Protein Science (2000), *9*:1423–1427. Cambridge University Press. Printed in the USA.

ACCELERATED COMMUNICATION

Crystal structure of viral serpin crmA provides insights into its mechanism of cysteine proteinase inhibition

MILJAN SIMONOVIC,[1] PETER G.W. GETTINS,[1] AND KARL VOLZ[2]

[1]Department of Biochemistry and Molecular Biology, College of Medicine, University of Illinois at Chicago, Chicago, Illinois 60612-7334
[2]Department of Microbiology and Immunology, College of Medicine, University of Illinois at Chicago, Chicago, Illinois 60612-7334

(RECEIVED June 6, 2000; FINAL REVISION June 8, 2000; ACCEPTED June 8, 2000)

Abstract

CrmA is an unusual viral serpin that inhibits both cysteine and serine proteinases involved in the regulation of host inflammatory and apoptosis processes. It differs from other members of the serpin superfamily by having a reactive center loop that is one residue shorter, and by its apparent inability to form SDS-stable covalent complexes with cysteine proteinases. To obtain insight into the inhibitory mechanism of crmA, we determined the crystal structure of reactive center loop-cleaved crmA to 2.9 Å resolution. The structure, which is the first of a viral serpin, suggests that crmA can inhibit cysteine proteinases by a mechanism analogous to that used by other serpins against serine proteinases. However, one striking difference from other serpins, which may be significant for in vivo function, is an additional highly charged antiparallel strand for β sheet A, whose sequence and length are unique to crmA.

Keywords: apoptosis; crmA; crystal structure; cysteine proteinase; inhibition; serpin; viral; X-ray

Interest in crmA first arose from it being identified as a cowpox viral inhibitor of inflammation, mediated through the specific inhibition of interleukin 1β-converting enzyme (ICE), a cysteine proteinase of the caspase family (Ray et al., 1992). It has subsequently been shown that crmA can inhibit other caspases involved in both host inflammatory and apoptotic processes (Gagliardini et al., 1994; Garcia-Calvo et al., 1998). Such inhibition may be employed by the virus to overcome host defense mechanisms. CrmA is, however, a serpin, whose members are best known as inhibitors of serine proteinases (Gettins et al., 1996). The question, therefore, immediately arises as to how crmA might inhibit proteinases of a different mechanistic class. It is possible that crmA might employ a conformational change-based mechanism analogous to that of serpin inhibitors of serine proteinases (Stratikos & Gettins, 1999), but with formation of a thiol ester rather than an oxygen ester as the kinetically trapped intermediate. However, the failure to detect the SDS-stable complexes between crmA and cysteine proteinases that are the hallmark of serpin inhibition of serine proteinases, together with the shorter reactive center loop of crmA compared with most other serpins, raise questions whether such an inhibition mechanism is structurally possible. To address these questions, we have determined the X-ray crystal structure of a reactive center loop-cleaved form of crmA to 2.9 Å resolution and made comparisons between this and cleaved forms of other serpins that inhibit serine proteinases.

Reprint requests to: Karl Volz, University of Illinois at Chicago, College of Medicine, Department of Microbiology and Immunology, 835 S. Wolcott St., Rm. E-708, Chicago, Illinois 60612-7334; e-mail: kvolz@uic.edu.

Results and discussion

The overall X-ray structure of cleaved CrmA

The X-ray structure of reactive center loop-cleaved crmA (cleaved at the P4–P3 peptide bond) is shown in Figure 1. With the exception of the mobile N-terminal part of chain B (residues 301–309) and 20 mainly solvent accessible charged side chains (K12, K48, E49, D51, K52, K54, D55, I57, D83, N84, E117, D122, E144, S147, E157, K220, D326, K328, S337, and T340), all of the molecule is well ordered in the final crystal structure (R_{cryst} 22.4% and R_{free} 28.8%, Table 1). Comparison with known structures of other cleaved serpins (α_1-proteinase inhibitor (α_1PI), PAI-1, leukocyte elastase inhibitor, and α_1-antichymotrypsin) shows a high level of structural conservation despite the relatively low sequence identity among these family members (20–35% for these serpins). Root-mean-square deviation (RMSD) between cleaved crmA and cleaved α_1PI was found to be 1.9 Å for 292 α-carbons. Forty-two out of 51 mainly hydrophobic residues, which have been previously shown to be conserved throughout the serpin superfamily (Huber & Carrell, 1989), are present in crmA and perform the same role in structural stabilization. Although sequence homology is low in

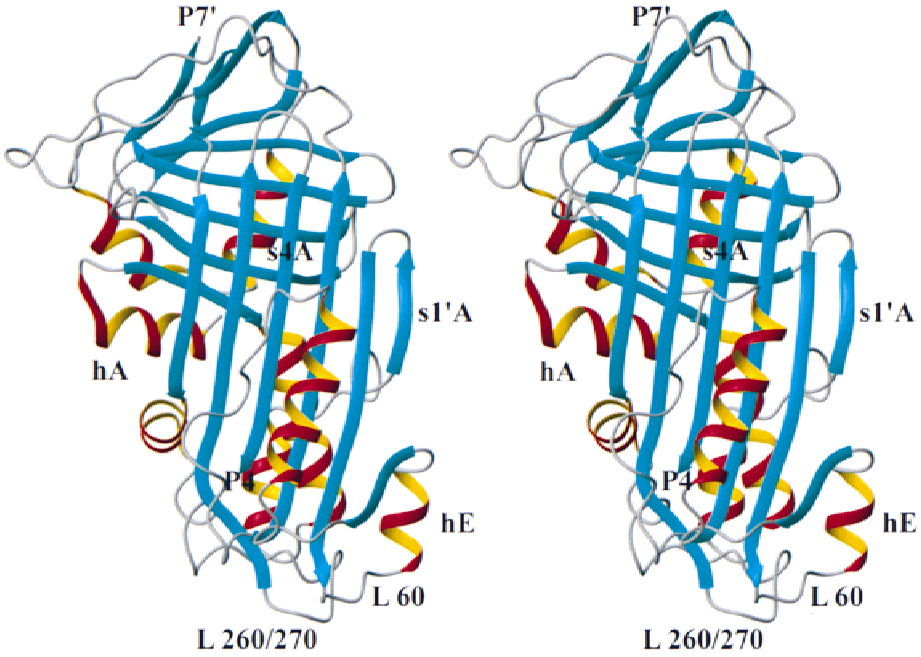

Fig. 1. Stereoview ribbon representation of the cysteine-free mutant of crmA in the cleaved form. β-Sheets cyan and α-helices magenta and yellow. Strands and helices are labeled as in the text. Loops 60, 260, and 270 are marked as L60, L260, and L270, respectively (produced using MOLMOL) (Koradi et al., 1996).

many of the areas of essential secondary structure, such as the strands of β-sheet A, structural alignment shows almost exact superpositioning (Fig. 2A). A striking feature of the structure of crmA is that the cleaved reactive center loop is completely inserted in the body of the molecule as strand s4A, as with other cleaved inhibitory serpins. The cleavage site in our structure is between P4 (L300) and P3 (V301) (equivalent to the P5–P4 bond in other serpins), rather than between P1 and P1′ residues. The P4 residue is well ordered and is clearly the last residue in s4A. Residues V301–T309, which are on the C-terminal side of the cleavage and which should remain at the "top" of the serpin, are disordered and not visible.

Unique structural features of crmA

There are several regions where the crmA molecule deviates from other known structures of cleaved inhibitory serpins. Alignment of the primary structure of crmA with those of various vertebrate serpins reveals large differences in the structurally conserved N-terminal helix A (hA) (Fig. 2A). CrmA has the shortest hA, missing up to four α-helical turns. Additionally, helix C (hC) is shorter in crmA by one helical turn.

Perhaps the most striking difference in length, sequence, and resulting structural features appears immediately after hC (Fig. 2A). In all other serpins, this region (residues V46 to D56 in crmA) is both much longer and contains a well-defined loop-helix-loop motif. In crmA, however, helix D (hD) and the two loops that connect hD with the preceding hC and subsequent strand s2A of the central β-sheet A are entirely absent. Instead, crmA has a short loop (residues 43–52) and a novel additional antiparallel β-strand of β-sheet A (residues 53–57), which we designate as s1′A. The new strand has characteristic antiparallel hydrogen bonding with s2A. This

Table 1. *Crystallographic statistics*

Diffraction data	
D_{min} (Å)	2.90
Total reflections	318,152
Unique reflections	9,406
Completeness (%)[a]	99.9 (91.1)
I/σ[a]	15.5 (4.9)
Multiplicity[a]	5.8 (5.7)
R_{merge} (%)[a,b]	11.2 (42.0)
Refinement	
Data range (Å)	20.00–2.90
Reflections ($F > 0$)	9,238
Completeness (%)	97.9%
Reflections in R_{free} set	968
Number of protein atoms	2,546
Number of solvent atoms	18
RMSDS	
Bond lengths (Å)	0.010
Bond angles (°)	1.25
B-factor for main-chain atoms (Å^2)	1.31
B-factors for side-chain atoms (Å^2)	1.83
Mean *B*-factor (Å^2)	35.8
R_{cryst} (%)[c]	22.4
R_{free} (%)[d]	28.8

[a]The statistics for the highest resolution shell are shown in parentheses.
[b]$R_{merge} = \sum\sum|I(h)_i - \langle I(h)\rangle|/\sum\langle I(h)\rangle$; where $I(h)$ is the observed intensity of the i^{th} measurement of reflection h, and $\langle I(h)\rangle$ the mean intensity of reflection h; calculated after loading and scaling.
[c]$R_{cryst} = \sum[|F_o| - |F_c|]/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively, and the summations are over all unique reflections.
[d]R_{free} is calculated as for R_{cryst} except the summation is over a test set of 10% of unique reflections omitted from the refinement.

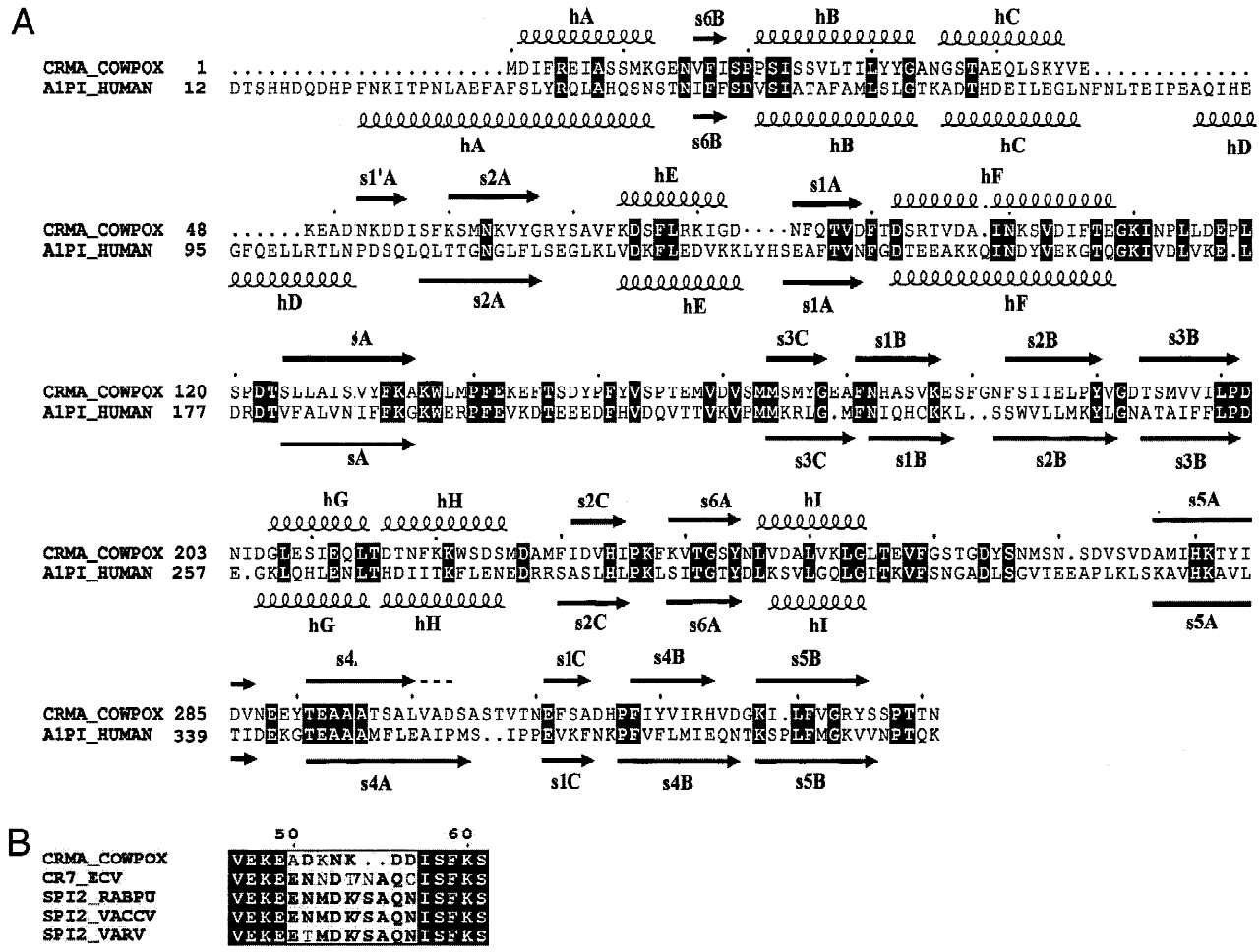

Fig. 2. A: Sequence alignment of cysteine-free crmA and α_1-PI. The upper numbering scheme is for crmA, while the lower one is derived from α_1-PI. Numbers used in the text are for the crmA sequence. Secondary structure elements are indicated for both proteins; arrows represent β strands, whereas coils represent α-helices. Dotted line in the extension of s4A designates the missing P3–P1 region in crmA, which is predicted to be in β conformation. Conserved residues are highlighted in black. **B:** Excerpt from the sequence alignment of viral serpins of the SPI-2 family (cowpox, ectromelia, rabbitpox, vaccinia, and variola virus) demonstrating the unique sequence length and amino acid composition in crmA in the region of the novel β-strand s1′A. Protein sequences are more than 90% identical in the rest of the molecule. Numbering is for crmA, and identical residues are highlighted in black.

region of crmA is also unique when compared to the very closely related viral serpins of the SPI-2 family. Compared with crmA, these other viral serpins have extremely high sequence identity throughout (>90%), with the sole exception of residues D51 to D56 (crmA numbering). This region in crmA is shorter by three residues and contains a completely unrelated and very highly charged sequence (DKNKDD in crmA vs. NMDKVSAQN, which is conserved in other SPI-2 serpins) (Fig. 2B). The electron density for residues 51 and 52 in our structure is weak, but the rest of the loop and the new s1′A β-strand are interpretable, thereby restricting the location of residues 51 and 52. At this resolution almost all the side chains in the region D51–D56 are disordered, and the backbone atoms have relatively high *B*-factors. Further inspection shows that crmA is more compact and adopts a somewhat flatter surface in this part of the molecule compared to other serpins. This flattening is extended to the region of helix E (hE; residues 72–83 in crmA), which is shorter in crmA by one turn. This is a direct consequence of a four residue deletion in the crmA sequence relative to α_1-PI. Consequently, the loop between hE and strand s1A has a different conformation than in other serpins, which thereby allows preservation of strand s1A.

Charge distribution analysis with the program GRASP (Nicholls et al., 1991) demonstrated that two areas of the surface of crmA are highly acidic. The first area is located around and above strand s1′A (residues E47, E49, D55, D56, E142, E169, and D224), whereas the second one is in the area of β-sheets B and C and hA (D2, E14, E40, D92, D97, D116, E117, D122, E157, E257, and D276). Such a markedly nonrandom charge distribution has no precedent in other known serpin structures, which are characterized by small charge patches distributed more randomly throughout the molecule. These differences, taken together, suggest a functional significance for strand s1′A and the acidic regions.

Significance of the structure for the mechanism of proteinase inhibition

In a key aspect, namely having the cleaved reactive center loop completely inserted into the central β-sheet A forming strand s4A

(Fig. 1), the structure of reactive center loop-cleaved crmA is analogous to that of other cleaved serpins that inhibit serine proteinases (Löbermann et al., 1984; Baumann et al., 1991, 1992; Mourey et al., 1993). This clearly establishes that crmA, like other inhibitory serpins, exists in a metastable state that can spontaneously convert to a much more stable state through insertion of the reactive center loop into β-sheet A. However, crmA has another structural feature that distinguishes it from most other inhibitory serpins and that is relevant to the mechanism of inhibition, namely that the reactive center loop of crmA is one residue shorter than in most serpins. Because it has been suggested that the length of the reactive center loop is of critical significance for serpin functioning (Wright & Scarsdale, 1995), we have examined the consequences of the shorter reactive center loop for the inhibitory mechanism of crmA by comparing the position of the P1 residue in crmA relative to the body of the serpin with that of P1 in other serpins. Because our crmA was cleaved at P4–P3 rather than P1–P1′, we modeled the P3–P1 residues (equivalent to P4–P2 in other serpins) according to their average positions in cleaved α_1-PI, PAI-1, leukocyte elastase inhibitor, and α_1-antichymotrypsin. Our measurements show that the P1 residue in crmA, in relation to the strands of β-sheet A, is shifted upward by more than 2 Å compared to the P1 position in other serpins (Fig. 3A). However, the 260s and 270s loops adopt an altered conformation, also being raised upward and inward (Fig. 3A). As a result, the position of the P1 residue in crmA relative to the bottom of the molecule is similar to that of other serpins (Fig. 3B), suggesting that crmA has adapted the conformation of the two loops at the bottom of the molecule (260s and 270s loop) in such a way that the reactive center loop is long enough to have the covalently linked active site cysteine of the proteinase in covalent complexes be at the same position as the active site serine in other inhibitory serpins, and hence, be kinetically trapped by the same kind of structural distortion of the proteinase.

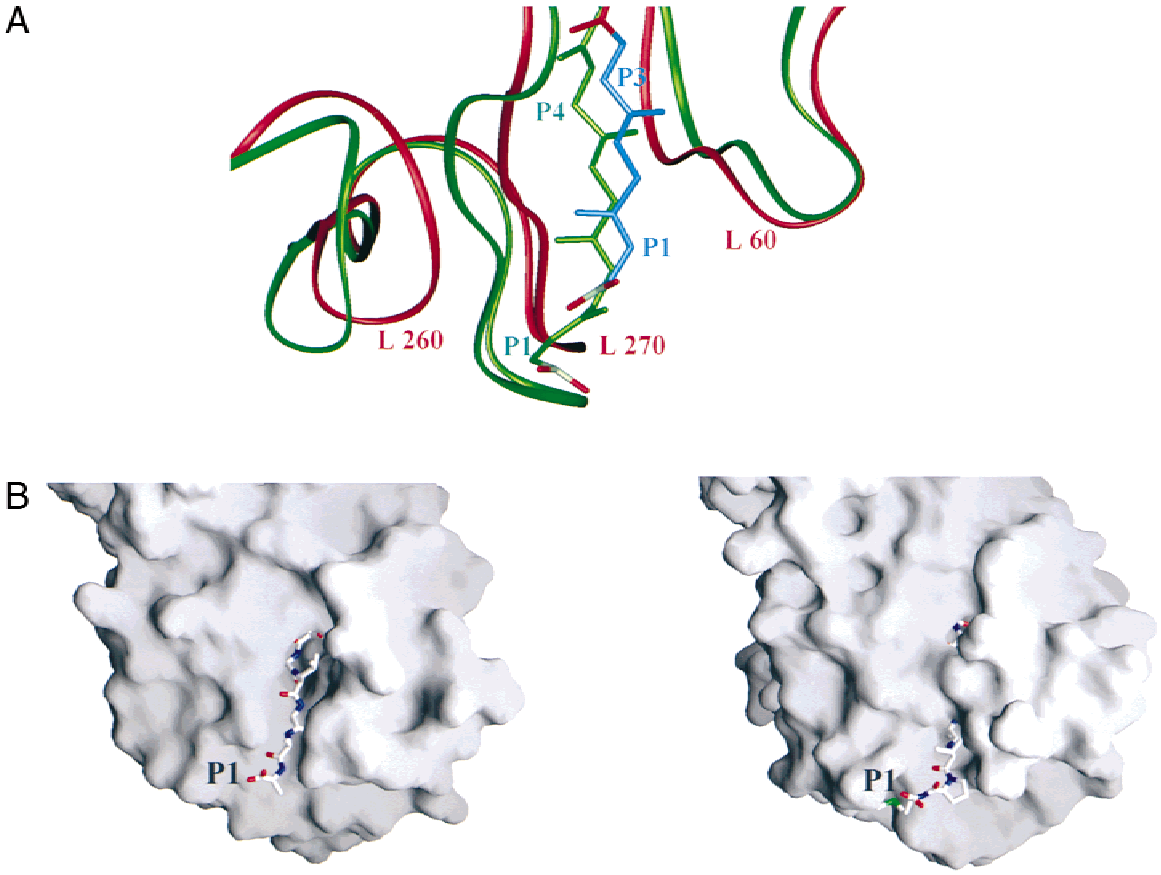

Fig. 3. A: Superposition of cleaved crmA (modeled backbone of P3–P1 cyan sticks, red thin ribbon elsewhere) and cleaved α_1-PI (7API) (backbone of P4–P1 green sticks, green thin ribbon elsewhere) showing the distance between P1 of crmA and P1 of α_1-PI, and changed conformation of 60s, 260s, and 270s loops in cleaved crmA (marked as L60, L260, and L270). C-terminal ends of s4A (P1 residue) are color coded: oxygen atoms in red and carbon atoms in white. P3–P1 residues are modeled as polyalanine in crmA according to their average position in other cleaved serpins, because the cleavage site in our structure is between P4 and P3 (see text). The view is rotated approximately 90° about the vertical axis relative to Figure 1. Figure was prepared using the program QUANTA (1997). **B:** Comparison of the surface representations of cleaved crmA (left) and cleaved α_1-PI (right), showing the height between the C-terminus of s4A and the bottom of the serpin body in each case. The β-strand s4A was excluded from the molecular surface calculation and is represented by sticks and color coded: oxygen in red, nitrogen in blue, sulfur in green, and carbon in white. P3–P1 residues are modeled as polyalanine in crmA. The view is rotated ~90° about the vertical axis relative to Figure 1. Figure was prepared using the program GRASP (Nicholls et al., 1991).

Materials and methods

Protein expression and purification

To prevent disulfide-mediated aggregation, we prepared a cysteine-free variant of crmA, which was found to retain normal inhibitory activity against ICE (data not shown). This crmA variant was expressed as inclusion bodies in *Escherichia coli* at 37 °C using the pQE-60 expression system (Qiagen, Hilden, Germany). Inclusion bodies were isolated, and the protein was refolded and purified as previously described for α_1PI (Kwon et al., 1995). The activity was assayed by inhibition of the activity of ICE.

Crystallization

Although native cysteine-free crmA was used, the crystals obtained were of reactive center loop-cleaved crmA, presumably from contamination with a bacterially derived proteinase. Cleavage was shown to be between residues P4 (L300) and P3 (V301) by dissolution of the crystals followed by N-terminal sequencing and electrospray mass spectroscopy. The latter gave two peaks corresponding to residues 1–300 (predicted 33,437.6 Da; measured 33,438 Da) and residues 301–341 (predicted 4,513 Da; measured 4,513.5 Da).

Crystallization conditions were identified in Hampton crystallization screens (Jancarik & Kim, 1991). Crystals were obtained at 20 °C using the hanging-drop vapor diffusion method, mixing equal volumes of the protein (20 mg/mL) and the well solution (1.6 M Na/K–H_2PO_4, 0.1 M HEPES, pH 7.50). Rice grain-shaped crystals grew within a month to average dimensions of $0.15 \times 0.06 \times 0.04$ mm^3. The crystals belong to space-group $P2_12_12_1$ ($a = 42.67$ Å, $b = 93.15$ Å, and $c = 101.63$ Å) with one molecule per asymmetric unit and approximately 56% solvent content.

Data collection and processing

A single crystal was harvested from the drop, soaked for 15–30 s in mineral oil (Sigma, St. Louis, Missouri) as a cryoprotectant, mounted on a nylon loop (Hampton Research, Laguna Niguel, California), and flash frozen in liquid nitrogen. X-ray diffraction data to 2.9 Å were collected at liquid nitrogen temperature on the BioCARS 14BMC synchrotron beamline source (Advanced Photon Source, Argonne National Laboratories) at a wavelength of 1 Å, on a CCD-Q4 detector. Details of the data collection were described earlier (Simonovic et al., 2000). The data were indexed with DENZO, and scaled and reduced with SCALEPACK (Otwinowski & Minor, 1997).

Structure determination and refinement

Molecular replacement calculations were performed with CNS (Brünger, 1998). A search model was constructed based on the coordinates of cleaved α_1-PI (7API), cleaved PAI-1 (9PAI), cleaved α_1-antichymotrypsin (1AS4), and cleaved leukocyte elastase inhibitor (1HLE), all without hD. The rotation search gave a 6.25σ peak, while the translation search gave a solution with a correlation coefficient of 0.64 (next highest peak 0.46). The oriented search model was then divided into several segments and refined as rigid bodies followed by cycles of simulated annealing. Electron-density maps were interpreted using the program QUANTA (1997). Regions of the model with poor density were deleted and regions of insertions were identified. Ten percent of the data were randomly assigned to an R_{free} test for cross-validation (Brünger, 1998). The model was progressively refined using simulated annealing protocols (5,000 K) with all data, followed by energy minimization and manual inspection and rebuilding. Once R_{cryst} dropped below 35%, bulk solvent and isotropic *B*-factor corrections were introduced. The final stages of refinement, including individual *B*-factor refinement, were performed using PROFFT (Hendrickson, 1985), and the final model statistics were calculated in CNS (Brünger, 1998). Coordinates have been deposited in the Protein Data Bank, with accession code 1C8O.

Acknowledgments

The plasmid pQE-60-crmA was generously provided by Dr. David Ucker. We thank Jill Bayliss for mutagenesis of crmA, Dr. Bob Lee, for N-terminus determination, Dr. R. Van Breemen and coworkers for collection of the electrospray mass spectrometry data, and Dr. K. Moffat and coworkers at the BioCARS beamline, APS, Argonne National Laboratory, for access and beam time. This work was supported by grants from the National Institutes of Health (GM47522 to K.V. and HL49234 and HL64013 to P.G.W.G.) and by funds from the Vice Chancellor for Research, University of Illinois at Chicago. M.S. is the recipient of a University fellowship from the University of Illinois at Chicago.

References

Baumann U, Huber R, Bode W, Grosse D, Lesjak M, Laurell CB. 1991. Crystal structure of cleaved α_1-antichymotrypsin at 2.7 Å resolution and its comparison with other serpins. *J Mol Biol 218*:595–606.

Baumann U, Bode W, Huber R, Travis J, Potempa J. 1992. Crystal structure of cleaved equine leucocyte elastase inhibitor determined at 1.95 Å resolution. *J Mol Biol 226*:1207–1218.

Brünger AT. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D 54*:905–921.

Gagliardini V, Fernandez PA, Lee RKK, Drexler HCA, Rotello RJ, Fishman MC, Yuan J. 1994. Prevention of vertebrate neuronal death by the *crmA* gene. *Science 263*:836–828.

Garcia-Calvo M, Peterson EP, Leiting B, Ruel R, Nicholson DW, Thornberry NA. 1998. Inhibition of human caspases by peptide-based and macromolecular inhibitors. *J Biol Chem 273*:32608–32613.

Gettins PGW, Patston PA, Olson ST. 1996. *Serpins: Structure, function and biology*. Austin, Texas: R.G. Landes Co.

Hendrickson WA. 1985. Stereochemically restrained refinement of macromolecular structures. *Methods Enzymol 115*:252–270.

Huber R, Carrell RW. 1989. Implications of the three-dimensional structure of α_1-antitrypsin for structure and function of serpins. *Biochemistry 28*:8951–8966.

Jancarik J, Kim SH. 1991. Sparse matrix sampling: A screening method for crystallization of proteins. *J Appl Crystallogr 24*:409–411.

Koradi R, Billeter M, Wüthrich K. 1996. MOLMOL: A program for display and analysis of macromolecular structures. *J Mol Graphics 14*:52–55.

Kwon K-S, Lee S, Yu M-H. 1995. Refolding of α_1-antitrypsin expressed as inclusion bodies in *Escherichia coli*: Characterization of aggregation. *Biochim Biophys Acta 1247*:179–184.

Löbermann H, Tokuoka R, Deisenhofer J, Huber R. 1984. Human α_1-proteinase inhibitor. Crystal structure analysis of two crystal modifications and preliminary analysis of the implications for function. *J Mol Biol 177*:731–757.

Mourey L, Samama JP, Delarue M, Petitou M, Choay J, Moras D. 1993. Crystal structure of cleaved bovine antithrombin-III at 3.2-Ångstrom resolution. *J Mol Biol 232*:223–241.

Nicholls A, Sharp KA, Honig B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins 11*:281–296.

Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol 276*:461–472.

QUANTA basic operations. 1997. San Diego: Molecular Simulations Inc.

Ray CA, Black RA, Kronheim SR, Greenstreet TA, Sleath PR, Salvesen GS, Pickup DJ. 1992. Viral inhibition of inflammation. Cowpox virus encodes an inhibitor of the interleukin-1β converting enzyme. *Cell 69*:597–604.

Simonovic M, Gettins PGW, Volz K. 2000. Crystallization and preliminary X-ray diffraction analysis of a recombinant cysteine-free mutant of crmA. *J Acta Crystallogr Sect D.* Forthcoming.

Stratikos E, Gettins PGW. 1999. Formation of the covalent serpin-proteinase complex involves translocation of the proteinase by more than 70Å and full insertion of the reactive center loop into β-sheet A. *Proc Natl Acad Sci USA 96*:4808–4813.

Wright HT, Scarsdale JN. 1995. Structural basis for serpin inhibitor activity. *Proteins 22*:210–225.