The EMBO Journal vol.12 no.11 pp.4151-4157, 1993

First structure of a snake venom metalloproteinase: a prototype for matrix metalloproteinases/collagenases

Franz-Xaver Gomis-Rüth[1,2], Lawrence F.Kress[3] and Wolfram Bode[1]

[1]Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany, [2]Institut de Biología Fonamental, Universitat Autónoma de Barcelona, E-08193 Bellaterra, Spain and [3]Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Buffalo, NY 14263, USA

Communicated by R.Huber

Adamalysin II, a 24 kDa zinc endopeptidase from the snake venom of *Crotalus adamanteus*, is a member of a large family of metalloproteinases isolated as small proteinases or proteolytic domains of mosaic haemorrhagic proteins from various snake venoms. Homologous domains have recently been detected in multimodular mammalian reproductive tract proteins. The 2.0 Å crystal structure of adamalysin II reveals an ellipsoidal molecule with a shallow active-site cleft separating a relatively irregularly folded subdomain from the calcium-binding main molecular body composed of a five-stranded β-sheet and four α-helices. The folding of the peptide fragment containing the zinc-binding motif HExxHxxGxxH bears only a distant resemblance to thermolysin, but is identical to that found in astacin, with the three histidines and a water molecule (linked to the glutamic acid) likewise constituting the zinc ligand; adamalysin II lacks a fifth (tyrosine) zinc ligand, however, leaving its zinc ion tetrahedrally co-ordinated. Furthermore, adamalysin II and astacin share an identical active-site basement formed by a common Met-turn. Due to their virtually identical active-site environment and similar folding topology, the snake venom metalloproteinases (hitherto called adamalysins) and the astacins (and presumably also the matrix metalloproteinases/mammalian collagenases and the *Serratia* proteinase-like large bacterial proteinases) might be grouped into a common superfamily with distinct differences from the thermolysin family.

Key words: collagenases/metalloproteinases/snake venom proteinases/X-ray crystal structure/zinc endopeptidases

Introduction

Snake venom contains a number of different pharmacologically active peptides and proteins (Bjarnasson and Fox, 1989; Kini and Evans, 1992). Most of the haemorrhagic proteins of snake venoms are metalloproteinases, with molecular weights ranging from 20 to 100 kDa. The high molecular weight proteinases are multimodular proteins, with a central zinc-containing catalytic domain preceded by an N-terminal putative pro-domain and followed by a disintegrin domain and a cysteine-rich domain (Neeper and Jacobson, 1990; Takeya *et al.*, 1990a, 1992; Hite *et al.*, 1992; Kini and Evans, 1992; Paine *et al.*, 1992). The low molecular weight proteinases consist of the highly conserved 200–204 amino acid residue catalytic domain alone, which often seems to be derived from larger precursors via post-translational proteolytic events. This catalytic domain comprises an extended HisGluXxxXxxHisXxxXxxGlyXxxXxxHis consensus sequence (see Table I), which in analogy to astacin has been implicated in zinc-binding (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993). Both the high and the low molecular weight forms exhibit proteolytic activity towards various matrix proteins. Similar to the matrix metallo proteinases (see Woessner, 1991; Murphy and Docherty, 1992; Birkedal-Hansen *et al.*, 1993) and to the astacins (Titani *et al.*, 1987; Stöcker *et al.*, 1988), however, their hydrolytic activity towards short peptide substrates can be extremely low; only long peptides extending three or more residues beyond the scissile peptide bond are effectively cleaved (Fox *et al.*, 1986). In further accordance with matrix metalloproteinases, the snake venom enzymes seem to cleave peptides preferentially N-terminal to medium sized hydrophobic residues (Fox *et al.*, 1986; Bjarnasson and Fox, 1989). Recently, a few reproductive tract proteins and their respective DNAs have been analysed. These proteins seem to be implicated in sperm maturation [EAP I from rat and monkey (Blobel *et al.*, 1992; Blobel and White, 1992) and cyritestin from mouse testis (U.A.O.Heinlein, EMBL accession no. X644227)] and in sperm–egg fusion [PH-30-α and -β (Perry *et al.*, 1992)], and share similarities in sequence and domain organization with the snake venom metalloproteinases (Blobel and White, 1992; Blobel *et al.*, 1992; Perry *et al.*, 1992).

We have determined the high resolution X-ray crystal structure of adamalysin II, a 203 amino acid residue snake venom metalloproteinase. In the following we present its essential structural features and show that it can be grouped into a common superfamily together with matrix metalloproteinases, astacins and *Serratia*-like proteinases. Adamalysin II (EC 3.4.24.46), previously known as proteinase II, was isolated from the venom of the eastern diamondback rattlesnake *Crotalus adamanteus* (Kurecki *et al.*, 1978; Kress, 1986; Kress and Catanese, 1981). This enzyme has been shown to exhibit no significant haemorrhagic activity, but to inactivate serpins by limited proteolysis of their reactive-site loops (Kurecki *et al.*, 1978; Kress, 1986; Kress and Catanese, 1981). It is a single-chain 24 kDa protein, which requires zinc and calcium for its activity. As expected from its amino acid composition, our 'crystallographic sequencing' revealed a close sequential similarity to the low molecular weight metalloproteinases Ht-c and Ht-d [which themselves degrade type IV collagen and various gelatins and are considered as haemorrhagic (Shannon *et al.*, 1989)] isolated from the venom of *Crotalus atrox* (see Table I).

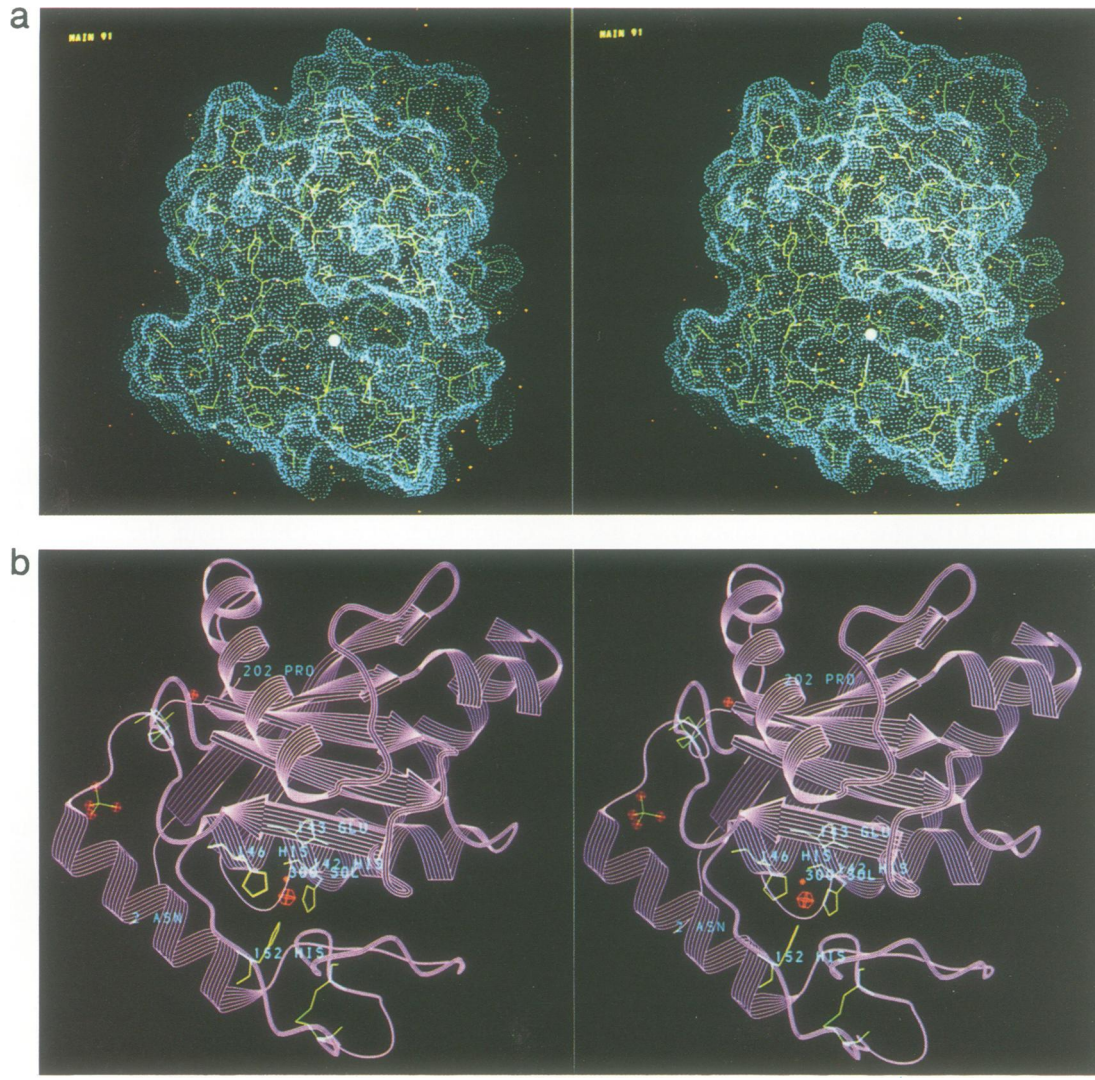

Fig. 1. 'Front view' of adamalysin II (**a**) displayed together with the Connolly dot surface [calculated with a rolling sphere of 1.4 Å radius, MAIN (Turk, 1992)], or (**b**) shown as a ribbon model (RIBBON; Priestle, 1988). An 'upper' main domain and a 'lower' subdomain are separated by the active-site cleft, which accommodates the zinc ion [displayed as a white (a) or red sphere (b), respectively]. A bound peptide substrate would presumably bind to the cleft running from left to right (Matthews, 1988). The second red blob (back of panel b) represents the bound calcium ion.

Results and discussion

We have crystallized adamalysin II (Bode and Kress, 1983) and solved its X-ray crystal structure by multiple isomorphous replacement at 2.0 Å resolution (see Table II). The structure reveals a compact molecule of oblate ellipsoidal shape notched at the periphery to give a relatively flat active-site cleft separating the small 'lower' subdomain from the 'upper' main body (Figure 1a). The peptide chain starts at the subdomain surface and forms the regularly folded 'upper' domain (residues 6–150; see Figure 1b). This domain consists essentially of a central, highly twisted five-stranded β-pleated sheet of $-1\times$, $2\times$, 2, -1 topology flanked by a long surface-located helix and large loop segments on its convex side, and by two long helices, one of which represents the central 'active-site helix', on its concave side (Figure 1b). The 'lower' subdomain, comprising the last 50 residues, is organized in multiple turns cross-connected by a disulfide bridge, Cys157–Cys164, with the chain ending in a long C-terminal helix and an extended segment clamped to the 'upper' domain via Cys117–Cys197 (Figure 1b). The N- and the C-termini are both located on the molecular surface and are thus [in contrast to the astacins (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993), but probably in agreement with the collagenases, see below] not involved in specific intradomainal stabilization.

In the active-site cleft, the catalytic zinc ion is co-ordinated by His142, His146 and His152 and by a water molecule (Sol300, anchored to Glu143), which probably acts as a catalytic base (Matthews, 1988), in a nearly tetrahedral manner, with the three Nϵ2 nitrogen atoms and the solvent oxygen atom located 2.1 Å and 2.3 Å away from the zinc ion (Figure 2 and Table I). His142, Glu143 and His146 are part of the long 'active-site helix', which extends up to Gly149, where it turns sharply toward His152. Asp153, which is strictly conserved in snake venom metalloproteinases (see Table I), is buried in the subdomain and seems to stabilize the hydrophobic active-site basement.

A calcium ion, identified on the basis of high electron density and short distances to neighbouring oxygen atoms, is located on the surface of adamalysin II close to the C-terminus opposite to the active site (Figure 1b). This calcium ion is liganded by the side chains of Asp93, Asn200 and Glu9, by an internal and an external water molecule and the

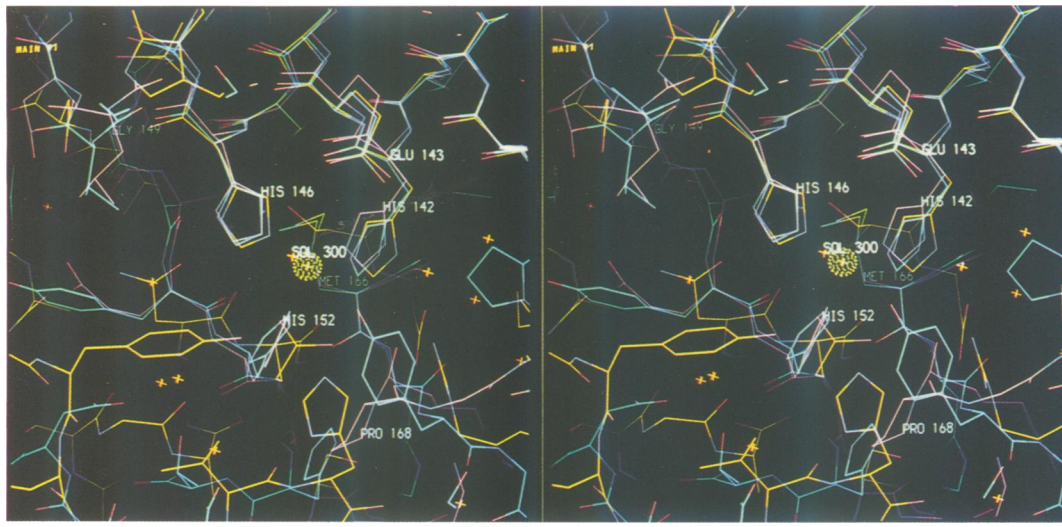

Fig. 2. Stereo figure showing the corresponding zinc-binding sites of adamalysin II (magenta), astacin (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993) (blue) and thermolysin (Holmes and Matthews, 1982) (yellow); orientation as in Figures 1 and 3. The adamalysin II structure is displayed together with the zinc ion (dense sphere of radius 0.5 Å) and a few localized solvent molecules (crosses). The structures of the non-liganded thermolysin (PDB code 3TLN; Holmes and Matthews, 1982; Bernstein *et al.*, 1977) and of astacin (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993) are superimposed on the adamalysin II structure after a least-squares fit of their central active-site helix residues (r.m.s. = 0.3 Å). In all three peptidases the active-site helix (above) provides two zinc-liganding histidines (His142 and His146 in adamalysin II and thermolysin, His92 and His96 in astacin) and the glutamic acid (Glu143 and Glu93) binding the zinc-liganding water (Sol300). However, in adamalysin II and astacin this helix terminates three residues later (at Gly149 and Gly99) with the chain turning back toward the third liganding histidine (His152 and His102, lower left corner). Only these two peptidases exhibit 'Met-turns', with the side chains of Met166 and Met147 forming hydrophobic basements for the three histidine ligands and the zinc ion. The 'fifth' zinc-liganding position (right lower corner), occupied by Tyr149 in astacin, is free in adamalysin II; the equivalent Pro168 is buried more deeply in the adamalysin II structure.

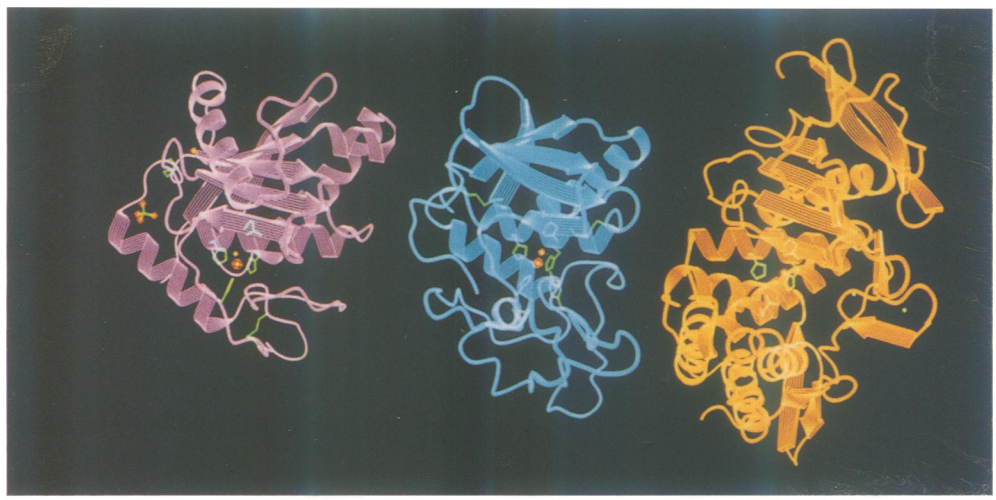

Fig. 3. Adamalysin II structure (**a**; magenta) shown as a ribbon model (Priestle, 1988) compared with ribbon plots of astacin (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993) (**b**; blue) and thermolysin (Holmes and Matthews, 1982) (**c**; yellow); all three peptidases are displayed in equivalent orientations, with the active-site helix at the same height. The zinc ions and the calcium ions (adamalysin II and thermolysin) as well as the zinc-liganding side chains and the disulfide bridges are displayed. The five-stranded β-sheet and the two helices on its concave side in the 'upper' domain and the C-terminal helix in the 'lower' domain are topologically similar in adamalysin II and astacin; the larger thermolysin exhibits a more extended β-sheet scaffold and similarly placed helices in its 'upper' domain, but a completely unrelated multi-helical 'lower' domain.

carbonyl group of Cys197. The aspartate at position 93 is strictly invariant, and the other calcium ligands are highly conserved in other venom metalloproteinases.

The zinc-binding sequences of 16 adamalysin II-related proteins (comprising 13 snake venom proteinases and three mammalian reproductive tract proteins), hereafter referred to as adamalysins, are shown in Table I. For most of the 200–204 residue catalytic domains of the adamalysins, sequence alignment is possible with no or with just one single insertion (Takeya *et al.*, 1990a,b; Hite *et al.*, 1992; Paine *et al.*, 1992). Adamalysin II, with the two disulfides (Cys117–Cys197 and Cys157–Cys164) is a typical '2-disulfide proteinase' (Figure 1b), as are also HT-2 (Takeya *et al.*, 1990b) and Ht-d/c (Neeper and Jacobson, 1990). The majority of the adamalysins (the '3-disulfide proteinases'), however, exhibit disulfide connectivities Cys157–Cys181 and Cys159–Cys164 (Table I) in their 'lower' domain (Takeya *et al.*, 1989, 1990a; Miyata *et al.*,

Table I. Sequence alignment

				142	143			146			149			152				156				159
Adamalysin II	T	M	A	**H**	**E**	L	G	**H**	N	L	**G**	M	E	**H**	**D**	G	K	D	–	**C**	L	R
HT-2	T	M	A	**H**	**E**	L	G	**H**	N	L	**G**	M	E	**H**	**D**	G	K	D	–	**C**	L	R
Ht-d/c	T	M	A	**H**	**E**	L	G	**H**	N	L	**G**	M	E	**H**	**D**	G	K	D	–	**C**	L	R
Ht-e	T	M	T	**H**	**E**	L	G	**H**	N	L	**G**	I	H	**H**	**D**	T	D	S	–	**C**	S	C
LHFII	T	M	A	**H**	**E**	L	G	**H**	N	L	**G**	M	K	**H**	**D**	E	N	H	–	**C**	H	C
Rhodostomin	I	M	A	**H**	**E**	M	G	**H**	N	L	**G**	V	R	**H**	**D**	G	E	Y	–	**C**	T	C
Jararhagin	I	M	A	**H**	**E**	M	G	**H**	N	L	**G**	I	H	**H**	**D**	T	G	S	–	**C**	S	C
HR1B	I	M	A	**H**	**E**	M	G	**H**	N	L	**G**	I	P	**H**	**D**	G	N	S	–	**C**	T	C
HR2a	T	M	T	**H**	**E**	I	G	**H**	N	L	**G**	M	E	**H**	**D**	D	K	D	K	**C**	K	C
Trigramin	T	M	T	**H**	**E**	M	G	**H**	N	L	**G**	M	H	**H**	**D**	E	D	K	–	**C**	N	C
H2	T	M	T	**H**	**E**	L	G	**H**	N	L	**G**	M	E	**H**	**D**	D	K	D	K	**C**	K	C
Fibrolase	T	M	A	**H**	**E**	L	G	**H**	N	L	**G**	M	N	**H**	**D**	G	N	Q	–	**C**	H	C
RVV-XH	I	M	A	**H**	**E**	L	S	**H**	N	L	**G**	M	Y	**H**	**D**	G	K	N	–	**C**	I	C
Rat EAP I	R	M	A	**H**	**Q**	L	G	**H**	S	L	**G**	M	R	**H**	**D**	D	F	P	–	**C**	T	C
Monkey EAP I	R	M	A	**H**	**Q**	L	G	**H**	N	L	**G**	M	Q	**H**	**D**	E	F	P	–	**C**	T	C
Cyritestin	V	L	S	Q	L	L	G	I	N	L	**G**	L	A	Y	**D**	D	V	Y	N	**C**	F	C
Astacin	T	I	I	**H**	**E**	L	M	**H**	A	I	**G**	F	Y	**H**	E	H	T	R	M	D	R	D
Serratia prot.	T	F	T	**H**	**E**	I	G	**H**	A	L	**G**	L	S	**H**	P	G	D	Y	N	A	G	E
Stromelysin	V	A	A	**H**	**E**	I	G	**H**	S	L	**G**	L	F	**H**	S	A	N	T	E	A	L	M
Thermolysin	V	V	G	**H**	**E**	L	T	**H**	A	V	T	D	Y	T	A	G	L	V	Y	Q	N	E

Sequence alignment (shown in the one-letter code) around the zinc-binding consensus sequence for adamalysin II (as obtained from the electron density and from partial peptide sequencing) and the related adamalysins: namely, the snake venom haemorrhagic proteins HT-2 (Takeya *et al.*, 1990b) (from *Crotalus ruber ruber*), Ht-d/c (Shannon *et al.*, 1989) (*Crotalus atrox*), Ht-e (Hite *et al.*, 1992) (*C.atrox*), LHFII (Sánchez *et al.*, 1991) (*Lachesis muta muta*), rhodostomin (Au *et al.*, 1991) (*Calloselasma rhodostoma*), jararhagin (Paine *et al.*, 1992) (*Bothrops jararaca*), HR1B (Takeya *et al.*, 1990a) and HR2a (Miyata *et al.*, 1989) (*Trimeresurus flavoviridis*), the trigramin precursor (Neeper and Jacobson, 1990) (*Trimeresurus gramineus*), the snake venom metalloproteinases H2 (Takeya *et al.*, 1989) (*T.flavoviridis*), fibrolase (Randolph *et al.*, 1992) (*Agkistrodon contortrix contortrix*) and the factor X activating enzyme RVV-XH from Russell's viper venom (Takeya *et al.*, 1992), as well as the epididymal apical proteins EAP I from rat (*Rattus norvegicus*) and monkey (*Macaca fascicularis*) (Perry *et al.*, 1992), and cyritestin (U.Heinlein, personal communication). In addition, one representative protein each of the astacins (astacin; Titani *et al.*, 1987), large bacterial proteinases (*Serratia* proteinase; Nakahama *et al.*, 1986), and matrix metalloproteinases (human stromelysin, MMP-3; Whitham *et al.*, 1986), which contain the consensus sequence H(E,Q)XXHXXGXXH involved in zinc ligation, is displayed. The shorter HEXXH consensus sequence of thermolysin (Titani *et al.*, 1972) is also shown. The numbering refers to the HT-2 sequence (Takeya *et al.*, 1990b).

1989) in addition to the Cys117–Cys197 linkage. Cys164, placed in a conserved turn Cys164–Pro168 (see below), and residue 181, at the third position of the C-terminal helix, seem to be restrained to their respective sites as found in adamalysin II. The disulfide connectivity in these '3-disulfide proteinases' would therefore require some conformational shifts in the lower domain.

Crystal structures of zinc-endopeptidases are available for only three thermolysin-like bacterial proteinases (Matthews, 1988; Holland *et al.*, 1992) and for astacin (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993). This latter metalloproteinase derived from the crayfish *Astacus astacus* is the prototype of the astacin family (Stöcker *et al.*, 1988); up to now its structure has been the best starting point for modelling matrix metalloproteinases. The five-stranded β-pleated sheet of adamalysin II and the two α-helices attached to its concave side are topologically similar to elements in astacin (and presumably also to the catalytic domain of matrix metalloproteinases) and to a much lesser degree also to thermolysin (Matthews, 1988; Holland *et al.*, 1992) (Figure 3). The small 'lower' subdomain of adamalysin II bears some weaker similarity to the larger 'lower' domain of astacin, in that both enzymes exhibit a similarly oriented but shifted C-terminal helix, which ends with a surface located C-terminal segment disulfide-linked to the 'upper' domain (Figure 3). Of strikingly similar arrangement and conformation is a turn containing a conserved methionine (the 'Met-turns' Cys164-Ile165-Met166-Arg167-Pro168 of adamalysin II, and Ser145-Ile146-Met147-His148-Tyr149 of astacin, with Tyr149 constituting the fifth zinc ligand), which forms a hydrophobic basement for the active-site groups in both proteinases (Figure 2). The matrix metalloproteinases (matrixins) and the *Serratia*-like proteinases (serralysins) possess similar conserved methionine sequences (Ala-Leu-Met-Tyr-Pro and Ser-Ile-Met-Ser-Tyr, respectively) C-terminal to their zinc-binding segments, also implying that there are similar active-site basements in these metalloproteinases. In contrast to the astacins and to serralysins, adamalysin II exhibits no fifth tyrosine zinc ligand; rather, the topologically equivalent residue is Pro168 of the above mentioned Met-turn, the side chain of which is only weakly exposed and points away from the catalytic zinc (see Figure 2). The equivalent proline residue in the matrixins, which appears to be fully conserved, is probably arranged in a similar conformation.

The central 'active-site helix' of adamalysin II (residues 137–148) is almost exactly superimposable on the equivalent helices in astacin and thermolysin. In particular, the first two histidine ligands, His142 and His146, are identically arranged, while the carboxylate oxygen atoms of the enclosed Glu143 are much more distant (>4 Å) from the equivalent zinc-liganding Sol300 (see Figure 2). Following this helix, the adamalysin II peptide chain turns sharply away from the helix axis at Gly149 towards the third zinc ligand, His152 (Figures 1b and 2). As previously predicted (Stöcker *et al.*, 1988), this conformation is identical to that of astacin and very probably also to the matrixins and serralysins (see Table I), but unlike that of thermolysin. The Gly109–Asn114 segment of adamalysin II forms the upper edge of the active-site cleft in a similar fashion to the edge

Table II. Crystallographic and refinement data

Crystal data (native protein)		Current refinement parameters	
Space group	$P3_212$	Protein atoms	1619
Unit cell (Å)		Water molecules	171
a and b	73.6	Reflections	16 776
c	96.4	r.m.s. deviation from target values	
		bonds (Å)	0.016
Resolution (Å)	2.0	angles (degrees)	1.983
Total measured reflections	24 244		
Unique reflections	17 021	Resolution range	8.0–2.0 Å
% Completeness to 2.0 Å	83%		
% Completeness last shell		1790 individual B-factors	
2.09–2.03 Å	42%	R_{cryst}[a]	0.183
R_{merge}[b]	0.0544	R_{cryst}[a](2.09–2.00 Å; 985 refl.)	0.27

Derivative data Compound/concentration	Soaking time	Number of sites	Phasing power[c]	Limiting resolution (Å)
$HgAc_2$ (3 mM)/$SH\text{-}(CH_2)_2\text{-}NH_3Cl$ (12 mM)/$CaCl_2$ (9 mM) + $HgCl_2$ (1 mM)/$CaCl_2$ (2 mM)	30 min + 15 min	3	1.16	20–2.2
K_2PtCl_4 (0.2 mM)	7 days	2	1.12	20–3.0
K_2PtCl_6 (<2.5 mM)	11 days	2	0.91	20–3.0
$Rb_3Re_3Cl_{12}$ (<<2 mM)	12 days	2	0.80	20–3.0
$HgCl_2$ (2.4 mM)/K_2PtCl_4 (1.1 mM)	4.5 h	3	1.31	20–2.5
OP:1,10-phenanthroline (apoenzyme)	24 h	3	0.95	20–2.5
Overall figure of merit (20.0–2.2 Å)	0.52			

[a] $R_{cryst} = \Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$.
[b] $R_{merge} = \Sigma\Sigma|I(h)_i - \langle I(h)\rangle|/\Sigma\langle I(h)\rangle$ where $I(h)_i$ 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, scaling, and merging for separately averaged Friedel pairs.
[c] Phasing power = $\langle F_H\rangle/E$, i.e. the r.m.s. heavy-metal structure factor amplitude divided by the residual lack of closure error.

β-strands in astacin and thermolysin; thus the carbonyl group of Gly109 presumably plays a similar role in hydrogen bonding and polarizing the amino group of the P_1' residue of a bound substrate (Matthews, 1988; Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993). By further analogy, Glu143 and the zinc ion probably polarize water molecule Sol300, promoting its attack on the carbonyl group of the scissile peptide bond (Figure 2). As mentioned before, however, adamalysin II lacks in its 'fifth' zinc liganding position any proton donor group [such as Tyr149 of astacin (Bode *et al.* 1992; Gomis-Rüth *et al.*, 1993) or His231 of thermolysin (Matthews, 1988; Holland *et al.*, 1992)], which could either act as a shuttle or further assist in polarization.

A hydrophobic pocket (Figure 1a) forms the putative S_1' subsite of adamalysin II, which is therefore (in contrast to astacin, but probably similar to the matrixins due to a similar sequence preference) well equipped to accommodate bulky, hydrophobic P_1' residues such as leucine. Due to a barrier to the right side of the cleft (see Figure 1a), peptidyl substrates must probably bind in a kinked conformation at P_1', similar to that observed for phosphonamides in thermolysin complexes (Matthews, 1988; Holland *et al.*, 1992). Further studies are under way to identify the extended substrate recognition site of adamalysin II and to analyse the detailed nature of substrate interactions; these must be different from thermolysin, but may resemble those expected for matrixins, due to a similar recognition pattern about X-Leu bonds (Hite *et al.*, 1992) and the common requirements for elongated peptidic substrates.

A comparison of the crystal structure of adamalysin II with those of astacin and thermolysin reveals that ~106 (78) out of the 202 adamalysin II residues can be considered as topologically equivalent to those in astacin (thermolysin), with relatively high r.m.s. deviations of 3.2 Å (3.5 Å) for their α-carbon atomic positions. However, of these only 14 (12) residues are identical, seven (five) of which are situated in the zinc-binding consensus region (Table I). Thus, the adamalysins exhibit an identical active-site environment and (despite low sequence homology) some overall topological similarity to the astacins and probably also to the matrixins, with which they should be grouped into a common superfamily (see Murphy *et al.*, 1991; Jiang and Bond, 1992; Rawlings and Barrett, 1993).

The few amino acid residues which have been suggested to be related to induction of haemorrhagic activity in the 'haemorrhagic' snake venom proteinases (Takeya *et al.*, 1990b; Hite *et al.*, 1992), are in adamalysin II not clustered together in space, but are partly on the molecular surface and partly buried within the molecule. They are therefore very probably not the real structural origins of haemorrhage. The larger proteinases are generally much more haemorrhagic, i.e. are capable of provoking instantaneous bleeding even at low doses (Bjarnasson and Fox, 1989). Thus, the appending domains of snake venom metalloproteinases might support the catalytic domain in recognizing and specifically cleaving macromolecular substrates such as collagens (Hite *et al.*, 1992), similar to the haemopexin-like domains in collagenases (see Woessner, 1991). The surface location of the N- and the C-termini of the polypeptides in adamalysin II suggests that the flanking domains of

multimodular species are probably flexibly linked to the catalytic domain.

The mammalian reproductive tract proteins EAP I (Perry *et al.*, 1992), cyritestin and PH-30-α and -β (Blobel *et al.*, 1992; Blobel and White, 1992) also exhibit an adamalysin II-like domain of >30% sequence identity. Interestingly, the highly conserved active-site Glu143 of adamalysin II is replaced in rat and monkey EAP I by a glutamine (Perry *et al.*, 1992) (see Table I), a substitution somewhat reminiscent of the Asp to Asn substitution in the active-site triad on going from trypsin-like serine to papain-like cysteine proteinases. Mouse testis cyritestin shows a similar degree of homology within its adamalysin II-like domain; however, it possesses neither zinc-liganding His nor Glu residues in the 'zinc binding consensus' region (see Table I), and can therefore neither bind a zinc ion nor exhibit proteolytic activity.

Materials and methods

The protein sample was prepared as previously described (Bode and Kress, 1983) and crystallized as trigonal crystals of space group $P3_212$ (cell constants a = b = 73.6 Å, c = 96.4 Å) from 1.8 M ammonium sulfate at pH 5. The crystals were harvested with 2.85 M ammonium sulfate, 5 mM $ZnCl_2$, 5 mM $CaCl_2$ at pH 5.0; data collection of native and derivative crystals was performed with the same buffer, except that zinc was omitted in the case of the heavy metal derivatives (see Table II). X-ray intensity data from a native crystal (to 2.0 Å resolution) and from three heavy-atom derivatives were collected at 4°C on a FAST detector (Enraf Nonius, Delft) using monochromatic CuKα radiation from one single crystal each and processed on-line with MADNES (Messerschmidt and Pflugrath, 1987). Data for three more derivatives were collected from two or three crystals using X-ray films scanned on a rotating-drum densitometer. Reflections with $I/\sigma I$ of <2 were rejected. The heavy atom sites (see Table II) and the phases were refined using MIRPH (J.Remington). An adamalysin II starting model built using PSFRODO (Jones, 1978) on an ESV-30 Graphic system workstation (Evans and Sutherland) against an initial multiple isomorphous replacement map smoothed by solvent flattening (Wang, 1983) was improved

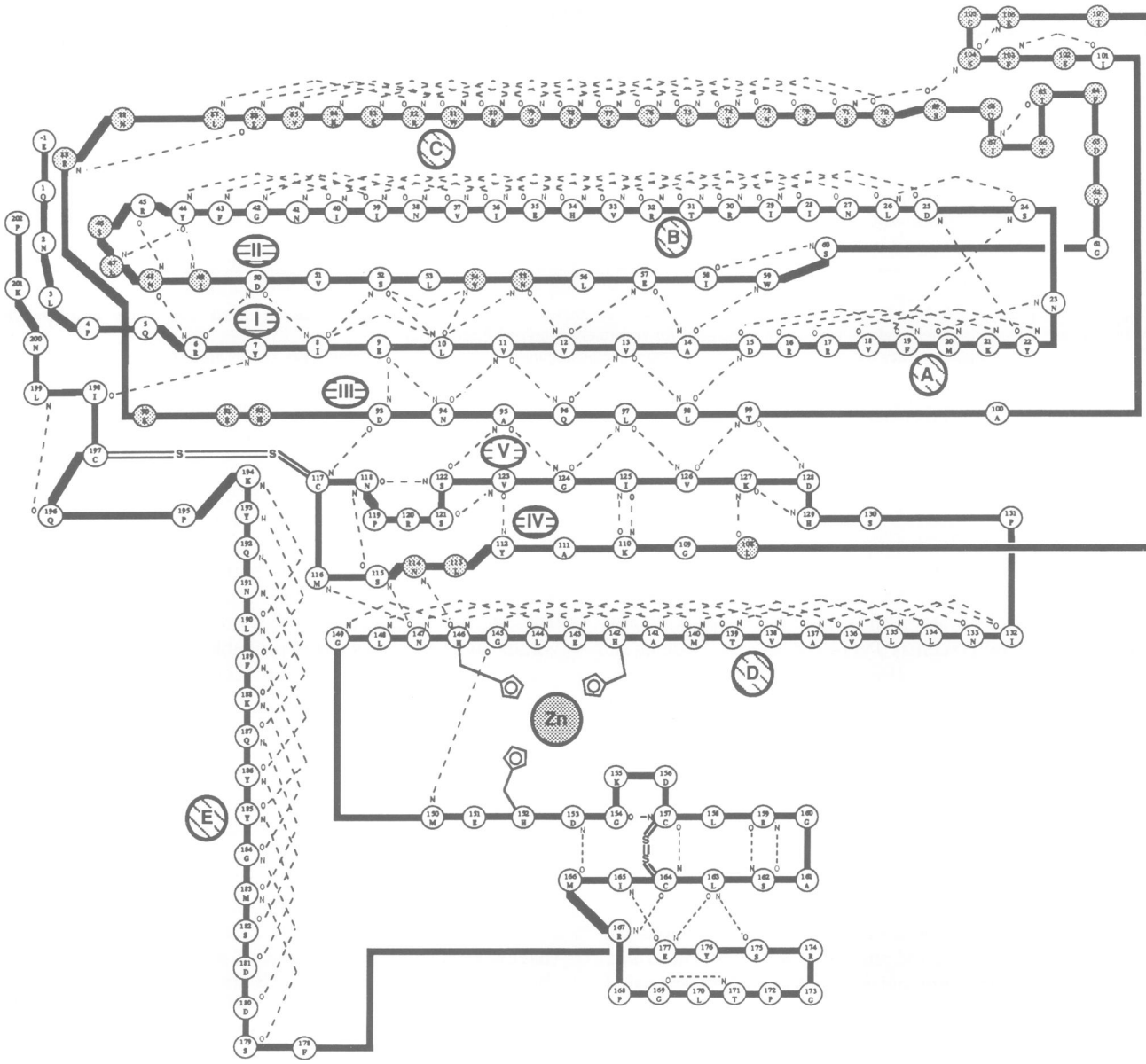

Fig. 4. Secondary structure plot of adamalysin II, with the 'upper' and the 'lower' domains arranged above and below the active-site cleft. Intra main-chain hydrogen bonds are selected according to criteria given by Kabsch and Sander (1983) using an energy threshold of −0.7 kcal/mol, and are indicated by dashed lines. Disulfide connections are displayed by double lines. Helices are labelled with capital letters (A to E) and β-strands with Roman numerals (I to V). All amino acid residues labelled with open circles were confirmed by amino acid sequencing; those in shaded circles are consistent with the current electron density.

and completed in iterative cycles, each cycle comprising modelling against Fourier maps calculated with combined or (in the final steps) model calculated phases (PROTEIN; Steigemann, 1974), and energy-restrained crystallographic refinement with X-PLOR (Brünger, 1990) using parameters derived from the Cambridge database (Engh and Huber, 1991). The primary structure of adamalysin II was first derived from the electron density; three-quarters of this sequence has now been confirmed by protein sequencing (open circles in Figure 4). Accordingly, the adamalysin II molecule exhibits ~80% sequence identity with Ht-d (Shannon *et al.*, 1989), extending from pyroglutamate −1 to Pro202 [HT-2 nomenclature (Takeya *et al.*, 1990b)]. The adamalysin II chain is fully defined by electron density from Asn2 to Pro202 and exhibits acceptable main chain geometry. Only loop Lys104−Glu106 is less well positioned, and its side chains are only partially localizable. The crystallographic R-value of the current model is 0.183 for all 16 776 independent reflections between 8.0 and 2.0 Å resolution. A summary of the crystallographic and the refinement data is provided in Table II. After completion of protein sequencing and refinement, this structure will be deposited in the Brookhaven data bank.

The three structures (adamalysin II, astacin and thermolysin) were compared using the program OVRLAP (Rossmann and Argos, 1976), which minimizes spatial positions as well as segment orientations. Quite generous equivalence probability weights are chosen for the spatial and the orientational part ($E_1 = E_2 = 4.0$ Å). Pictures and plots were performed with the graphics facilities of RIBBON (Priestle, 1988; modified by A.Karshikov) and MAIN (Turk, 1992).

Acknowledgements

We thank Mrs I.Mayr for crystallization, Dr F.X.Li for assistance in data collection and derivative search, Dr J.Kellermann for peptide sequencing, Prof. R.Huber for continuous support of this work, and Dr M.Stubbs for careful reading. The financial support of the SFB207 of the Universität München and of the Fonds der Chemischen Industrie (W.B.), of an EMBO long-term fellowship (F.-X.G.R.) and of NIH grant HL22996 (L.F.K.) is acknowledged.

References

Au,L.C., Huang,Y.B., Huang,T.F., Teh,G.W., Lin,H.H. and Choo,K.B. (1991) *Biochem. Biophys. Res. Commun.*, **181**, 585−593.

Bernstein,F.C., Koetzle,T.F., Williams,G.J.B., Meyer,E.F., Brice,M.D., Rogers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) *J. Mol. Biol.*, **122**, 535−542.

Birkedal-Hansen,H., Moore,W.G.I., Bodden,M.K., Windsor,L.J., Birkedal-Hansen,B., De Carlo,A. and Engler,J.A. (1992) *Crit. Rev. Oral Biol. Medicine*, **4**, 197−250.

Bjarnasson,J.B. and Fox,J.W. (1989) *J. Toxicol. Toxin Rev.*, **7**, 121−209.

Blobel,C.P. and White,J.M. (1992) *Curr. Opin. Cell Biol.*, **4**, 760−765.

Blobel,C.P., Wolfsberg,T.G., Turck,C.W., Myles,D.G., Primakoff,P. and White,J.M. (1992) *Nature*, **356**, 248−252.

Bode,W. and Kress,L.F.(1983) *FEBS Lett.*, **154**, 274−276.

Bode,W., Gomis-Rüth,F.X., Huber,R., Zwilling,R. and Stöcker,W. (1992) *Nature*, **358**, 164−167.

Brünger,A.T. (1990) *X-PLOR Manual version 2.1* (Yale University, New Haven).

Engh,R.A. and Huber,R (1991) *Acta Crystallogr.*, **A47**, 392−400.

Fox,J.W., Campbell,R., Beggerly,L. and Bjarnasson,J.B. (1986) *Eur. J. Biochem.*, **156**, 65−72.

Gomis-Rüth,F.X., Stöcker,W., Huber,R., Zwilling,R. and Bode,W. (1993) *J. Mol. Biol.*, **229**, 945−968.

Hite,L.A., Shannon,J.D., Bjarnasson,J.B. and Fox,J.W. (1992) *Biochemistry*, **31**, 6203−6211.

Holland,D.R., Tronrud,D.E., Pley,H.W., Flaherty,K.M., Stark,W., Jansonius,J.N., McKay,D.B. and Matthews,B.W. (1992) *Biochemistry*, **31**, 11310−11316.

Holmes,M.A. and Matthews,B.W. (1982) *J. Mol. Biol.*, **160**, 623−639.

Jiang,W. and Bond,J.S.(1992) *FEBS Lett.*, **312**, 110−114.

Jones,T.A. (1978) *J. Appl. Crystallogr.*, **15**, 23−31.

Kabsch,W. and Sander,C. (1983) *Biopolymers*, **22**, 2577−2637.

Kini,R.M. and Evans,H.J. (1992) *Toxicon*, **30**, 265−293.

Kress,L.F. (1986) *J. Cell. Biochem.*, **32**, 51−58.

Kress,L.F. and Catanese,J.J. (1981) *Biochemistry*, **20**, 7432−7438.

Kurecki,T., Laskowski,M.,Sr and Kress,L.F. (1978) *J. Biol. Chem.*, **253**, 8340−8345.

Matthews,B.W. (1988) *Accts. Chem. Res.*, **21**, 333−340.

Messerschmidt,A. and Pflugrath,J.W. (1987) *J. Appl. Crystallogr.*, **20**, 306−315.

Miyata,T., Takeya,H., Ozeki,Y., Arakawa,M., Tokunaga,F., Iwanaga,S. and Omori-Satoh,T. (1989) *J. Biochem.*, **105**, 847−853.

Murphy,G. and Docherty,A.J.P. (1992) *Am. J. Respir. Cell Mol. Biol.*, **7**, 120−125.

Murphy,G.J.P., Murphy,G. and Reynolds,J.J. (1991) *FEBS Lett.*, **289**, 4−7.

Nakahama,K., Yoshimura,K., Marumoto,R., Kikuchi,M., Lee,I.S., Hase,T. and Matsubara,H. (1986) *Nucleic Acids Res.*, **14**, 5843−5855.

Neeper,M.P. and Jacobson,M.A. (1990) *Nucleic Acids Res.*, **18**, 4255−4255.

Paine,M.J.I., Desmond,H.P., Theakston,R.D.G. and Crampton,J.M. (1992) *J. Biol. Chem.*, **267**, 22869−22876.

Perry,A.C.F., Jones,R., Barker,P.J. and Hall,L. (1992) *Biochem. J.*, **286**, 671−675.

Priestle,J.P. (1988) *J. Appl. Crystallogr.*, **21**, 572−576.

Randolph,A., Chamberlain,S.H., Chu,H.L.C., Retzios,A.D., Markland,F.S.,Jr and Masiarz,F.R. (1992) *Protein Sci.*, **1**, 590−600.

Rawlings,D. and Barrett,A.J. (1993) *J. Biochem.*, **290**, 205−218.

Rossmann,M.G. and Argos,P. (1976) *J. Mol. Biol.*, **105**, 75−95.

Sánchez,E.F., Diniz,C.R. and Richardson,M. (1991) *FEBS Lett.*, **282**, 178−182.

Shannon,J.D., Baramova,E.N., Bjarnasson,J.B. and Fox,J.W. (1989) *J. Biol. Chem.*, **264**, 11575−11583.

Steigemann,W. (1974) Ph.D. Thesis. Technische Universität München.

Stöcker,W., Wolz,R.L., Zwilling,R., Strydom,D.J. and Auld,D.S. (1988) *Biochemistry*, **27**, 5026−5032.

Takeya,H., Arakawa,M., Miyata,T., Iwanaga,S. and Omori-Satoh,T. (1989) *J. Biochem.*, **106**, 151−157.

Takeya,H., Oda,K., Miyata,T., Omori-Satoh,T. and Iwanaga,S. (1990a) *J. Biol. Chem.*, **265**, 16068−16073.

Takeya,H., Onikura,T., Sugihara,H. and Iwanaga,S. (1990b) *J. Biochem.*, **108**, 711−719.

Takeya,H., Nishida,S., Miyata,T., Kawada,S.-I., Saisaka,Y., Morita,T. and Iwanaga,S. (1992) *J. Biol. Chem.*, **267**, 14109−14117.

Titani,K., Hermodson,M. A., Ericsson,L. H., Walsh,K.A. and Neurath,H. (1972) *Nature*, **238**, 35−37.

Titani,K., Torff,H.-J., Hormel,S., Kumar,S., Walsh,K.A., Rödl,J., Neurath,H. and Zwilling,R. (1987) *Biochemistry*, **26**, 222−226.

Turk,D. (1992) Ph.D. Thesis. Technische Universität München.

Wang,B.C. (1983) *Lecture Notes for School on Direct Methods and Macromolecular Crystallography*. Medical Foundation of Buffalo, Buffalo, NY, pp. 1−7.

Whitham,S.E., Murphy,G., Hugel,P., Rahmsdorf,H.-J., Smith,B.J., Lyons,A., Harris,T.J.R., Reynolds,J.J., Herrlich,P. and Docherty,A.J.P. (1986) *Biochem. J.*, **240**, 913−916.

Woessner,J.F. (1991) *FASEB J.*, **5**, 2145−2154.

Received on June 6, 1993; revised on July 19, 1993

Note added in proof

Our newly determined catalytic domain crystal structure of a human collagenase indeed confirms that the adamalysins share not only a similar overall topology and an identical zinc environment with the collagenases, but also a very similar active site cleft geometry, i.e. they also topologically represent 'collagenase'.