J. Biochem. 119, 49-57 (1996)

Crystal Structure of H_2-Proteinase from the Venom of *Trimeresurus flavoviridis*

Takashi Kumasaka,*,†,1 Masaki Yamamoto,† Hideaki Moriyama,* Nobuo Tanaka,* Mamoru Sato,‡ Yukiteru Katsube,‡ Yoshio Yamakawa,§ Tamotsu Omori-Satoh,§ Sadaaki Iwanaga,‖ and Tatzuo Ueki†

Department of Life Science, Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Midori-ku, Yokohama 226; †Biophysics Laboratory, The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-01; ‡Institute for Protein Research, Osaka University, Suita, Osaka 565; §Department of Biochemistry and Cell Biology, National Institute of Health, Shinjuku-ku, Tokyo 162; and ‖ Department of Biology, Faculty of Science, Kyushu University, Higashi-ku, Fukuoka 812

Received for publication, July 21, 1995

The crystal structure of the zinc-protease, H_2-proteinase, isolated from the venom of *Trimeresurus flavoviridis* has been determined. The crystallographic *R* factor is 0.176 for 10,635 reflections with $F_{obs} > 2\sigma(F_{obs})$ in the 8.0 to 2.2 Å resolution range. The enzyme has two domains with a cleft in which a catalytic zinc atom is located. The N-terminal domain is composed of four helices around a central five-stranded β-sheet. The irregularly folded C-terminal domain contains one helix and two disulfide bridges. These two domains are linked by a disulfide bridge. In the zinc environment, the catalytic zinc atom forms a distorted tetrahedral coordination with three histidines and one catalytic water molecule, and the methionine-containing turn is structurally conserved. These are distinctive features of the metzincins, one of the zinc metalloprotease superfamilies. The entire structure shows good agreement with that of two *Crotalus* snake venom proteases, adamalysin II and atrolysin C. The H_2-proteinase, however, contains no structural calcium ions, and differences of disulfide configurations and the coordination of the catalytic water molecule exist as compared with the other two proteases.

Key words: crystal structure, snake venom, zinc-protease.

Metzincins comprise a zinc metalloprotease superfamily named after two distinct characteristics: a zinc-binding consensus sequence HExxHxxGxxH and a structurally conserved methionine-containing turn (the Met-turn) (*1-3*). The metzincins are further divided into four families, the astacin, serratia, matrixin, and reprolysin families. Astacin from *Astacus fluviatilis* was the first metzincin of which the crystal structure was determined (*4*). Subsequently, three-dimensional structures of seven metzincins in all four families were determined (*5-14*).

Metalloproteases contained in Crotalidae and Viperidae snake venoms are classified into the reprolysin family together with mammalian reproductive tract proteins (*1*). The snake venom proteases hydrolyze pericellular basement membrane proteins which are involved in adhesion among capillary endothelial cells, leading to local hemorrhage (*15*). These snake venom proteases are divided into four classes in terms of the number of domains (*16*). The highest molecular mass proteases have four domains, the metalloprotease, disintegrin-like, Cys-rich, and lectin-like domains. The metalloprotease domain located in the N-terminal region of their amino acid sequences is equivalent to an entire sequence of the lowest molecular mass proteases. In the reprolysin family, crystal structures of two protease from rattlesnake venoms, adamalysin II from *Crotalus adamanteus* venom (*5, 6*) and atrolysin C from *Crotalus atrox* venom (*7*), have been determined. These proteases are similar in terms of crystal structure to mammalian collagenases which belong to the matrixin family.

H_2-proteinase [EC 3.4.24.53] isolated from the venom of *Trimeresurus flavoviridis* (Habu snake) (*17*) consists of 201 amino acid residues and three disulfide bridges (*18*). H_2-proteinase cleaves oxidized insulin B chains, fibrinogen A_α-chains and some synthetic peptides (*17, 19*). These peptides, which are more than five amino acid residues in length, are cleaved at a peptide bond adjacent to a Leu or Met residue. The shorter peptides show competitive inhibition which varies in degree according to their length. The proteolytic activity is inhibited by the antihemorrhagic factor in the serum of the Habu snake (*20*) and by metal-chelators such as EDTA.

In the Habu venom, four hemorrhagic proteins, HR1A, HR1B, HR2a, and HR2b were also found (*19, 21-23*). H_2-proteinase and HR2a belong to the lowest molecular mass class. They are similar in terms of amino acid sequence and enzymatic properties, although HR2a only possesses the hemorrhagic activity and requires calcium ions for its activity (*17, 22*).

We report here the molecular structure of H_2-proteinase refined with data up to 2.2 Å resolution. The crystal structures of various snake venom proteins are being

[1] To whom correspondence should be addressed at: Biophysics Laboratory, The Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako, Saitama 351-01. Phone: +48-462-1111 (Ext. 3477), Fax: +48-462-4646

determined in order to examine the structural features which control the substrate recognition by snake venom proteases.

MATERIALS AND METHODS

Purification and Crystallization—H_2-proteinase was purified from Habu venom and crystallized as described previously (*24*). The protein solution was concentrated to 18 mg/ml in 50 mM Tris-HCl buffer (pH 8.0) containing 100 mM NaCl. Crystallization was carried out by the hanging drop vapor diffusion method. A drop formed from equal volumes of the protein solution and reservoir solution consisting of 2.2 M ammonium sulfate and Tris buffer was equilibrated against the reservoir solution at 20°C. Truncated bipyramidal crystals with approximate dimensions of $0.3\times0.3\times0.2$ mm^3, which were suitable for X-ray studies, were obtained in several weeks. The crystals belong to the space group $P4_32_12$ with cell dimensions $a=b=77.8$ Å, $c=82.3$ Å. (The correct enantiomorph was selected based on anomalous dispersion data.)

Data Collection and MIR Phasing—The crystals were harvested into 3.0 M ammonium sulfate solution buffered with 50 mM Tris-HCl (pH 8.0). Heavy-atom derivatives were prepared by soaking crystals in the appropriate solutions (Table I).

Diffraction data of three heavy-atom derivative crystals were collected on a Rigaku RAXIS-IIC (*25*). Ni-filtered CuK_α X-rays ($\lambda=1.5418$ Å) were generated from a rotating anode generator, Rigaku RU-200, operated at 50 kV and 80 mA, and focused by a triangular double mirror system. The oscillation images were indexed and integrated using the program Rigaku-PROCESS (*26*).

Native anomalous diffraction data were collected to obtain anomalous dispersion phasing from the native zinc atom. To determine the wavelength for maximum Bijvoet differences, the X-ray absorption near edge structure (XANES) of the crystal was measured at stations BL-6A and BL-18B at the Photon Factory, Tsukuba. The wavelength $\lambda=1.2825$ Å was selected, where the absorption component of anomalous scattering ($\Delta f''$) takes a maximum value. The anomalous diffraction data and higher resolution data (Native-2) for model refinement at $\lambda=1.07$ Å were obtained using the Weissenberg cameras of the above stations (*27*). The Weissenberg images were processed by the program WEIS (*28*), and scaled and averaged by the programs ROTAVATA and AGROVATA in the CCP4 program suite (*29*). A summary of data collection is presented in Table I.

A difference Patterson map for each derivative was calculated at 4 Å resolution. Possible heavy atom positions approximated from the Harker sections were assessed using cross-phased difference-Fourier maps, where minor sites of each derivative were found. The Bijvoet difference Patterson map calculated using the native anomalous data showed clear peaks derived from the zinc atom. Also, the Bijvoet difference Fourier map calculated using the native anomalous data and the phases from three derivative data set showed only one peak in an asymmetric unit. This result is consistent with the finding that the asymmetric unit contains a monomer of the protein at a optimal V_M value, as determined in the preliminary X-ray study. The absolute configuration of the crystals was determined to be $P4_32_12$ from the anomalous effect. The initial phase set was calculated at 2.5 Å resolution with the overall figure of merit of 0.59 for 5,854 reflections (Table I). The solvent flattened map was helpful in the model building. The phase determination by means of multiple isomorphous replacement was performed using the program package PHASES (*30*).

Model Building and Refinement—The polypeptide chain was traced on minimaps calculated using the initial phases and the phases modified by means of the solvent flattening. In the electron density maps, the zinc atom was located near helical density distributions considered to be an "active site helix." Model building to electron density maps was carried out using the program Turbo-FRODO (*31*). Initially, 84% of the residues was built, excluding the amino- and carboxy-terminal and a loop region (residue 155-175).

The model was refined using the program X-PLOR (*32*) with the force-field parameters derived from the Cambridge database (*33*). The restraints with the initial phases

TABLE I. **Statistics for data collection and phase determination.**

Derivatives	$KAu(CN)_4$	EMTS[a]	$Er_2(SO_4)_3$	Native-anomalous[b]	Native-1	Native-2[b]
Concentration (mM)	1.0	1.0	5.0			
Soak time (days)	2	1	4			
Instrument	RAXIS-IIc	RAXIS-IIc	RAXIS-IIc	PF/1.2825 Å	RAXIS-IIc	PF/1.0700 Å
Resolution limit (Å)	2.5	2.5	2.5	2.5	2.5	2.2
Observations	35,117	44,158	24,144	49,573	39,954	53,845
Unique reflections	7,401	8,353	5,676	7,295	7,722	14,262
Completeness (%)	80.0	90.3	61.4	79.1	83.6	86.2
$R_{merge}(I)$[c]	0.088	0.071	0.086	0.079	0.080	0.069
R_{iso}[d]	0.090	0.088	0.088	–		
R_{Cullis}[e]	0.64	0.65	0.54	–		
R_{Kraut}[f]	0.049	0.040	0.052	0.085		
Number of sites	1	1	2	1		
Phasing power[g]	1.40	1.26	2.03	1.72		
Mean overall figure of merit		0.59				

[a]Ethylmercurithiosalicyclic acid sodium salt. [b]Each of the data sets collected at Photon Factory was scaled and merged using the programs of the CCP4 suite, while the others were done using the program Rigaku-PROCESS and only included amplitudes of $I>1\sigma(I)$. [c]$R_{merge}(I)=\Sigma_h\Sigma_j|I_{hj}/G_j-\langle I\rangle_h|/\Sigma_h\Sigma_{hj}\langle I\rangle_h$, where $\langle I\rangle_h$ is the averaged intensity of reflection, h, derived from the measurements, I_{hj}, and G_j is the inverse scale factor for the jth plate. [d]$R_{iso}=2\Sigma_h|F_{PH}-F_P|/\Sigma_h|F_{PH}+F_P|$. [e]$R_{Cullis}=\Sigma_h|F_{PH}\pm F_P-F_H|/\Sigma_h|F_{PH}-F_P|$ for centric reflections. [f]$R_{Kraut}=\Sigma_h|F_{PH(obs)}-F_{PH(calc)}|/\Sigma_h|F_{PH}|$ for acentric reflections. [g]Phasing power is the root-mean-square (r.m.s.) value of the heavy atom structure factor amplitude divided by the r.m.s. residual lack of closure errors.

were applied until the loop region could be assigned. Water molecules were added to residual peaks on $2F_o - F_c$ or $F_o - F_c$ maps to match the stereochemical conditions. Temperature factors of non-hydrogen atoms were refined under isotropic individual restraints. The accuracy of the model was assessed by calculating the omit maps. The structural parameters for 1,646 non-hydrogen atoms were converged with a crystallographic R factor of 17.6% for 10,635 reflections (Native-2) with $F > 2\sigma(F)$ in the 8.0 to 2.2 Å resolution range (Table II).

RESULTS AND DISCUSSION

Overall Structure—The H_2-proteinase is a flat ellipsoid with dimensions of 50×45×30 Å (Fig. 1a). It is composed of two domains separated by a cleft. Each domain is composed of segments from both the N-terminal 150 residues and the C-terminal 50 residues. The N-terminal domain has a central core of a five-stranded β-sheet with four α helices. The central β-sheet is parallel except for strand IV, which faces the cleft. The secondary structure arrangement of $\beta_I\alpha_A\alpha_B\beta_{II}\alpha_C\beta_{III}\beta_{IV}\beta_V\alpha_D$ is similar to that of other metzincins except for a large insertion of loop II-C and the helix C (Fig. 1b). The C-terminal domain has one helix (α_E) and an irregularly folded region, which is presumably important in substrate recognition. The region contains the Met-turn and is stabilized by two disulfide bridges. The bottom of the cleft is occupied by a catalytic center consisting of the zinc atom and its chelators.

For the model having 198 residues and lacking the two amino-terminal and one carboxy-terminal residues, the Ramachandran plot (*34*) shows that most residues are in the allowed main chain conformation except Cys117 and Asn193 (Fig. 2). The two residues are not taken in a regular secondary structure, and are located in the C-terminal portion, which includes a disulfide, Cys117-Cys196, and contacts with a symmetrically related molecule. The Luzzati plot (*35*) suggests a mean coordinate error of 0.28 Å. The $F_o - F_c$ omit map (Fig. 3) shows good agreement with the refined model. Electron densities corresponding to Asp154-Lys155 and Ser170-Lys172 are relatively weak.

Zinc Environment and Substrate Recognition—The zinc atom is located at the bottom of the cleft and is coordinated in a distorted tetrahedral manner to His142, His146, His152 and the putative catalytic water molecule (Fig. 4). This water molecule is anchored to the catalytic base residue, Glu143. The imidazole rings of the three histidine residues are arranged in the vicinity of Met165 of the Met-turn, which provides a hydrophobic environment and stabilizes the zinc coordination. These hydrogen bond networks in the zinc environment are summarized in Table III. The successive residue to the third histidine chelator is conserved in each metzincin subfamily (*1*). Asn153, conserved in the reprolysin family, is located within possible hydrogen bonding distance from the highly conserved Ser178.

Substrates may interact with the cleft between the two domains in an anti-parallel arrangement against strand IV (*6*). The protease recognizes the P3-P2-P1-P1′-P2′-P3′ sequence of a substrate in the S3, S2, S1, S1′, S2′, S3′ subsites, respectively, and cleaves it at the peptide bond between P1 and P1′. The S1′, S2′, and S3′ subsites (*36*), are located at the deep valley mainly consisting of the protu-

TABLE II. **Refinement statistics (8.0-2.2 Å).**

Crystallographic R factor (%)	17.6
Number of reflection [$F > 2\sigma(F)$]	10,635
Completeness in this range (%)	76.0
r.m.s.d. bond lengthes (Å)	0.014
r.m.s.d. bond angles (°)	1.734
Number of non-hydrogen atoms	1,646
Protein	1,582
Water	63
Zinc	1
Average temperature factors (Å²)	28.2
Protein atoms	27.9
Main-chain atoms	26.0
Water molecules	37.0
Zinc atom	26.9

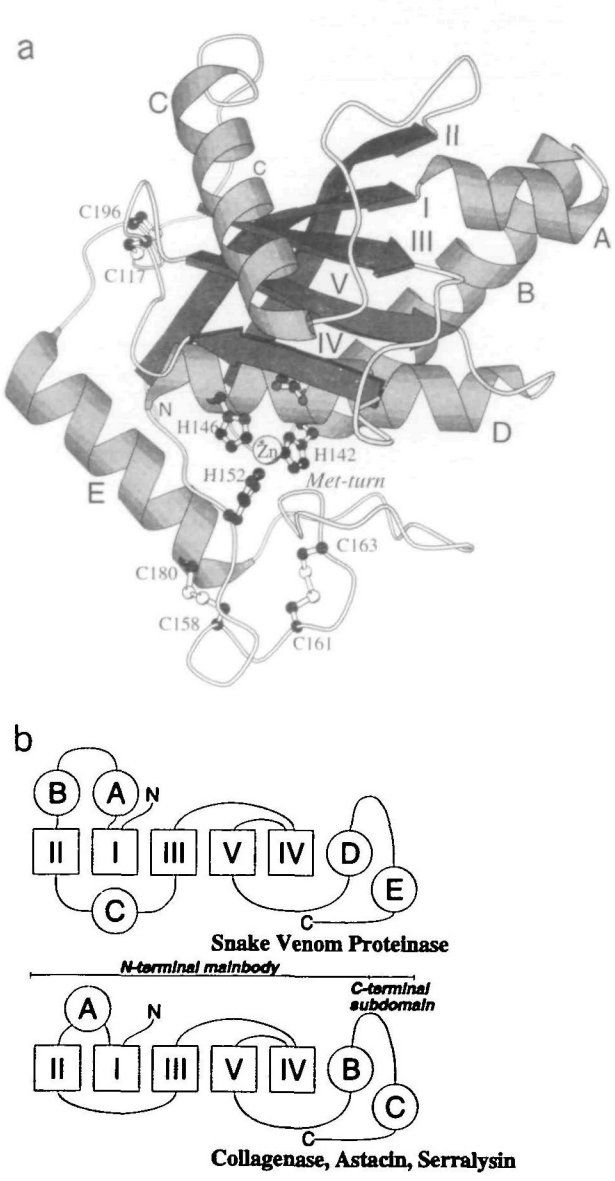

Fig. 1. **(a) Schematic drawing of H_2-proteinase structure.** Zinc atom (solid circle), its chelators and disulfide bridges (ball and stick) are also shown. Figs. 1a, 4, 6a, 7, and 8 were made using the program MOLSCRIPT (*40*). **(b) Metzincins topology.** Topological packing diagrams for snake venom proteinases and other metzincins.

berant loop III-IV and C-terminal loop D-E. The S1, S2, and S3 subsites are relatively flat.

The S1′ subsite, which recognizes the C-terminal side residue of the scissored peptide bond, is located adjacent to the zinc atom and plays the dominant role in substrate recognition. The wall of the S1′ pocket consists of Val138, Ile169, and Ser166, and recognizes hydrophobic bulky residues such as Leu and Met. The enzyme cleaves peptides longer than pentamers, and thus has a complex substrate recognition system involving several sites.

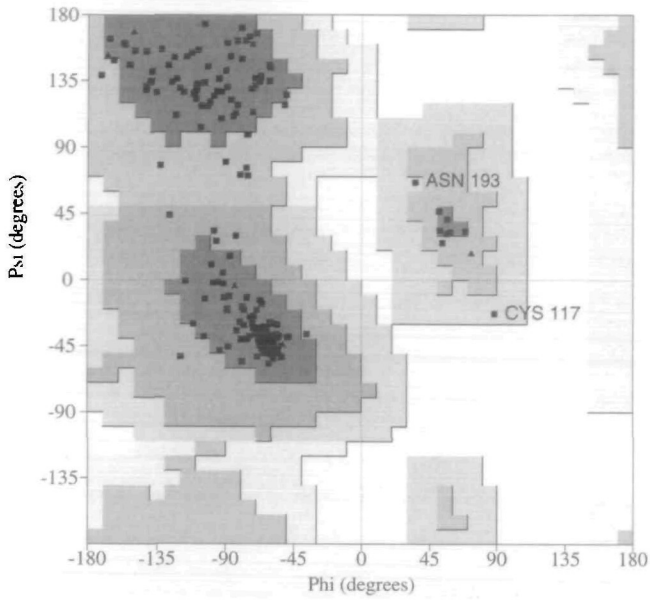

Fig. 2 **Ramachandran plot of main-chain torsion angles.** Non-glycine residues are indicated with a box and glycines are indicated with a triangle. The allowed regions are indicated by solid boundaries The plot was made using PROCHECK (*29*)

Configurations of Disulfide Bridges—Snake venom proteases contain various numbers of disulfide bridges, ranging from two to four. A highly conserved disulfide, Cys117 and Cys196, connects two domains. The variation of the number of disulfide bridges occurs in the C-terminal domain. H_2-proteinase has two disulfide bridges in this domain. One disulfide bridge connects Cys158 on loop D-E and Cys180 in helix E and stabilizes the C-terminal domain The other disulfide between Cys160 and Cys163 serves to direct Met165 to the zinc atom and stabilize its position. In the snake venom proteases possessing three disulfide bridges, various numbers of amino acid residues (three to six) are inserted between the latter cysteine residues (Fig. 5). Since the residues between the cysteines project to a solvent area, these variations do not cause loss of structural integrity.

Comparisons with Adamalysin II—The sequence alignments of H_2-proteinase with adamalysin II and atrolysin C

TABLE III **Geometry of the zinc environment.**

Distances	(Å)	Angles	(°)
ZN-H142 $N^{\epsilon 2}$	2 39	H142 $N^{\epsilon 2}$-ZN-H146 $N^{\epsilon 2}$	88 1
ZN-H146 $N^{\epsilon 2}$	2 26	H142 $N^{\epsilon 2}$-ZN-H152 $N^{\epsilon 2}$	93 0
ZN-H152 $N^{\epsilon 2}$	2 19	H146 $N^{\epsilon 2}$-ZN-H152 $N^{\epsilon 2}$	94 9
ZN-WAT	2 44	H142 $N^{\epsilon 2}$-ZN-WAT	106 4
WAT-E143 $O^{\epsilon 2}$	3.27	H146 $N^{\epsilon 2}$-ZN-WAT	138 8
ZN-M165 $C^{\epsilon 1}$	4 90	H152 $N^{\epsilon 2}$-ZN-WAT	121 7

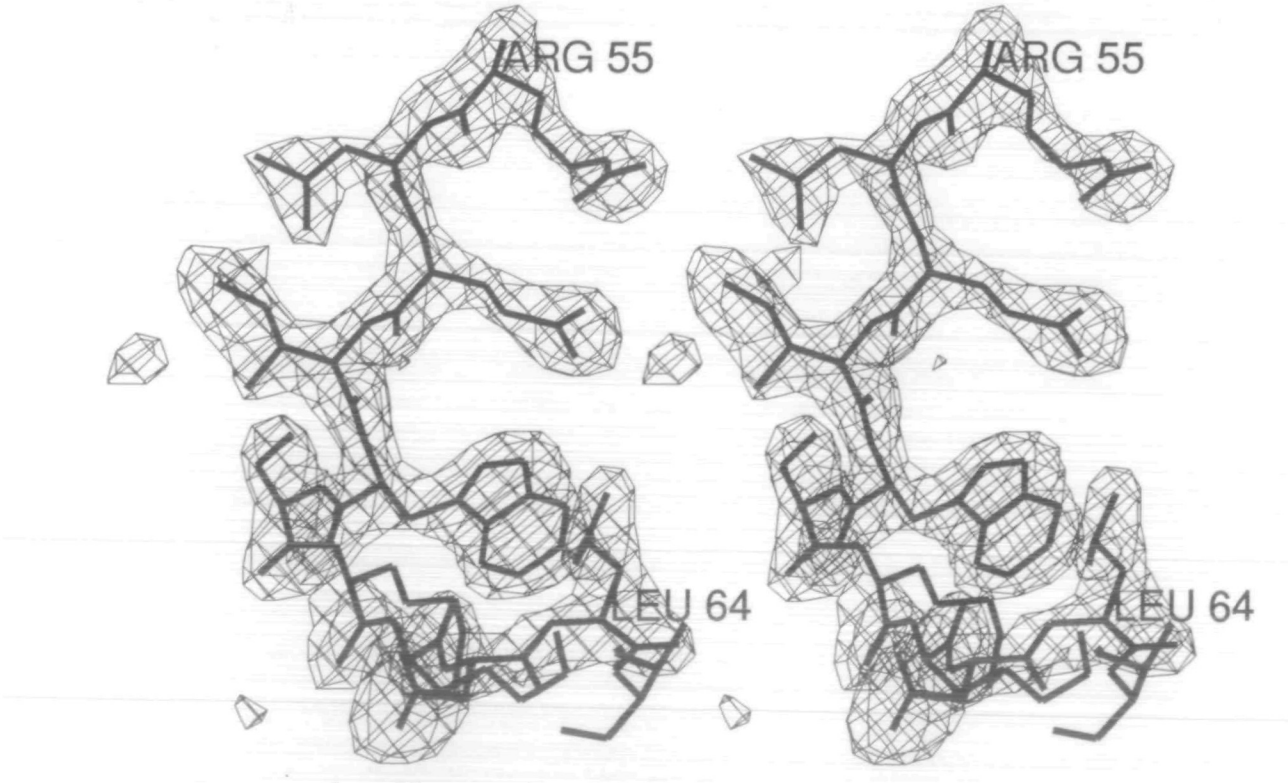

Fig 3. **$F_o - F_c$ omit map for residues 55-64 (RLQIWSKKDL), contoured at 3.0σ.** Refined atomic coordinates are superposed.

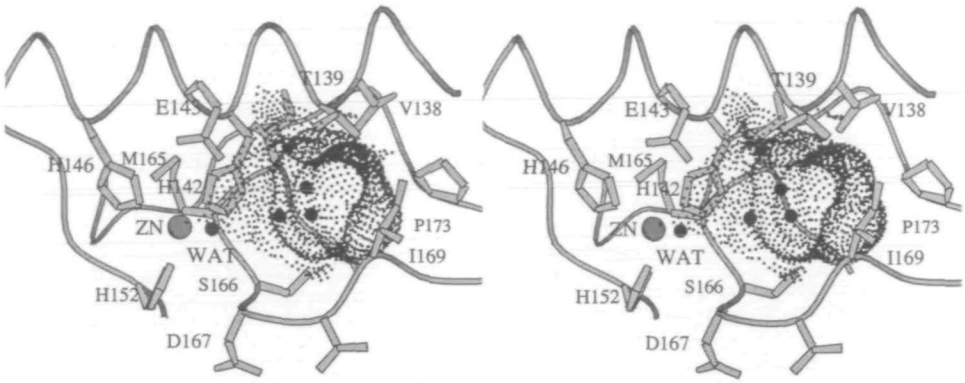

Fig 4. **The zinc environment.** The catalytic zinc atom (ZN) and the catalytic water molecule (WAT) are indicated with gray circles The S1′ pocket, which is located beside the zinc-chelator, is mapped with a Connolly surface obtained by the program Turbo-FRODO. Black circles indicate water molecules which were found in the pocket

```
          140        150        160        170        180
H2   :    MTHELGHNLGMEHDDKDKCKC...EACIMSDVISDKPSKLFSDCSK
HR2a:    MTHEIGHNLGMEHDDKDKCKC...EACIMSAVISDKPSKLFSDCSK
HR1B:    MTHEMGHNLGIPHDGNS.CTC.GGFPCIMSPMISDPPSELFSNCSK
RHOD:    MAHEMGHNLGVRHDGEY.CTCYGSSECIMSSHISDPPSKYFSNCSY
         ----D----->                              <--E--
AtrC:    MAHELGHNLGMEHDGKD.CLR.GASLCIMRPGLTKGRSYEFSDDSM
Ad   :   MAHELGHNLGMEHDGKD.CLR.GASLCIMRPGLTPGRSYEFSDDSM
                           (157)    (164)
```

Fig 5. **Sequence alignment of H_2-proteinase with other snake venom proteases in the loop D-E.** Italicized letters indicate the zinc-binding motif and the Met-turn. Bold letters mean three histidine chelators and the methionine of the Met-turn. The upper four proteins contain two disulfides (thick lines) in this region, while the lower two proteins have only one. The sequences were aligned with the aid of the disulfide configurations. [H2· H_2-proteinase (*18*), HR2a (*23*), HR1B (*21*) from the Habu venom, RHOD rhodostomin (*41*) from the venom of *Calloselasma rhodostoma*, Ad: adamalysin II (*6*), AtrC· atrolysin C (*39*)].

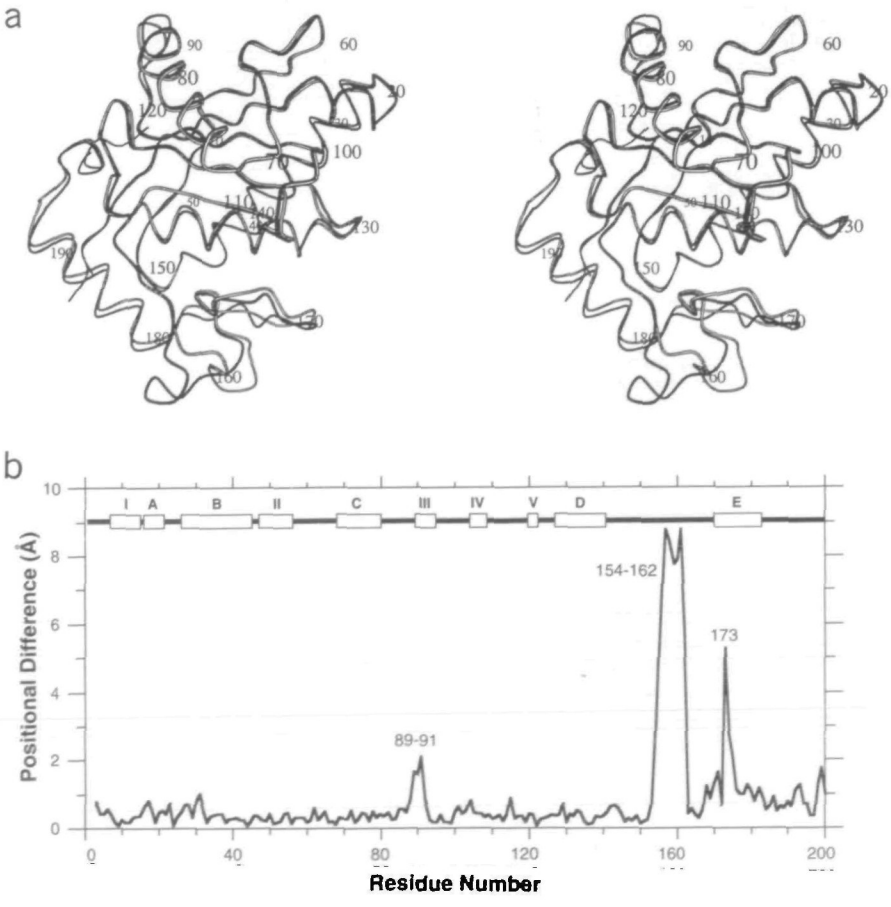

Fig. 6 **(a) Stereo figures of the C_α atoms of H_2-proteinase (solid line) and adamalysin II (hollow line).** The coordinates of the adamalysin II structure were from the Protein Data Bank (Code: 1IAG). Each 10 residues are numbered **(b) Positional differences between corresponding C_α atoms of H_2-proteinase and adamalysin II.** Residue numbers reflect alignment of the sequence of adamalysin II onto that of H_2-proteinase with the gap between Ala162-Cys163. Above lines and boxes show regions of secondary structures.

show 53 and 52% homologies, with one gap at Ala162-Cys163 in the sequence of H_2-proteinase, respectively. The level of sequence homology between adamalysin II and atrolysin C is 80%. Although the overall topology of the H_2-proteinase structure is similar to that of adamalysin II and atrolysin C (Fig. 6a), there are significant differences caused by differences in the configuration of the disulfide bridges in the C-terminal domain (Fig. 5). The positional differences of C_α atom positions are shown in Fig. 6b. The average of the values only in the N-terminal domain (residues 3-152) is 0.50 Å. The differences in phi-psi angles of Arg89 in loop C-III and the positional shift of helix E were caused by differences in the crystal packing, because each of the conformations is reasonable complementally.

Adamalysin II has only one disulfide, Cys157-Cys164, in the C-terminal domain (Fig. 7a). These cysteine residues correspond to Cys158 and Cys163 of H_2-proteinase in the alignment of the amino acid sequences (Fig. 5). In H_2-proteinase, the region from Asp154 to Cys163 is located closer than in adamalysin II to helix E, due to the introduction of two cysteines (Cys160 and Cys180). The largest positional difference of 8.79 Å is between Lys157 of H_2-proteinase and Cys157 of adamalysin II (Fig. 6b).

Residues from Ser170 to Pro173 of H_2-proteinase form the turn of the loop connecting the Met-turn with helix E. Pro173, which is conserved in snake venom proteases possessing three disulfide bridges, distorts a β-conformational loop and faces the bottom of the S1′ pocket. In adamalysin II, this depression is complemented by orienting a Tyr176 side chain toward the inside of the corresponding loop.

The zinc environment in H_2-proteinase is different from that of the adamalysin II molecule in terms of the positions of the catalytic water molecule and the side chain of the catalytic base glutamate residue, Glu143 (Fig. 7b). An O^δ atom of Glu143 in H_2-proteinase is located within possible

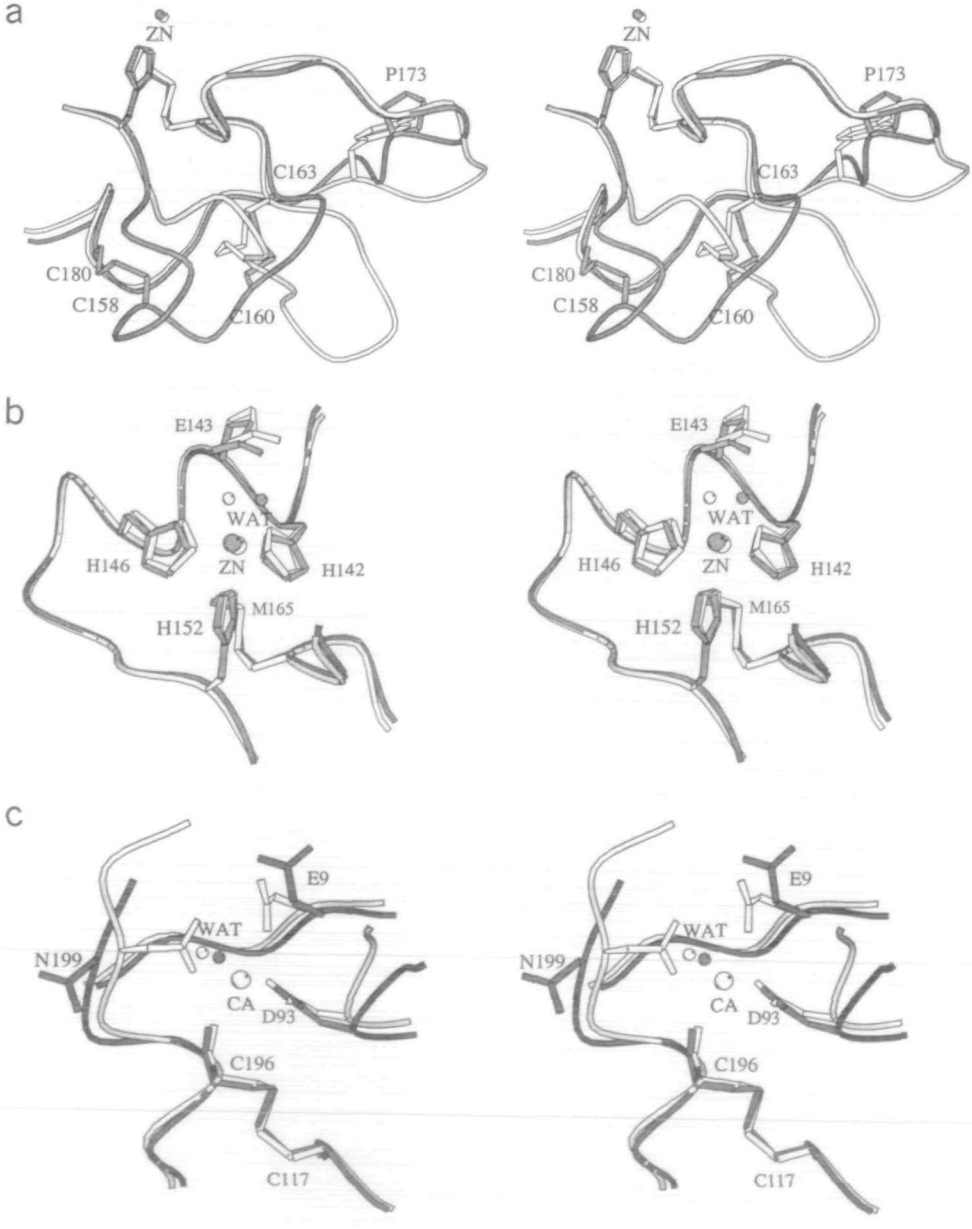

Fig 7. **Stereo figures of superposition of H_2-proteinase and adamalysin II.** (a) Disulfide bridges in the C-terminal domain, (b) zinc coordination, (c) calcium binding site. The calcium ion is only included in adamalysin II. Gray lines and white lines indicate H_2-proteinase and adamalysin II, respectively

hydrogen bonding distance (3.42 Å) from the water molecule, while the glutamate residue of adamalysin II is far from the water molecule (4.11 Å). Most zinc-proteases show optimal activity under neutral or weakly alkaline conditions. The crystals of H_2-proteinase were obtained at pH 8.0, near the optimum pH for the proteolytic activity, while those of adamalysin II were obtained at pH 5.0 (*6*). The pH dependency of the activity may be related to the pK_a of the γ-carboxyl groups of the glutamates. This conformation is required for deprotonating the water molecule and cleaving substrates.

The arrangement of the zinc atom and the three histidine residues corresponds well with that in the adamalysin II molecule, although the water coordination is different and there is a large positional differences of loop D-E located near the zinc environment.

A structural calcium ion in adamalysin II binds tetrahedrally to three residues (Glu9, Asp93, and Asn200 in adamalysin II) and a water molecule (Fig. 7c). Despite the conservation of these residues in H_2-proteinase, no areas representing calcium ions could be found in the electron density map. In H_2-proteinase, Glu9 contacts Ser52, and Asn199 (Asn200 in adamalysin II) is linked to a water molecule and Lys21 of a symmetrically related molecule through hydrogen bonding. Although H_2-proteinase can be readily inactivated by the addition of some divalent cations to the calcium-free solution, HR2a from Habu venom, which also possesses these binding residues, requires calcium ions for its proteolytic activity and is not inhibited by these cations in the presence of calcium ions. These features suggest that structural calcium ions affect the stability of the zinc coordination in HR2a, because the

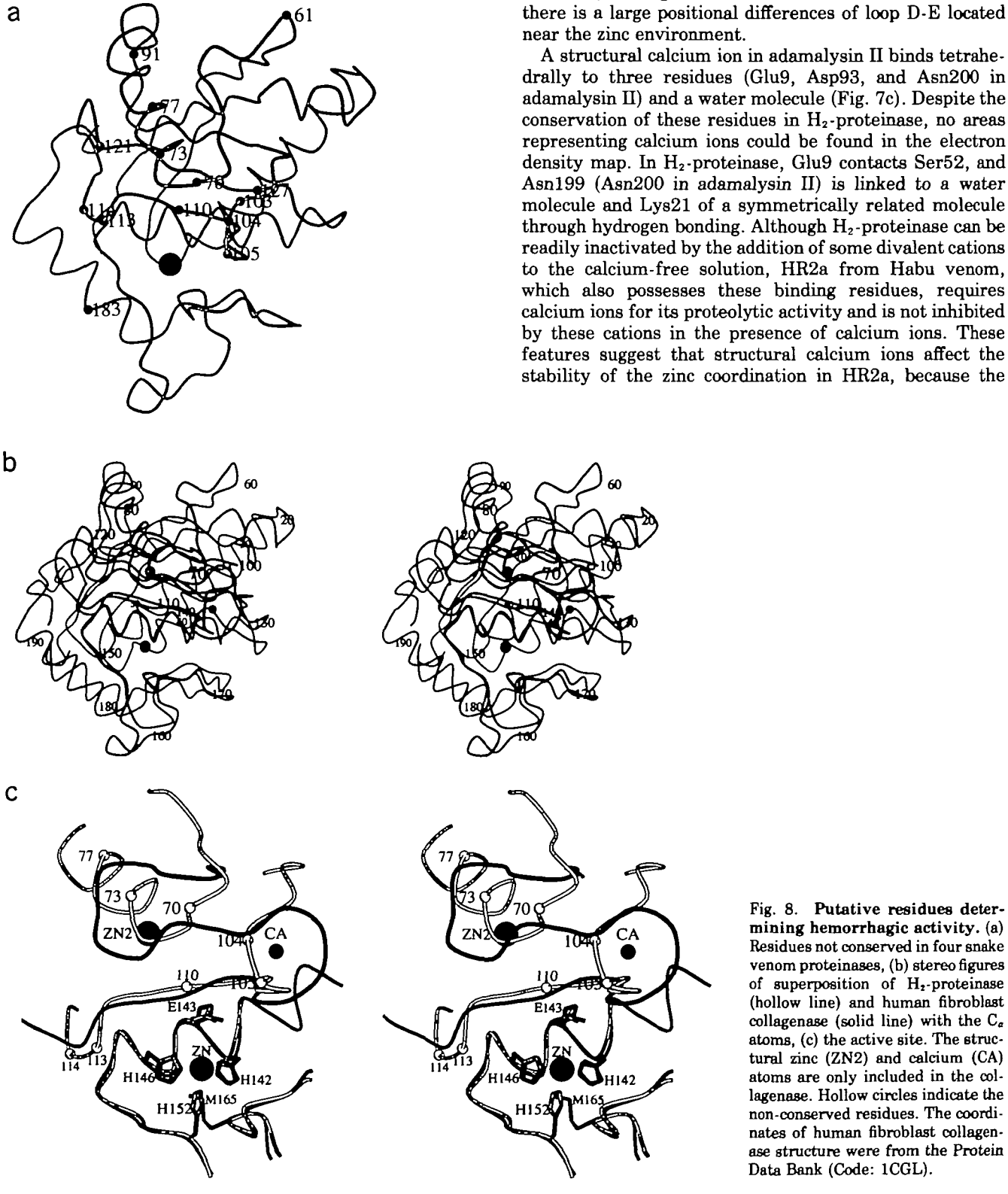

Fig. 8. **Putative residues determining hemorrhagic activity.** (a) Residues not conserved in four snake venom proteinases, (b) stereo figures of superposition of H_2-proteinase (hollow line) and human fibroblast collagenase (solid line) with the C_α atoms, (c) the active site. The structural zinc (ZN2) and calcium (CA) atoms are only included in the collagenase. Hollow circles indicate the non-conserved residues. The coordinates of human fibroblast collagenase structure were from the Protein Data Bank (Code: 1CGL).

substitution of the zinc ion reduces the activity by changing the coordination number around the substituted cation (*37, 38*).

Comparisons of Metzincins—Superpositions of the crystal structure of H_2-proteinase and metzincins whose structure have been determined were calculated with respect to the C_α atoms from Val138 on the active site helix D to the third histidine ligand His152 (the region for the fitting was extended along helix D). All of them show agreement in the zinc environment, especially, human fibroblast collagenase (*8-10*) and human neutrophil collagenase (*11*), belonging to the matrixin family. The astacin and serratia families have a tyrosine residue as a fifth zinc ligand while in the reprolysin and matrixin families it is replaced by a proline residue, which is not a zinc chelator. H_2-Proteinase has a chelatable aspartate residue (Asp167) in that position. However, this side chain in H_2-proteinase is oriented away from the cleft and cannot chelate the zinc atom.

H_2-proteinase (*18*) and adamalysin II (*6*) as non-hemorrhagic proteases were compared with hemorrhagic proteases, HR2a (*22*) and atrolysin C (*39*), in terms of primary structure. Different amino acid residues in each group may take part in the recognition of matrix proteins. Most of these residues lie on the surface of the cleft side of loop II-C (Lys61, Ala70), helix C (Val73, Ser77), loop III-IV (Asn104, Asp105), strand IV (Leu110), and loop IV-V (Lys113, Lys114) (Fig. 8a). Cleavage of the matrix proteins by matrixins or snake venom hemorrhagic proteases might be carried out *via* direct interaction of the cleft of the enzyme with the matrix proteins. Because there are few differences in the specificities for synthetic peptides between snake venom hemorrhagic and non-hemorrhagic proteases, residues on the surface of the enzymes might determine the affinities for the matrix proteins. Superposition of the crystal structure of a matrixin, human fibroblast collagenase (*8, 9*), on that of H_2-proteinase shows positional agreements for residues 70, 73, 77, 104, and 105, although the amino acid sequence of the collagenase is different in the N-terminal domain, with large deletions compared with that of H_2-proteinase (Fig. 8, b and c). In this region of the collagenase, two additional cations stabilize the conformation. This supports the hypothesis that these molecular surface regions are related to the enzyme affinity for matrix proteins. However, these residues may not be dominant in substrate recognition, because the lowest molecular mass proteases show weaker proteolytic activities than the higher molecular mass ones. Also, crystal structures of matrixins were determined only at the catalytic domain. The elucidation of the crystal structure of the whole enzymes is required in order to understand the mechanisms of enzyme affinity for substrates.

We are grateful to Professor S.S. Hasnain, SERC Daresbury Laboratory, and Dr. M. Nakasako, The Institute of Physical and Chemical Research (RIKEN), for helpful discussions. We also thank Professor N. Sakabe and Drs. A. Nakagawa and N. Watanabe for advice on data collection at the Photon Factory, Tsukuba. Data collection at the Photon Factory was performed with the approval of the Photon Factory Program Advisory Committee (Proposal No. 92-G219).

REFERENCES

1. Hooper, N.M. (1994) Families of zinc metalloproteases. *FEBS Lett.* **354**, 1-6
2. Bode, W., Gomis-Rüth, F.X., and Stöcker, W. (1993) Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins.' *FEBS Lett.* **331**, 134-140
3. Stöcker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Rüth, F.-X., McKay, D.B., and Bode, W. (1995) The metzincins —Topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Prot. Sci.* **4**, 823-840
4. Gomis-Rüth, F.X., Stöcker, W., Huber, R., Zwilling, R., and Bode, W. (1994) Refined 1.8 Å X-ray crystal structure of astacin, a zinc-endopeptidase from the crayfish *Astacus astacus* L. *J. Mol. Biol.* **229**, 945-968
5. Gomis-Rüth, F.X., Kress, L.F., and Bode, W. (1993) First structure of a snake venom metalloproteinase: A prototype for matrix metalloproteinases/collagenases. *EMBO J.* **12**, 4151-4157
6. Gomis-Rüth, F.X., Kress, L.F., Kellermann, J., Mayr, I., Lee, X., Huber, R., and Bode, W. (1994) Refined 2.0 Å X-ray crystal structure of the snake venom zinc-endopeptidase adamalysin II. *J. Mol. Biol.* **239**, 513-544
7. Zhang, D., Botos, I., Gomis-Rüth, F.X., Doll, R., Blood, C., Njoroge, F.G., Fox, J.W., Bode, W., and Meyer, E.F. (1994) Structural interaction of natural and synthetic inhibitors with the venom metalloproteinase, atrolysin C (form d). *Proc. Natl. Acad. Sci. USA* **91**, 8447-8451
8. Lovejoy, B., Cleasby, A., Hassell, A.M., Longley, K., Luther, M.A., Weigl, D., McGeehan, G., McElroy, A.B., Drewry, D., Lambert, M.H., and Jordan, S.R. (1994) Structures of the catalytic domain of fibroblast collaganase complexed with an inhibitor. *Science* **263**, 375-377
9. Lovejoy, B., Hassell, A.M., Luther, M.A., Weigl, D., and Jordan, S.R. (1994) Crystal structures of recombinant 19-kDa human fibroblast collaganase complexed to itself. *Biochemistry* **33**, 8207-8217
10. Borkakoti, N., Winkler, F.K., Williams, D.H., D'Arcy, A., Broadhurst, M.J., Brown, P.A., Johnson, W.H., and Murray, E.J. (1994) Structure of the catalytic domain of human fibroblast collagenase complexed with an inhibitor. *Nature Struct. Biol.* **1**, 106-110
11. Stams, T., Spurlino, J.C., Smith, D.L., Wahl, R.C., Ho, T.F., Qoronfleh, M.W., Banks, T.M., and Rubin, B. (1994) Structure of human neutrophil collagenase reveals large S1′ specificity pocket. *Nature Struct. Biol.* **1**, 119-123
12. Gooley, P.R., O'Connell, J.F., Marcy, A.I., Cuca, G.C., Salowe, S.P., Bush, B.L., Hermes, J.D., Esser, C.K., Hagmann, W.K., Spinger, J.P., and Johnson, B.A. (1994) The NMR structure of the inhibited catalytic domain of human stromelysin-1. *Nature Struct. Biol.* **1**, 111-118
13. Baumann, U., Wu, S., Flaherty, K.M., and McKay, D.B. (1993) Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: A two-domain protein with a calcium binding parallel beta roll motif. *EMBO J.* **12**, 3357-3364
14. Baumann, U. (1994) Crystal structure of the 50 kDa metallo protease from *Serratia marcescens*. *J. Mol. Biol.* **242**, 244-251
15. Ohsaka, A., Just, M., and Habermann, E. (1973) Action of snake venom hemorrhagic principles on isolated glomerular basement membrane. *Biochim. Biophys. Acta* **323**, 415-428
16. Iwanaga, S. and Takeya, H. (1993) Structure and function of snake venom metalloproteinase family in *Methods in Protein Sequence Analysis* (Imahori, K. and Sakiyama, F., eds.) pp. 107-115, Plenum Press, New York
17. Takahashi, T. and Ohsaka, A. (1970) Purification and characterization of a proteinase in the venom of *Trimeresurus flavoviridis*. *Biochim. Biophys. Acta* **198**, 293-307
18. Takeya, H., Arakawa, M., Miyata, T., Iwanaga, S., and Omori-Satoh, T. (1989) Primary structure of H_2-proteinase, a non-hemorrhagic metalloproteinase, isolated from the habu snake, *Trimeresurus flavoviridis*. *J. Biochem.* **106**, 151-157
19. Takeya, H., Nishida, S., Nishino, N., Makinose, Y., Omori-Satoh, T., Nikai, T., Sugihara, H., and Iwanaga, S. (1993)

Primary structures of platelet aggregation inhibitors (disintegrins) autoproteolytically released from snake venom hemorrhagic metalloproteinases and new fluorogenic peptide substrates for these enzymes. *J. Biochem.* **113**, 473-483
20. Omori-Satoh, T. (1977) Antihemorrhagic factor as a proteinase inhibitor isolated from the serum of *Trimeresurus flavoviridis*. *Biochim. Biophys. Acta* **495**, 93-98
21. Takeya, H., Oda, K., Miyata, T., Omori-Satoh, T., and Iwanaga, S. (1990) The complete amino acid sequence of the high molecular mass hemorrhagic protein HR1B isolated from the venom of *Trimeresurus flavoviridis*. *J. Biol. Chem.* **265**, 16068-16073
22. Takahashi, T. and Ohsaka, A. (1969) Purification and some properties of two hemorrhagic principles (HR2a and HR2b) in the venom of *Trimeresurus flavoviridis*; complete separation of the principles from proteolytic activity. *Biochim. Biophys. Acta* **207**, 65-75
23. Miyata, T., Takeya, H., Ozeki, Y., Arakawa, M., Tokunaga, F., Iwanaga, S., and Omori-Satoh, T. (1989) Primary structures of hemorrhagic protein, HR2a, isolated from the venom of *Trimeresurus flavoviridis*. *J. Biochem.* **105**, 847-853
24. Kumasaka, T., Takeya, H., Yamamoto, M., Yamakawa, Y., Omori-Satoh, T., Moriyama, H., Tanaka, N., Sato, M., Katsube, Y., and Iwanaga, S. (1995) Crystallization and preliminary X-ray study of H_2-proteinase from the venom of *Trimeresurus flavoviridis*. *J. Biochem.* **177**, 929-930
25. Sato, M., Yamamoto, M., Imada, K., Katsube, Y., Tanaka, N., and Higashi, T. (1991) Development of a high-speed data-collection system with an imaging plate in consideration of large-unit-cell crystals. *J. Appl. Cryst.* **23**, 348-357
26. Higashi, T. (1990) Auto-indexing of oscillation images. *J. Appl. Cryst.* **23**, 253-257
27. Sakabe, N. (1991) X-ray diffraction collection system for modern protein crystallography with a Weissenberg camera and an imaging plate using synchrotron radiation. *Nucl. Instrum. Methods. Phys. Res.* **A303**, 448-463
28. Higashi, T. (1989) The processing of diffraction data taken on a screenless Weissenberg camera for macromolecular crystallography. *J. Appl. Cryst.* **22**, 9-18
29. Collaborative Computational Project, Number 4 (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* **D50**, 760-763
30. Furey, W. and Swaminathan, S. (1990) PHASES—A program package for the processing and analysis of diffraction data from macromolecules. *Am. Crystallogr. Assoc. Annu. Mtg. Program Abst.* **18**, 73
31. Bio-Graphics (1994) Turbo-FRODO (version 5.02), Manual
32. Brünger, A.T. (1992) X-PLOR (version 3.0), Manual
33. Engh, R.A. and Huber, R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst.* **A47**, 392-400
34. Ramakrishnan, C. and Ramachandran, G.N. (1965) Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. *Biophys. J.* **5**, 909-933
35. Luzzati, P.V. (1952) Traitement statistique des erreurs dans la determination des structures cristallines. *Acta. Cryst.* **5**, 802
36. Schechter, I. and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157-162
37. Hakånsson, K., Wehnert, A., and Liljas, A. (1994) X-ray analysis of metal-substituted human carbonic anhydrase II derivatives. *Acta Cryst.* **D50**, 93-100
38. Gomis-Rüth, F.-X., Grams, F., Yiallouros, I., Nar, H., Küsthardt, U., Zwilling, R., Bode, W., and Stöcker, W. (1994) Crystal structures, spectroscopic features, and catalytic properties of cobalt(II), copper(II), nickel(II), and mercury(II) derivatives of the zinc endopeptidase astacin. *J. Biol. Chem.* **269**, 17111-17117
39. Hite, L., Jia, L.-G., Bjanason, J.B., and Fox, J.W. (1994) cDNA sequences for four snake venom metalloproteinases: Structure, classification, and their relationship to mammalian reproductive proteins. *Arch. Biochim. Biophys.* **308**, 182-191
40. Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950
41. Au, L.C., Huang, Y.B., Huang, T.F., Teh, G.W., Lin, H.H., and Choo, K.B. (1991) A common precursor for a putative hemorrhagic protein and rhodostomin, a platelet aggregation inhibitor of the venom of *Calloselasma rhodostoma*: Molecular cloning and sequence analysis. *Biochem. Biophys. Res. Commun.* **181**, 585-593