a strategy unavailable to a substance that is solely constructed of standard amino acid residues.

The arginine of A90720A occupies trypsin's deep, narrow specificity pocket, making full use of the hydrogen-bonding capabilities of the guanidinium (Fig. 4b). The guanidinium has two hydrogen bonds with Asp189 at the bottom of the pocket (2.75 and 2.87 Å), and forms hydrogen bonds to the hydroxy side chain of Ser190 (2.62 Å), the carbonyl oxygen of Gly219 (2.99 Å), and water-mediated hydrogen bonds to Gly216 and Gly219 (Fig. 4b).

If A90720A were a trypsin substrate, the arginine carbonyl would be attacked by the nucleophilic hydroxyl of the side chain of Ser 195, which is rendered more nucleophilic by losing a proton to His57 (Fig. 4c). The side-chain hydroxyl is hydrogen bonded to His57 (2.69 Å), but is 2.77 Å away from the arginine carbonyl carbon; the electron density between the two atomic centers shows a pronounced discontinuity (Fig. 2). The arginine carbonyl oxygen is hydrogen bonded to the amide NHs of both Ser195 and Gly193 in the oxyanion hole (Fig. 4c).

Finally, the hydrophobic region of the inhibitor, defined by the *N*-methyl-tyrosine and its adjacent leucine (Fig. 4d), provides further binding interactions. The leucine sits in the center of an aromatic binding pocket, leading to the prediction that replacing this residue with an aromatic one would generate an even more potent trypsin inhibitor. The other residues in the aromatic pocket are Tyr39, Tyr151 and the *N*-methyl-tyrosine of A90720A. There is a strong hydrogen bond between the phenolic hydroxyls of Tyr151 and 1 (2.77 Å) (Fig. 4d).

Mechanism of inhibition

It is instructive to compare A90720A with another cyclic peptide-based serine proteinase inhibitor, cyclotheonamide A [15–17]. Cyclotheonamide A inhibits by forming a covalent complex; A90720A forms a noncovalent substrate-like complex. A90720A inhibits by two complementary strategies: stabilizing a substrate-like complex and preventing dissociation. All of the inter-actions mentioned above — the hydrogen bonds to the side chain, the arginine in the specificity pocket, and the hydrophobic binding pocket for Leu and *N*-methyl-tyrosine — stabilize the interaction of A90720A with trypsin. This exquisite complementarity of the enzyme–inhibitor interaction leads to a deep energy minimum for the complex, and a correspondingly high activation barrier in the free-energy profile for hydrolysis [12,14,18]. But even if the transition state for proteolysis were reachable, the overall hydrolysis reaction is unlikely to proceed because the amine leaving group is held in place by the cyclic nature of the depsipeptide, reinforced by the two Ahp hydrogen bonds. In this sense A90720A is most like the so-called 'small' protein inhibitors of serine proteinases. These inhibitors are proteins of 29 to 180 residues that have a hydrophobic core and an exposed binding loop held in the 'canonical' binding conformation [12–14,18]. The Ahp residue thus has an essential role in binding; its transannular hydrogen bonds both determine the binding conformation of the inhibitor and prevent its dissociation.

Significance

Cells degrade proteins for many reasons: to remove unwanted proteins, to catabolize proteins for energy, to prepare peptide antigens, to degrade regulatory peptides, and to regulate the activity of enzymes. The activity of proteinases must be controlled, and proteinase inhibitors are essential for the normal function of the body. Regulation of proteinases is an important therapeutic approach to a variety of diseases, and serine proteinases have become particularly important as therapeutic targets. For example, an inhibitor of human leukocyte elastase, which normally degrades cellular debris in the lungs, could be useful in treating emphysema. Thrombin, a trypsin-like serine proteinase, converts fibrinogen to fibrin; thus, thrombin inhibitors could be useful antithrombotic drugs for heart attack and strokes. Finding the ideal proteinase inhibitor has not proved easy. Endogenous proteinase inhibitors are relatively large proteins with excellent specificity and potency, but they are metabolically unstable and potentially immunogenic. Small molecule proteinase inhibitors rarely exhibit the desired selectivity. Screening natural products is an effective way to discover small, yet specific, inhibitors with novel structures, which may act as leads for drug development, as well as providing insight into the mechanisms of proteinase inhibition. A90702A was found in a screen for novel thrombin inhibitors. Here we have determined the high-resolution structure of A90702A bound to trypsin. This structure has allowed us to determine the absolute stereochemistry of several parts of the molecule, and to show that this inhibitor binds tightly to trypsin in a substrate-like manner. The method that it uses to inhibit proteinase activity is similar to that of the 'small' proteinase inhibitors (which are 5–20 times larger than A90702A). Hydrogen bonds involving Ahp, an amino acid not found in proteins, cause A90702A to take up a conformation that is appropriate for binding; the 'small' proteinase inhibitors have a similar loop, but use a large protein core to dictate its conformation.

The determination of the structure of a natural product is normally performed in isolation, separately from studies intended to determine how the molecule has its effect. In the past, important advances in natural products chemistry have resulted from new strategies that allow the combination of previously separated analytical steps. For example, until quite recently, a natural products chemist would isolate compounds from a selected organism and submit pure compounds for biological testing; today, virtually all laboratories use bioassays to guide their purification strategy, combining the isolation step with that of biological testing. Here we have combined the step of structure determination with an experiment that has revealed how the molecule operates in a biological context. As is often the case in science, this initial study resulted from serendipity and desperation. But with the profusion of receptor- and enzyme-based assays to guide the isolation of natural products, the stage is set for many more such studies.

Materials and methods

Crystallization of the complex

Lyophilized bovine trypsin (Sigma, lot 62F-0765) was dissolved in 50.0 mM Tris, pH 8.00 buffer to a concentration of approximately 12 mg ml^{-1} and incubated overnight at 4 °C with a 10-fold molar excess of A90720A in methanol. Droplets (10 µl) containing equal volumes of trypsin–A90720A solution and reservoir solution (1.6 M ammonium sulfate, 0.1 M sodium chloride, 0.1 M HEPES buffer, pH 7.5) were equilibrated against reservoir solution using the hanging drop method. Single crystals appeared within two days and grew to their maximum size of 0.3 x 0.3 x 0.3 mm in a week. Crystals belong to space group $P2_12_12_1$ with a = 63.66, b = 67.53 and c = 69.20 Å and contain one trypsin-A90720A complex in the asymmetric unit with a solvent content of approximately 60 % [19].

Data collection

Because of instability, crystals were mounted in siliconized glass capillaries directly from the drop. X-ray data were collected on a SDMS Mark II area detector coupled to an Rigaku RU2000 rotating Cu-anode X-ray generator. The intensity data were processed by SCALEPACK [20] resulting in an R_{sym} of 5.1 % for 21,586 unique reflections (94 % completeness)

Solution and refinement

The structure was isomorphous with a previously reported bovine β-trypsin structure (PDB entry 1TLD) and this model, stripped of solvent molecules and with uniform thermal and occupancy parameters, was used for the initial X-PLOR refinement [21]. Initially rigid body refinement lowered the R-factor to 30.1 % for the 10–2.8 Å resolution data with $|F_o| > 2\sigma$. Simulated annealing with slow cooling [22] and conjugate gradient positional refinement was used with 8.0–2.4 Å, 2σ data reduced the R-factor to 22.4 %. The $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ maps at 2.4 Å were used manually to adjust the model using the CHAIN [23] graphics program. A90720A was modeled into the $|F_o|-|F_c|$ map. The data were gradually extended to 1.9 Å resolution with periodic manual refitting between refinement cycles. At the later stages of refinement, water molecules were added to fit 3σ electron density in the $|F_o|-F_c$ map. Restrained, individual temperature values were refined for each residue. Average B-values were 11.8, 12.8, and 31.5 $Å^2$ for trypsin, A90720A, and solvent, respectively. Refinement used the X-PLOR slow-cooling simulated annealing [22] and least-squares energy refinement

technique. The final model contains the protein, A90720A, 168 water molecules, and a calcium ion. The final model has an R-factor of 16.2 % for all 2σ data in the 8.0–1.9 Å resolution range and root-mean-square deviations from ideality for bond lengths and bond angles of 0.01 Å and 2.2°, respectively. Coordinates have been deposited with the Brookhaven Protein Data Bank.

Acknowledgements: This work was partially supported by NIH grant CA24487 (J.C.) and NIH Training Grant GM08267 in Molecular Biophysics to A.Y.L. Richard Gillilan of Cornell's Theory Center helped in preparing illustrations. Support for the blue-green algal programs of R.E. Moore and G.M. Patterson of the University of Hawaii (NSF CHE-9024748) is gratefully acknowledged.

References

1. Smith, G.F. (1980). The mechanism of fibrin–polymer formation in solution. *Biochem. J.* **185**, 1–11.
2. Harker, L.A. (1994). Strategies for inhibiting the effects of thrombin. *Blood Coagulation Fibrinolysis* **S1**, S44–58.
3. Johnson, P.H. (1994). Hirudin: clinical potential of a thrombin inhibitor. *Annu. Rev. Med.* **45**, 165–177.
4. Gold, H.K. (1990). Conjunctive antithrombotic and thrombolytic therapy for coronary occlusion. *N. Engl. J. Med.* **323**, 1483–1485.
5. Smitka, T.A., *et al.*, & Patterson, G.M.L. (1992). Ambiguine isonitriles, fungicidal hapalindole-type alkaloids from three genera of blue–green algae belonging to the stigonemataceae. *J. Org. Chem.* **57**, 857–861.
6. Bonjouklian, R., *et al.*, & Patterson, G.M.L. (1994). A90720A, A serine protease inhibitor isolated from a terrestrial blue–green alga *Microchaete loktakensis*. *J. Org. Chem.* **59**, in press
7. Pettit, G.R., *et al.*, & Kizu, H. (1989). Isolation and structure of the cytostatic depsipeptide dolastatin 13 from the sea hare *Dolabella auricularia*. *J. Am. Chem. Soc.* **111**, 5015–5017.
8. Harada, K., Mayumi, T., Shimada, T., Suzuki, M., Kondo, F. & Watanabe, M. (1993). Occurrence of four depsipeptides, aeruginopeptins, together with microcystins from toxic cyanobacteria. *Tetrahedron Lett.* **34**, 6091–6094.
9. Martin, C., Oberer, L., Ino, T., König, W.A., Busch, M. & Weckesser, J. (1993). Cyanopeptolins, new depsipeptides from the cyanobacterium *Microcystis sp.* PCC 7806. *J. Antibiot.* (Tokyo) **46**, 1550–1556.
10. Tsukamoto, S., Painuly, P., Young, K.A., Yang, X. & Shimizu, Y. (1993). Microcystolide A: a novel cell-differentiation-promoting depsipeptide from *Microcystis aeruginosa* NO-15-1840. *J. Am. Chem. Soc.* **115**, 11046–11047.
11. Okino, T., Murakami, M., Haraguchi, R., Munekata, H., Matsuda, H. & Yamaguchi, K. (1993). Micropeptins A and B, plasmin and trypsin inhibitors from the blue-green Alga *Microcystis aeruginosa*. *Tetrahedron Lett.* **34**, 8131–8134.
12. Read, R.J. & James, M.N.G. (1986). Introduction to the protein inhibitors: X-ray crystallography. In *Proteinase Inhibitors*. (Barrett, A.J. & Salvesen, G., eds), pp. 301–336, Elsevier, Amsterdam.
13. Bode, W. & Huber, R. (1991). Ligand binding: proteinase–protein inhibitor interactions. *Curr. Opin. Struct. Biol.* **1**, 45–52.
14. Bode, W. & Huber, R. (1992). Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.* **204**, 433–451.
15. Lee, A.Y., Hagihara, M., Karmacharya, R., Albers, M.W., Schreiber, S.L. & Clardy, J. (1993). Atomic structure of the trypsin–cyclotheonamide A complex: lessons for the design of serine protease inhibitors. *J. Am. Chem. Soc.* **115**, 12619–12620.
16. Maryanoff, B.E., *et al.*, & Fusetani, N. (1993). Molecular basis for the inhibition of human alpha-thrombin by the macrocyclic peptide cyclotheonamide A. *Proc. Natl. Acad. Sci. USA* **90**, 8048–8052.
17. Lewis, S.D., Ng, A.S., Baldwin, J.J., Fusetani, N., Naylor, A.M. & Shafer, J.A. (1993). Inhibition of thrombin and other trypsin-like serine proteinases by cyclotheonamide A. *Thromb. Res.* **70**, 173–190.
18. Laskowski, M.J. & Kato, I. (1980). Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**, 593–626.
19. Matthews, B.W. (1968). Solvent content of crystals. *J. Mol. Biol.* **33**, 491–497.
20. Otwinski, Z. (1992). *The SCALEPACK Manual.* Howard Hughes Medical Institute, Yale University.
21. Brünger, A.T. (1992). *X-PLOR Manual, version 3.1.* Yale University Press, New Haven.
22. Brünger, A.T., Krukowski, A. & Erickson, J. (1990). Slow-cooling protocols for crystallographic refinement. *Acta Crystallogr. A.* **46**, 585–593.
23. Sack, J.S. (1988). CHAIN — A crystallographic modelling program. *J. Mol. Graph.* **6**, 224–225.
24. Bacon, D.J. & Anderson, W.F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219–220

Received: **15 Aug 1994**; revisions requested: **24 Aug 1994**; revisions received: **6 Sep 1994**. Accepted: **9 Sep 1994**