Protein Science (2000), *9*:1365–1373. Cambridge University Press. Printed in the USA.

The disulfide bond pattern of catrocollastatin C, a disintegrin-like/cysteine-rich protein isolated from *Crotalus atrox* venom

JUAN J. CALVETE,[1] M. PAZ MORENO-MURCIANO,[1] LIBIA SANZ,[1] MICHAEL JÜRGENS,[2] MICHAEL SCHRADER,[2] MANFRED RAIDA,[3] DAVID C. BENJAMIN,[4] AND JAY W. FOX[5]

[1]Instituto de Biomedicina, C.S.I.C., Jaime Roig 11, E-46010 Valencia, Spain
[2]BioVisioN GmbH & Co. KG, Feodor-Lynen-Strasse 5, D-30625 Hannover, Germany
[3]Niedersächsisches Institut für Peptid-Forschung GmbH, Feodor-Lynen-Strasse 31, D-30625 Hannover, Germany
[4]Beirne B. Carter Center for Immunology Research, University of Virginia Health Sciences Center, MR4 Box 4012, Charlottesville, Virginia 22908
[5]Department of Microbiology, University of Virginia Health Sciences Center, Box 441 Jordan Hall, Charlottesville, Virginia 22908

(RECEIVED December 7, 1999; FINAL REVISION April 17, 2000; ACCEPTED April 28, 2000)

Abstract

The disulfide bond pattern of catrocollastatin-C was determined by N-terminal sequencing and mass spectrometry. The N-terminal disintegrin-like domain is a compact structure including eight disulfide bonds, seven of them in the same pattern as the disintegrin bitistatin. The protein has two extra cysteine residues (XIII and XVI) that form an additional disulfide bond that is characteristically found in the disintegrin-like domains of cellular metalloproteinases (ADAMs) and PIII snake venom Zn-metalloproteinases (SVMPs). The C-terminal cysteine-rich domain of catrocollastatin-C contains five disulfide bonds between nearest-neighbor cysteines and a long range disulfide bridge between CysV and CysX. These results provide structural evidence for a redefinition of the disintegrin-like and cysteine-rich domain boundaries. An evolutionary pathway for ADAMs, PIII, and PII SVMPs based on disulfide bond engineering is also proposed.

Keywords: ADAM; catrocollastatin-C; cysteine-rich domain; disintegrin-like domain; disulfide bond pattern; mass spectrometry; Reprolysin protein family; snake venom protein

Snake venoms contain a complex mixture of pharmacologically active peptides and proteins (Markland, 1998). Venoms of the *Elapidae* and *Hydrophiidae* families are primarily neurotoxic. On the other hand, the distinctive characteristic of *Viperidae* snake envenomation is local and, in severe cases, systemic hemorrhage. Crotalid and viperid snake venoms are rich sources of hemorrhagic toxins, which act in a synergistic manner to cause hemorrhage. The toxins primarily responsible for hemorrhage are Zn^{2+}-dependent metalloproteinases that digest components of the extracellular matrix surrounding capillaries giving rises to hemorrhage (Baramova et al., 1989; Bjarnason & Fox, 1994). The resultant hemorrhage is exacerbated by the presence in the venoms of disintegrins; the RGD-containing peptides that function as potent inhibitors of platelet aggregation (Niewiarowski et al., 1994; Calvete, 1997; McLane et al., 1998). Hemorrhagic toxins (hemorrhagins) belong to the Reprolysin family of metalloproteinases (Fox & Bjarnason, 1995; Rawlings & Barrett, 1995). According to their structural organization, snake venom metalloproteinases (SVMPs) can be divided into four classes (Jia et al., 1996). PI SVMP are single metalloproteinase domains; PII SVMPs exhibit a disintegrin domain carboxy terminal to the proteinase domain; PIII SVMPs possess metalloproteinase, disintegrin-like and a C-terminal cysteine-rich domain. In addition to the domain structure of PIII metalloproteinases, SVMPs of the PIV class have lectin domains at the carboxyl terminus. It has been proposed that disintegrins, and the non-RGD-containing, disintegrin domain of atrolysin E, are derived by proteolytic processing of PII class SVMP precursors (Au et al., 1991; Kini & Evans, 1992; Tsai et al., 1994; Fox & Bjarnason, 1995; Shimokawa et al., 1996; Shimokawa et al., 1998). The PIII class of SVMPs is related to the ADAM/MDC (A Disintegrin And Metalloproteinase/Metalloproteinase-Disintegrin-Cysteine-rich) group of type I integral membrane proteins. However, in addition to homologous proteinase, disintegrin-like, and cysteine-rich domains, the ADAM/MDC proteinases possess a carboxyl-terminal extension comprised of growth factor-like, transmembrane, and cytoplasmic domains (Blobel et al., 1992; Perry et al., 1992; Wolfsberg & White, 1996). Class PIII SVMPs are significantly more toxic than the PI class hemorrhagins (Fox & Bjarnason,

Reprint requests to: Juan J. Calvete, Instituto de Biomedicina, C.S.I.C., Jaime Roig 11, E-46010 Valencia, Spain; e-mail: jcalvete@ibv.csic.es.

1995), suggesting that the metalloproteinase and the disintegrin-like/cysteine-rich domains synergistically contribute to hemorrhage production.

Disintegrins impair platelet aggregation by inhibiting in a competitive manner fibrinogen binding to the $\alpha_{IIb}\beta_3$ integrin on platelets. They use a common inhibitory mechanism involving the (K/R)GD tripeptide presented at the tip of a highly mobile loop. The potent activity of disintegrins depends on the appropriate disulfide bond pairings, which maintain the RGD-containing loop in its active conformation (Niewiarowski et al., 1994; Calvete, 1997; Calvete et al., 1997, and references therein; McLane et al., 1998). It is noteworthy that the disintegrin-like domains of PIII SVMPs and ADAM/MDC molecules lack the RGD cell-binding motif, which is replaced by XXCD sequences. Although the functional significance of the sequence degeneracy in the disintegrin-like loops is not clear, the recombinant disintegrin-like/cysteine-rich domains of *Crotalus atrox* hemorrhagin atrolysin A, as well as cyclic synthetic peptides containing the sequence RSECD, inhibit collagen- and ADP-stimulated platelet aggregation (Jia et al., 1997; Markland, 1998). Similarly, jararhagin-C and catrocollastatin-C, two MSECD-containing disintegrin-like and cysteine-rich domain fragments naturally derived, respectively, from PIII hemorrhagic toxins from *Bothrops jararaca* and *Crotalus atrox*, inhibit collagen-induced and ADP-stimulated (jararhagin) platelet aggregation (Paine et al., 1992; Zhou et al., 1995; Gichuhi et al., 1997; Shimokawa et al., 1997). Recent reports indicate that the disintegrin-like domains of mouse sperm fertilin (ADAM-1 and ADAM-2) might play a role in fertilization (Cho et al., 1998; Evans et al., 1998; Chen & Sampson, 1999, and references cited). However, the involvement of the disintegrin-like domain in integrin recognition is controversial. For example, Ivaska and coworkers (Ivaska et al., 1999) have reported that synthetic peptides displaying the RSECD sequence of the disintegrin-like domain of jararhagin failed to inhibit recombinant α2 I-domain binding to collagen, whereas a basic peptide from the metalloproteinase domain proved to be functional. In addition, Chen et al. (1998) examined the role of the EGF-like domains of mouse fertilin α and β, and found that a peptide corresponding to the cystine-constrained N-terminal subdomain of fertilin β had an activating effect on fertilization. The authors suggest that the EGF-like domain of fertilin β may have a function in sperm-egg binding and fusion, the same biological role proposed for the disintegrin-like domain.

The function of the cysteine-rich domain in the context of the disintegrin-like domain is unknown, although there is evidence that adhesion of atrolysin A to collagen I and von Willebrand factor is mediated through the cysteine-rich domain. Furthermore, recombinant cysteine-rich domain from atrolysin a has been demonstrated to inhibit collagen-stimulated platelet aggregation (Jia et al., 2000). Thus, this domain may serve to localize the hemorrhagic toxin to the subendothelial matrix (Fox & Long, 1998). It has also been suggested that a putative fusion peptide within the cysteine-rich domain of sperm fertilin α (ADAM 1) could have a role during gamete membrane fusion at fertilization (Blobel et al., 1992). The cysteine-rich domains of SVMPs and ADAM/MDC molecules are characterized by the presence of 13 strictly conserved cysteine residues in a distinctive sequence pattern (Jia et al., 1996). The odd number of cysteine residues in both the disintegrin-like (15 cys) and the cysteine-rich domains of Reprolysins suggests that one (or more) potentially unpaired cysteine residue in the cysteine-rich domain could be involved in disulfide bonding with the odd cysteine in the adjacent disintegrin-like domain. To ascertain this point, we have determined the disulfide bond pattern of the two-domain SVMP-PIII-derived fragment catrocollastatin-C.

Results

Quantitation of disulfide bonds

No carboxymethylcysteine was detected by amino acid analysis of catrocollastatin-C, which had been incubated with iodoacetamide under denaturing but nonreducing conditions. However, after reduction and alkylation with vinylpyridine, cysteine was quantitatively recovered as the pyridylethyl derivative, indicating that catrocollastatin-C does not have free thiol groups. This was confirmed by ESI mass spectrometry. Thus, the molecular masses of native catrocollastatin-C and catrocollastatin-C treated under nonreducing conditions with iodoacetamide were both 23,532 $\pm$ 2 Da, which corresponds to the isotope-averaged molecular mass calculated from the cDNA-deduced amino acid sequence of catrocollastatin-C (23,560.3 Da) with all 28 cysteine residues involved in disulfide bonds (23,532.3 Da). On the other hand, the molecular masses of reduced and carbamidomethylated and reduced and pyridylethylated catrocollastatin-C derivatives were 25,157 $\pm$ 2 Da and 26,504 $\pm$ 2 Da, respectively. These values are in excellent agreement with the isotope-averaged molecular masses calculated for catrocollastatin-C with 28 carbamidomethyl cysteine residues (25,157.6 Da) and 28 pyridylethyl cysteine residues (26,504.1 Da), respectively. As a whole, these results clearly showed that catrocollastatin-C does not possess any free sulfydryl groups, and that its 28 cysteine residues form 14 disulfide bonds.

Determination of the disulfide bond pattern of the cysteine-rich domain

Twenty-one proteolytic fragments were isolated by reversed-phase high-performance liquid chromatography (HPLC) after degradation of catrocollastatin-C with a mixture of trypsin and chymotrypsin. Each of these fractions was characterized by amino acid analysis, N-terminal sequencing, and mass spectrometry (Table 1). These fragments contained the necessary structural information for the unambiguous deduction of the disulfide bridge pattern of the cysteine-rich domain of catrocollastatin-C (Figs. 1, 2).

HPLC fraction TC1 contained a mixture of peptides corresponding to catrocollastatin-C sequence 167–180 (TC1a, 1,533.0 Da) and the disulfide-bonded peptides 167–168 and 168–180 (TC1b, 1,550.8 Da) generated by tryptic cleavage of TC1a (Table 1). This fragment defined a disulfide bond between neighboring cysteine residues 167 and 178. The same disulfide bond combination was found in fragments TC6, TC9, and TC10, which differed from TC1b only in two, one, and one C-terminal residues, respectively (Table 1).

Amino acid and N-terminal sequence analyses of TC8 showed that this fragment corresponded to the sequence 203–216, which includes cysteine residues 204 and 209. The facts that catrocollastatin-C does not have free cysteines, and that the molecular mass of TC8 (1,436.5 Da) is 2 Da less than that calculated for the isotope-averaged molecular mass of the linear sequence 203–216, are only compatible with the existence of an intramolecular disulfide bond in TC8.

The third disulfide bridge between nearest-neighbor cysteine residues 117 and 132 was characterized in fragments TC11 and

Table 1. *Characterization of proteolytic fragments isolated by reversed-phase HPLC of a tryptic/chymotryptic digest of catrocollastatin-C*

Fragment		N-terminal amino acid sequence	Position	Mass (Da)	S-S
TC1	a	CKDNSPGQNNPCKM	167–180	1,533.0	167–178
	b	DNSPGQNNPCKM	169–180	1,550.8	167–178
		CK	167–168		
TC2		NQKGNYY	136–142	886.5	
TC3		ERNQKGNYY	134–142	1,171.5	
TC4		FY	181–182	328.1	
TC5		ERNQK	134–138	673.4	
TC6		CK	167–168		
		DNSPGQNNPCKMFY	169–182	1,861.6	167–178
TC7		GMVLPGTK	190–197		
TC8		VCSNGHCVDVATAY	203–216	1,436.5	204–209
TC9	a	CKDNSPGQNNPCKMF	167–181	1,680.0	167–178
	b	DNSPGQNNPCKMF	169–181	1,697.6	167–178
		CK	167–168		
TC10		DNSPGQNNPCKMF	169–181	1,697.8	167–178
		CK	167–168		
TC11		HQCY	115–118		
		EAEDSCF	127–133	1,347.5	117–132
TC12		HQCY	115–118		
		EAEDSCFERNQKGNYY	127–142	2,500.6	117–132
TC13	a	CYNGNCPIMY	105–114	1,175.5	105–110
	b	NGNCPIMY	107–114	1,192.5	105–110
		CY	105–106		
TC14		SNEDEHK	183–189	857.6	
TC15	a	GYCRKENGNK	143–152		
		VLPGTKCADGK	192–202	2,254.6	145–198
	b	DLFGADVY	119–126	899.0	
	c	CYNGNCPI…AEDSCF	105–133	3,367.2	(105, 110, 117, 132)
TC16		CYNGNCPI…AEDSCF	105–133	3,367.2	(105, 110, 117, 132)
TC17		GYCRK..IPCAPEDVKCGR	143–164		
		VLPGTKCADGK	192–202	3,798.0	(145, 155, 162, 198)
TC18		GYCRKEN…NPCK	143–179		
		VLPGTKCADGK	192–202	5,213.4 ($M + Na^+$)	(145, 155, 162, 198)
TC19	a	IPCAPEDVKCGRL	153–164	1,398.1	155–162
	b	GYCY	103–106		
		NGNCPIMY	107–114	1,413.0	105–110
TC20		IPCAPEDVKCGRL	153–164	1,398.1	155–162
TC21		LGTDIIS…CLDNY	1–102	10,773.0	Disintegrin domain

TC12 (Table 1; Figs. 1, 2). Thus, the latter fragment was built by peptide 115–118 disulfide-bonded to 127–142 (2,500.6 Da). Chymotryptic cleavage of TC12 at the Phe133–Glu144 bond released fragment TC3 (134–142, 1,171.5 Da) (Table 1) and generated fragment TC11 (1,347.5 Da).

Fragment TC13 had major (TC13a, 1,175.5 Da) and minor (TC13b, 1,192.5) quasimolecular ions. Amino acid composition clearly showed that TC13a corresponded to the sequence 105–114. N-terminal sequencing yielded a major (>90%) sequence, XYN GNXPIMY, further confirming the amino acid analysis interpretation. In addition, the presence of the minor sequence NGXXPIMY was clearly identified. This, along with the mass spectrometric result, strongly indicated that TC13b may result from a single cleavage of TC13a at the Y106–N107 bond, thereby producing peptides 105–106 and 107–114 covalently linked by intermolecular cystine residue Cys105–Cys110 (Table 1). Moreover, a minor ion at *m/z* 1,413.0 in fragment TC19 had N-terminal sequence (G + N, Y + G, N, Y, P, I, M, Y) and was assigned to disulfide-bonded peptides 103–106 and 107–114 (Table 1). Thus, fragments TC19b and TC13b are identical except that the former has two extra N-terminal residues. Fragment TC16 had the amino acid composition, N-terminal sequence, and molecular mass corresponding to the polypeptide stretch 105–133 of catrocollastatin-C with two intramolecular disulfide bonds (Table 1). Although the actual pairing of the four cysteines could not be determined, a disulfide bond between Cys117–Cys132 has been characterized in fragments TC11 and TC12 (Table 1). Hence, the combined results from characterization of fragments TC13, TC16, and TC19b unambiguously showed the existence of a disulfide bond between neighboring cysteines 105 and 110 (Figs. 1, 2).

Fragment TC15 had a major ion with a molecular mass of 2,254.6 Da (TC15a, Table 1). N-terminal sequence analysis yielded equi-

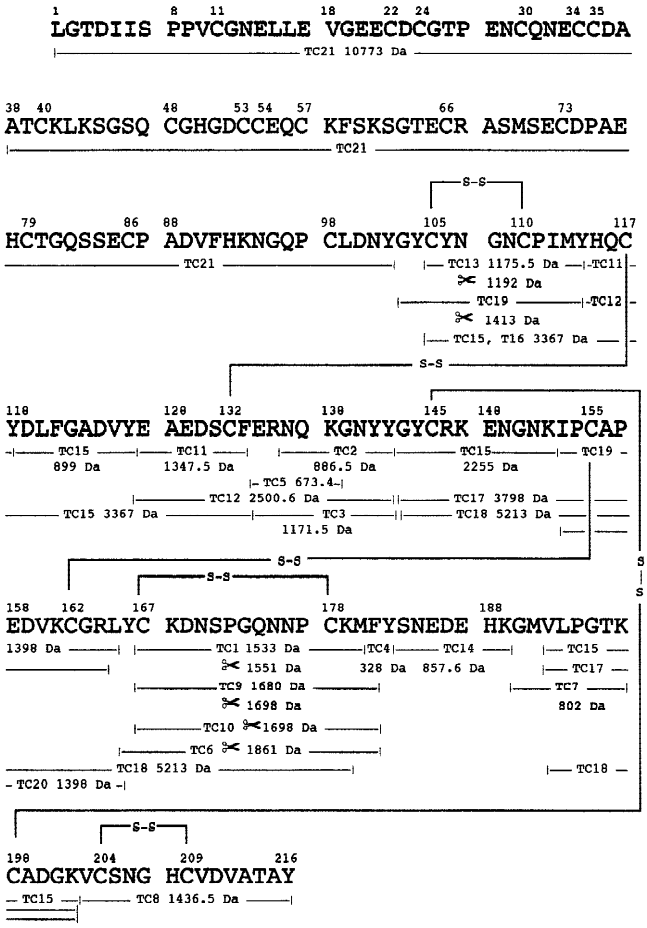

Fig. 1. Amino acid sequence of catrocollastatin-C (Shimokawa et al., 1997) showing the set of tryptic/chymotryptic fragments (TC-) from which the disulfide bond pattern of the cysteine-rich domain (residues 103–216) was established. Molecular masses determined by ESI mass spectrometry are indicated. -S–S–, disulfide bond; ✂, denotes an intramolecular cleavage within that fragment.

molar amounts of two PTH-derivatives in each cycle (except cycles 3 and 7), (G + V), (Y + L), P, (R + G), (K + T), (E + K), N, (G + A)…, which were interpreted as GYCRKENG… and VLPGTKCA… The mass spectrometric and sequencing results are only consistent with peptides 143–152 and 192–202 covalently linked by a disulfide bond involving Cys145 and Cys198 (Table 1; Figs. 1, 2).

Mass spectrometric analysis of fragment TC19 displayed a major ion at *m/z* 1,399, and N-terminal sequencing of this fraction showed the single sequence IPXAPEDVKXGRL. The same results were obtained with fragment TC20 (Table 1). This sequence corresponds to catrocollastatin-C 153–164 and includes cysteine residues 155 and 162. The calculated isotope-averaged molecular mass of the polypeptide stretch 153–164 is 1,401 Da. These data, and the fact that catrocollastatin-C does not have free thiol groups, demonstrated the existence of a disulfide bridge between the two cysteines.

Fragments TC17 and TC18 had the same N-terminal double sequence as TC15a, but their mass spectra showed major ions at *m/z* 3,798.0 and 5,213.4 (Table 1). This data indicated that TC17 and TC18 are derived from peptide 109–202 linked to 143–164 and 143–179, respectively, by two disulfide bonds involving cysteines 145, 155, 162, and 198. The actual pairing of these four cysteine residues could not be deduced, but taking into account the results obtained with fragments TC15a and TC19 (Table 1), they were assigned as 145–198 and 155–162.

Electrospray ionization mass spectrum of fragment TC21 showed three major ions at *m/z* 1,540.1, 1,796.5, and 2,155.5. These species correspond to the series $(M + 7H)^{7+}$, $(M + 6H)^{6+}$, and $(M + 5H)^{5+}$ of a molecule of averaged molecular mass 10,773.0 Da. The single N-terminal sequence LGTDIISPPVXGNELLEVG was obtained for TC21. These data clearly indicated that this fragment corresponds to residues 1–102 of catrocollastatin-C crosslinked by eight intramolecular disulfide bridges (calculated isotope-averaged molecular mass, 10,773.8 Da).

Determination of disulfide bonds of the disintegrin-like domain

To elucidate the pattern of disulfide bonds of the disintegrin-like domain of catrocollastatin-C, peptide TC21 was degraded with oxalic acid. Resulting fragments were separated by reversed-phase HPLC and characterized by MALDI-TOF-MS and N-terminal sequence analysis (Table 2).

Fragment Ox0 eluted as a shoulder of the salt peak and showed valine, glutamic acid, and cysteine by amino acid analysis, and the sequence VQ upon N-terminal sequence analysis (Table 2). These data are only compatible with peptides ^{10}VC^{11} and ^{30}CQ^{31} linked by a disulfide bond (Fig. 3).

Fragment Ox1 had two residues in roughly equimolar amounts in the first two Edman degradation cycles (G + S, T + E), one residue in cycle 3 (E) and traces of proline in cycle 4. No further sequence could be detected. The major sequence in fraction Ox2 was LDNY, although the same amino acid sequences of fraction Ox1 were also clearly present. MALDI-TOF-MS analysis showed an ion at *m/z* 842.4 in both Ox1 and Ox2 (Table 2). These data indicated that peptides ^{63}GTEC^{66} and ^{84}SECP^{87} (calculated isotope-averaged mass (Mav) of 841.9 Da) were disulfide-bonded (Fig. 3).

N-terminal sequencing of fragment Ox3 showed a major and a minor sequence. The major sequence corresponded to catrocollastatin-C ^{5}IISPPV^{10}. The minor sequence, which contained two PTH-amino acid residues in cycles 1 and 3, was (R + P) A (S + L) (M + D) (S + N) E X D. Mass spectrometry showed a prominent ion at *m/z* 1,459.3 Da (Table 2). Together, these data indicated the existence of a disulfide bond between ^{67}RASMSECD^{74} and ^{97}PCLDN^{101} (calculated Mav 1,458.6 Da) (Fig. 3). In agreement with this assignment, a minor ion at *m/z* 1,228.7, which could correspond to ^{67}RASMSECD^{74} disulfide-bonded to ^{97}PCL^{99}, was also detected in fraction 3.

Fraction Ox4 yielded two PTH-amino acids in each of the first three Edman degradation cycles followed by a blank (cycle 4) and small increase of PTH-Lys in cycle 5: (G + A) (E + A) (T + E) X K. This result, in conjunction with MALDI-MS showing an ion at *m/z* 927.6, demonstrated that peptide ^{19}GEEC^{22} was covalently linked to peptide ^{37}AATCK^{41} through cystine Cys22–Cys40 (Table 2; Fig. 3).

Fraction Ox5 contained a mixture of fragments as judged by N-terminal sequence analysis. Amino acid analysis showed this fraction to be particularly enriched in cysteine, suggesting that it could correspond to the cysteine-rich core of the disintegrin-like domain of catrocollastatin-C, which had been extensively and heterogeneously degraded by oxalic acid.

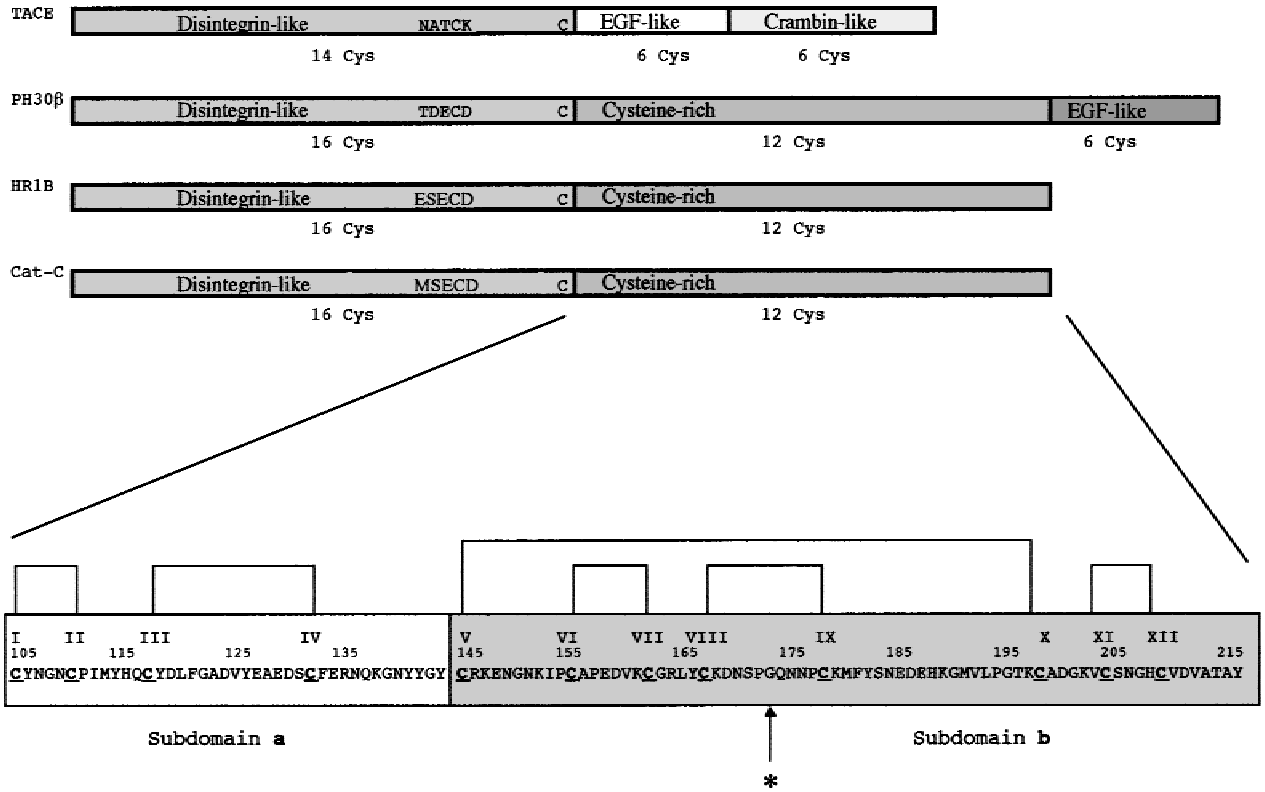

Fig. 2. Schematic representation of the modular structures of catrocollastatin-C, ADAM molecules TACE and PH30*β*, and the SVMP HR1B. The number of cysteine residues of the disintegrin-like, EGF-like, crambin-like, and cysteine-rich domains is indicated. Also shown are the sequences within the disintegrin-like domains that replace the RGD motif of snake venom disintegrins and the two extra cysteine residues that are not present in snake venom soluble disintegrins (one within the sequence that replaces the RGD motif and the other at the C-terminus of the disintegrin-like domain). A cartoon of the cysteine-rich domain of catrocollastatin-C showing its disulfide bond pattern and the proposed division into subdomains a and b (Hite et al., 1994). Cysteine residues are depicted in bold and underlined, and are numbered I–XII. The arrow points to the position of insertion of a putative fusion peptide in ADAM 1, which is marked with an asterisk.

Discussion

Redefinition of the domain structure of catrocollastatin-C

The first 102 residues of catrocollastatin-C fold into a proteolytically (tryptic and chymotryptic) resistant structure, which includes the disintegrin-like domain and what was previously proposed as the first cysteine residue of the cysteine-rich domain. Although the complete disulfide bond pattern of the disintegrin-like domain of catrocollastatin-C was not established, disulfide bonds between Cys11–Cys30, Cys21–Cys40, and Cys66–Cys86 (Fig. 3) are conserved in the structure of bitistatin (Calvete et al., 1997), the closest disintegrin proper homologue of catrocollastatin-C. Despite having conserved cysteine residues, snake venom disintegrins exhibit three different disulfide bond patterns (Calvete, 1997; Calvete et al., 1997; McLane et al., 1998): S-S(I), albolabrin; S-S(II), kistrin, flavoridin, and bitistatin; and S-S(III), echistatin and eristostatin (Fig. 3). Disulfide bond patterns I and II share three disulfides and differ in another three cystine linkages. The disulfide bond between cysteine residues II–VII is characteristic of S-S pattern II, strongly suggesting that catrocollastatin-C may fall into this disulfide bond pattern (Fig. 3).

The most important finding of this work is the demonstration of the existence of a disulfide bond between cysteine residues 73 and 98 (Fig. 3). These two cysteine residues are missing in snake venom disintegrins, which lack a C-terminal cysteine-rich domain, i.e., RGD-containing monomeric (Niewiarowski et al., 1994; Calvete, 1997; Calvete et al., 1997) and non-RGD dimeric disintegrins (Marcinkiewicz et al., 1999a, 1999b), but is characteristically found in the tandem disintegrin-like/cysteine-rich tandem domains of PIII SMVPs and ADAM molecules, the two members of the Re-

Table 2. *Characterization of fragments isolated by reversed-phase HPLC of oxalic acid degradation mixture of the isolated disintegrin-like domain (peptide TC21 from Table 1) of catrocollastatin-C*

Fragment	N-terminal amino acid sequence	Position	Mass (Da)	S-S
Ox0	VC	10–11		
	CQ	30–31		11–30
Ox1	GTEC	63–66		
	SECP	84–87	842.4	66–86
Ox2	LDNY	99–102		
Ox3	IISPPV	5–10		
	RASMSECD	67–74		
	PLCDN	97–101	1,459.3	73–98
Ox4	GEEC	19–22		
	AATCK	37–41	927.6	22–40

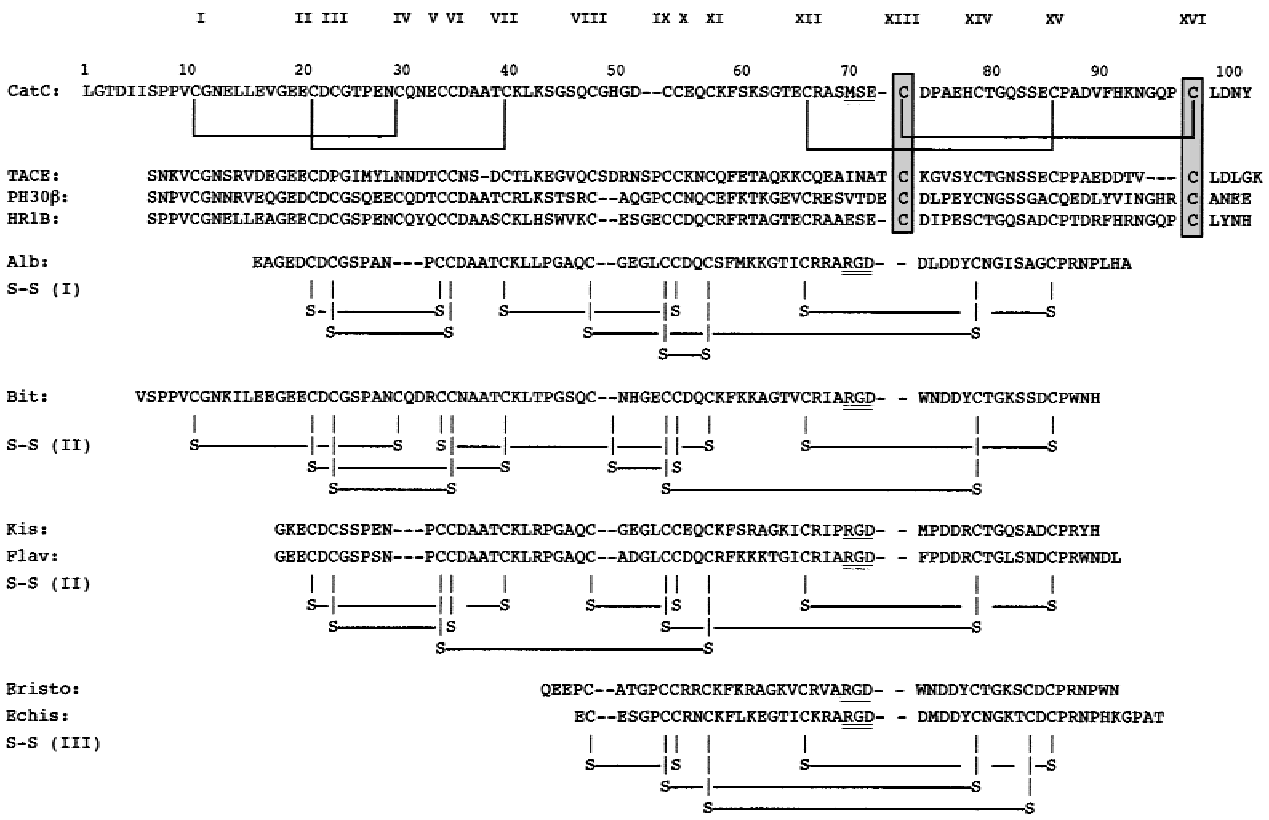

Fig. 3. Comparison of the amino acid sequences of the disintegrin domains of catrocollastatin-C, ADAM molecules TACE and PH30β, the SVMP HR1B, and representative members of the long (bitistatin, Bit), medium [albolabrin (Alb) kistrin (Kis), and flavoridin (Flav)], and short [eristostatin (Eristo) and echistatin (Echis)] RGD-containing disintegrins. Cysteine residues are numbered I–XVI. Disulfide bonds of catrocollastatin-C (this work) are compared with the three different disulfide bond patterns of found in long, medium, and short disintegrins: S-S(I), albolabrin (Calvete et al., 1991; Smith et al., 1996); S-S(II), bitistatin (Calvete et al., 1997), kistrin (Adler et al., 1991), and flavoridin (Calvete et al., 1992; Klaus et al., 1993); and S-S(III), eristostatin (McLane et al., 1996) and echistatin (Calvete et al., 1992; Bauer et al., 1993); bitistatin (Calvete et al., 1997). The two extra cysteines (CysXIII and CysXVI), which form a disulfide bridge and are characteristically found in PIII SVMPs and ADAM molecules, are boxed.

prolysin subfamily of zinc-dependent metalloproteinases (Bjarnason & Fox, 1995; Rawlings & Barrett, 1995; Wolfsberg & White, 1996). In these molecules, cysteine 73 is located just C-terminal of the sequence, which replaces the RGD motif of disintegrins derived from PII class venom metalloproteinases (Wolfsberg & White, 1996) (Fig. 3). Cysteine 98, on the other hand, has been proposed to represent the first cysteine residue of the cysteine-rich domain of Reprolysins. Our results show that it may be considered as part of the disintegrin-like domain of PIII SMVP and ADAM Reprolysin molecules. Hence, we propose to redefine the structural organization of catrocollastatin-C as an N-terminal (residues 1–102) disintegrin-like domain containing eight cystine residues, followed by a C-terminal cysteine-rich domain spanning residues 103–216, including 12 cysteine residues arranged into 6 disulfide bonds (Figs. 1, 2).

Comparison of the structures of catrocollastatin-C and the homologous domains from other modular proteins of the ADAM (A Disintegrin-like And Metalloproteinase) family (Wolfsberg & White, 1996; Fox & Long, 1998) is depicted in Figure 2. The 16 cysteines of the disintegrin-like domain and the 12 cysteines of the cysteine-rich domain are conserved in other P-III venom metalloproteinases, such as the hemorrhagic protein HR1B, isolated from the venom of *Trimeresurus flavoviridis* (Takeya et al., 1990) (see Fig. 8 in Shimokawa et al., 1997), as well as proteins of the ADAM family, i.e., PH30β (fertilin β), a type-I membrane protein of mouse spermatozoa (Stone et al., 1999). However, unlike other mammalian ADAM family members, which exhibit a cysteine-rich domain and an EGF-like domain, in TACE (Tumour-necrosis factor α Converting Enzyme, ADAM 17), the disintegrin-like domain is followed by an EGF-like domain and a crambin-like domain (Black et al., 1997; Moss et al., 1997). Noteworthy, the last cysteine residue of the disintegrin domain of TACE is topologically equivalent of the 16th cysteine residue of catrocollastatin-C. This is the cysteine that has been considered as the first cysteine of the cysteine-rich domain, and that we propose to redefined as the last cysteine of the disintegrin-like domain. Therefore, this supports our structural two-domain model. The new domain definition may be relevant for the recombinant expression of individual domains and for designing molecular biology experiments to test structure–function correlations. Previously, Hite et al. (1994) observed that the characteristic sequence pattern of the cysteine-rich domain of Reprolysins can be subdivided into two subdomains. Subdomain a, comprising the N-terminal 48 residues of the cysteine-rich domain, shows almost absolute conservation of the cysteine pattern (C X_4 C X_6 C X_{14} C X_{12}), while subdomain b, encompassing the C-terminal 62–72 residues, is less stringent in this

regard (C X_9 C X_6 C X_4 C X_{4-10} C X_{15-19} C X_5 C X_4 C X_{7-16}). Our results provide structural evidence for this subdomain classification. Thus, subdomain a contains two disulfide bonds between nearest-neighbor cysteine residues and is connected by a 12-residue polypeptide spacer to subdomain b, which is cross-linked by three short disulfide loops and a long-range disulfide bond (Fig. 2).

As the cysteine-rich domain is always present in conjunction with the disintegrin-like domain, its function could be as a spacer domain aiding in the correct positioning of the metalloproteinase and disintegrin-like domains. The hypothesis has been put forward that a relatively hydrophobic polypeptide stretch embedded in the cysteine-rich domain of fertilin α (ADAM 1), showing striking similarity to viral fusion peptide sequences, could play a role during sperm–egg plasma membrane fusion at fertilization (Blobel et al., 1992). Alignment of the cysteine-rich domains of ADAMs with those of PIII SVMPs indicate that the latter are lacking about 20 residues that correspond to the putative fusion peptide. This sequence, which in guinea pig fertilin α is LICTGISSIPPIRAL FAAIQIP, is located in subdomain b between the eighth and ninth cysteine residues, which are linked by a disulfide bond (Fig. 2). The inserted putative fusion peptide is thus likely to form a loop structure. However, no ADAM molecule has been formally shown to possess bona fide membrane fusion activity. Another recent report indicates that the recombinant subdomain b of the cysteine-rich domain of human ADAM 12 supports cell adhesion of a panel of carcinoma cell lines by engaging a cell surface heparan sulfate proteoglycan receptor (Iba et al., 1999). Also, recombinant cysteine-rich domain from atrolysin a supports collagen stimulated platelet aggregation (Jia et al., 2000). These reports clearly suggest that there may reside functional motifs in the cysteine-rich domain that contribute to the overall biological activity of the protein.

A structural model for the disintegrin-like domain of Reprolysins

Although the complete pairing of all 16 cysteine residues of the disintegrin-like domain of catrocollastatin-C has not been determined, we hypothesize that it will be the same as in the long snake venom disintegrin bitistatin (Calvete et al., 1997), with the two extra cysteines forming an additional disulfide bond (Cys73–Cys98, Fig. 3). The conservation of cysteine residues in cellular disintegrin-like domain-containing proteins strongly suggests that this will also be the case for other ADAM family member. If this holds true, the conformation of the putative integrin binding sequence is restricted by three disulfide bonds, CysIX–CysXIV, CysXII–CysXV, and CysXIII–CysXVI (Fig. 3). Zhang and coworkers (Zhang et al., 1998) have shown that the disintegrin-like domain of Metargidin (ADAM-15), which possesses an RGDC sequence, when expressed as a soluble, fusion protein with glutathione S-transferase in bacteria, interacts specifically with integrin $\alpha_v\beta_3$. In agreement with this paper, Iba et al. (1999) showed that $\alpha_v\beta_3$ integrin-expressing A375 melanoma cells were strongly adherent to recombinant ADAM-15 disintegrin-like domain, whereas MDA-MB-231 cells, which express only low levels of $\alpha_v\beta_3$ integrin, did not adhere to the protein. These data indicated that despite the extra disulfide bond involving the cysteine residue flanking the RGD sequence, this motif is accessible for engaging an integrin receptor. However, the role of the extra disulfide bond remains ambiguous. Evans et al. (1995) reported that a linear peptide from guinea pig fertilin β (CAQDEC) inhibited sperm–egg binding in a dose-dependent manner, whereas the cyclized peptide did not have inhibitory activity. Similarly, Pyluck et al. (1997) have shown that only linear and not cyclic synthetic peptides from the putative disintegrin-like active loop sequence of mouse fertilin β (ECD), inhibited in vitro fertilization at 500 μM concentrations. The authors suggested that the red-ox state of the "extra pair" of cysteines may distinctly affect the biological activity of the disintegrin-like domain of different ADAM molecules. On the other hand, Jia et al. (1997), using synthetic peptides, have provided evidence that the region of the disintegrin-like domain of atrolysin A that is positionally analogous to the RGD loop of disintegrins (RSECD) is capable of inhibiting collagen-induced platelet aggregation (i.e., inhibition of integrin $\alpha_2\beta_1$). For these peptides to have significant inhibitory activity, the cysteine residue must be constrained by participation in a disulfide bond. Additionally, the two acidic amino acids flanking the cysteine residue were found to be important for biological activity (Jia et al., 1997). However, it is not clear how the metalloproteinase, disintegrin-like, and cysteine-rich domains of ADAMs and SVMPs are spatially arranged, and how the molecular architecture of these multidomain proteins affects the protein binding capabilities of individual domains.

A hypothesis for the evolution of disintegrin domains

Phylogenetic analysis suggests that mammalian ADAMs and SVMPs have evolved from a common ancestor gene, already assembled as the multidomain structure, both by speciation and gene duplication (Moura-da-Silva et al., 1996). Disintegrins, and the short chain hemorrhagins arose from SVMPs late during evolution only after mammals and reptiles diverged (Moura-da-Silva et al., 1996). PIII SVMPs, composed of metalloproteinase, disintegrin-like and cysteine-rich domains, are the closest homologues of cellular ADAMs, and may have evolved from a common ancestor after loosing the membrane anchor. All known PIII SVMPs contain a disintegrin(-like) domain with eight disulfide bonds, followed by a cysteine-rich region. PII SVMPs contain disintegrin domains with seven disulfide bonds, and lack the cysteine-rich domain. Comparison of cDNAs coding for PII and PIII SVMPs precursors clearly indicates that the occurrence in PIII SVMPs of CysXIII and CysXVI is linked to the occurrence of a C-terminal cysteine-rich region. Here we hypothesize that the loss of the disulfide bond between cysteines XIII and XVI (Fig. 3), along with the loss of the cysteine-rich domain and emergence of the RGD sequence, gave rise to PII SVMPs, from which disintegrins are released proteolytically in snake venoms. The presence of these soluble disintegrins, capable of disrupting platelet aggregation, would certainly enhance the hemorrhagic potency of the venom and contribute to the hemorrhagic shock important for prey immobilization.

Materials and methods

Purification of catrocollastatin-C

Lyophilized crude venom from *Crotalus atrox* was purchased from Miami Serpentarium Laboratories (Salt Lake City, Utah). Catrocollastatin-C was purified from the crude venom by DE-52 ion-exchange chromatography followed by Sephadex G-50 gel filtration chromatography, as described (Shimokawa et al., 1997). The protein migrated as a single 32 kDa band in SDS-polyacrylamide gels

and had the single N-terminal sequence LGTDIISPPVXG (determined using an Applied Biosystems Procise 494 instrument following the manufacturer's operating protocols).

Determination of thiol groups and disulfide bonds

For quantitation of free cysteine residues and disulfide bonds in catrocollastatin-C, the protein (1 mg/mL in 150 mM Tris/HCl, pH 8.6, 1 mM ethylenediaminetetraacetic acid, 6 M guanidine hydrochloride) was incubated with either 10 mM iodoacetamide for 1 h at room temperature, or with 1% 2-mercaptoethanol for 2 min at 100 °C, followed by addition of a fivefold molar excess of 4-vinylpyridine or iodoacetamide over reducing agent and incubation for 1 h at room temperature. Samples were dialyzed against deionized (MilliQ) water and lyophilized. Aliquots of these samples were subjected to matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) using a PE Biosciences Voyager DE-Pro instrument and alpha-cyano-4-hydroxycinnamic acid as matrix. Samples were hydrolyzed with 6 M HCl at 110 °C for 24 h and subjected to amino acid analysis using an Alpha Plus amino acid analyzer (Pharmacia, Uppsala, Sweden).

Proteolytic cleavage and isolation and characterization of fragments

Catrocollastatin-C (2 mg/mL in 100 mM NH_4HCO_3, pH 8.6, 1 M guanidine hydrochloride) was digested with a mixture of TPCK-trypsin and α-chymotrypsin (Sigma, St. Louis, Missouri) (1:100 w/w enzyme:substrate ratio) for 18 h at 37 °C. Proteolytic peptides were isolated by reversed-phase HPLC on a Lichrospher RP100 C18 column (25 × 0.4 cm, 5 μm particle size) eluting at 1 mL/min with a mixture of 0.1% (v/v) trifluoroacetic acid in water (solvent A) and acetonitrile (solvent B) employing the following chromatographic conditions: first, isocratic (5% B) for 5 min, followed by gradients of 5–10% B for 5 min, 10–35% B for 100 min, and 35–70% B for 15 min. Peptide elution was monitored at 220 nm, and chromatographic fractions were collected manually and characterized by amino acid analysis (as above), N-terminal sequencing (using an Applied Biosystems Procise 494 instrument following the manufacturer's instructions), and electrospray ionization (ESI) mass spectrometry (using a Sciex API-III+ triple quadrupole mass spectrometer).

Oxalic acid cleavage of the disintegrin-like domain and isolation of fragments

Purified disintegrin-like domain (0.15 mg) of catrocollastatin-C (fraction TC21 of the reversed-phase HPLC separation of tryptic/chymotryptic peptide mixture corresponding to catrocollastatin-C residues 1–102, see Table 1) (10 mg/mL in water) was degraded with 250 mM oxalic acid (final concentration) for 4 h and 12 h at 100 °C in sealed, evacuated ampoules (Bauer et al., 1993). Hydrolyzates were dried in a Speed-Vac, dissolved in 0.1% TFA in 5% aqueous acetonitrile and fragments were isolated by reversed-phase HPLC on a Lichrospher RP100 C18 (250 × 4 mm, 5 μm particle size) column eluting at 1 mL/min with 0.1% (v/v) TFA in water (solvent A) and 0.1% (v/v) TFA in acetonitrile (solvent B). Peptides were eluted isocratically (5% B) for 5 min, followed by a gradient of 5–50% B for 90 min, and 50–70% B for 20 min. Elution was monitored by ultraviolet absorption at 220 nm, and chromatographic fractions were collected manually. Chromatographic fractions were characterized by amino acid analysis, N-terminal sequencing, and mass spectrometry.

Amino acid analysis and N-terminal sequencing

Oxalic acid-derived fragments were subjected to amino acid analysis using a Pharmacia AlphaPlus analyzer (after sample hydrolysis with 6 M HCl at 110 °C for 18 h in evacuated and sealed ampoules) and N-terminal amino acid sequence analysis [using an Applied Biosystems Procise 494 sequencer and a Porton LF3000 (Beckman) instrument] following the manufacturer's protocols.

MALDI-TOF-MS

Aliquots of 0.5 μL of reversed-phase HPLC fractions were applied automatically using the pipetting robot SymBiot I (PerSeptive Biosystems, Framingham, Massachusetts) to the inner 8 × 8 positions of a 10 × 10-position polished stainless steel sample plate. Each sample was mixed on the plate with the same volume of a matrix composed of 5–10 μg/μL of each alpha-cyano-4-hydroxycinnamic acid (CHC) and 6-desoxy-L-galactose (L(-)fucose) (Sigma-Aldrich, Deisenhofen, Germany) in a mixture of 1:1 (v/v) acetonitrile: 0.1% TFA in water. Samples were dried at ambient temperature using a microventilator. Measurements were performed in reflector mode with a Voyager-DE STR (PE Biosystems, Foster City, California) instrument equipped with a 2 m flight tube (3 m in reflector mode) and a 337 nm nitrogen laser. Positive ions were accelerated at 20 kV, and 256 laser shots were accumulated for a single spectrum.

Acknowledgments

This work was financed by grants PB97-1237 and PB98-0694 from the Dirección General de Investigación Científica y Técnica (DGICYT), Madrid, Spain.

References

Adler M, Lazarus RA, Dennis MS, Wagner G. 1991. Solution structure of kistrin, a potent platelet aggregation inhibitor and GPIIb-IIIa antagonist. *Science 253*:445–448.

Au LC, Huang YB, Huang TF, Teh GW, Lin HH, Choo KB. 1991. A common precursor for the 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.

Baramova EN, Shannon JD, Bjarnason JB, Fox JW. 1989. Degradation of extracellular matrix proteins by hemorrhagic metalloproteinases. *Arch Biochem Biophys 275*:63–71.

Bauer M, Sun Y, Degenhardt C, Kozikowski B. 1993. Assignment of all four disulfide bridges in echistatin. *J Protein Chem 12*:759–764.

Bjarnason JB, Fox JW. 1994. Hemorrhagic metalloproteinases from snake venoms. *Pharmacol Ther 62*:325–372.

Bjarnason JB, Fox JW. 1995. Snake venom metalloendopeptidases: Reprolysins. *Methods Enzymol 248*:345–368.

Black RA, Rauch CT, Kozlosky CJ, Peschon JJ, Slack JL, Wolfson MF, Castner BJ, Stocking KL, Reddy P, Srinivasan S, et al. 1997. A metalloproteinase disintegrin that releases tumor-necrosis factor-α from cells. *Nature 385*: 729–733.

Blobel CP, Wolfsberg TG, Turck CW, Myles DG, Primakoff P, White JM. 1992. A potential fusion peptide and an integrin ligand domain in a protein active in sperm-egg fusion. *Nature 356*:248–251.

Calvete JJ. 1997. Snake venom disintegrins and disintegrin-like domains: Soluble antagonists and cellular ligands of integrin receptors. In: Eble J, Kühn K, eds. *Integrin–ligand interactions*. Austin, Texas: Landes Co. Biomedical Publishers. pp 157–173.

Calvete JJ, Schäfer W, Soszka T, Lu W, Cook JJ, Jameson BA, Niewiarowski S. 1991. Identification of the disulfide bond pattern in albolabrin, an RGD-

containing peptide from the venom of *Trimeresurus albolabris*: Significance for the expression of platelet aggregation inhibitory activity. *Biochemistry 30*:5225–5229.

Calvete JJ, Schrader M, Raida M, McLane MA, Romero A, Niewiarowski S. 1997. The disulfide bond pattern of bitistatin, a disintegrin isolated from the venom of the viper *Bitis arietans*. *FEBS Lett 418*:197–202.

Calvete JJ, Wang Y, Mann K, Schäfer W, Niewiarowski S, Stewart GJ. 1992. The disulphide bridge pattern of snake venom disintegrins, flavoridin and echistatin. *FEBS Lett 309*:316–320.

Chen H, Pyluck AL, Janik M, Sampson NS. 1998. Peptides corresponding to the epidermal growth factor-like domain of mouse fertilin: Synthesis and biological activity. *Biopolymers 47*:299–307.

Chen H, Sampson NS. 1999. Mediation of sperm–egg fusion: Evidence that mouse egg $\alpha_6\beta_1$ integrin is the receptor for sperm fertilin β. *Chem Biol 6*:1–10.

Cho C, Bunch D, Faure J-E, Goulding EH, Eddy EM, Primakoff P, Myles DG. 1998. Fertilization defects in sperm from mice lacking fertilin β. *Science 281*:1857–1859.

Evans JP, Schultz RM, Kopf GS. 1995. Mouse sperm–egg plasma membrane interactions: Analysis of roles of egg integrins and the mouse homologue of PH-30 (fertlin) β. *J Cell Sci 108*:3267–3278.

Evans JP, Schultz RM, Kopf GS. 1998. Roles of the disintegrin domains of mouse fertilins α and β in fertilization. *Biol Reprod 59*:145–152.

Fox JW, Bjarnason JB. 1995. Snake venom metalloendopeptidases: Reprolysins. *Methods Enzymol 248E*:369–387.

Fox JW, Long C. 1998. The ADAMs/MDC family of proteins and their relationships to the snake venom metalloproteinases. In: Bailey GS, ed. *Enzymes from snake venom*. Fort Collins, Colorado: Alanken, Inc. pp 151–178.

Gichuhi PM, Ford WC, Hall L. 1997. Evidence that peptides derived from the disintegrin domain of primate fertilin and containing the ECD motif block the binding of human spermatozoa to the zona-free hamster oocyte. *Int J Androl 20*:165–170.

Hite LA, Jia L-G, Bjarnason JB, Fox JW. 1994. cDNA sequences for four snake venom metalloproteinases: Structure, classification, and their relationship to mammalian reproductive proteins. *Arch Biochem Biophys 308*:182–191.

Iba K, Albrechtsen R, Gilpin BJ, Loechel F, Wewer UM. 1999. Cysteine-rich domain of human ADAM 12 (meltrin α) supports tumor cell adhesion. *Am J Pathol 154*:1489–1501.

Ivaska J, Kapyla J, Pentikainen O, Hoffrén AM, Hermonen J, Huttunen P, Johnson MS, Heino J. 1999. A peptide inhibiting the collagen binding function of integrin α_2I domain. *J Biol Chem 274*:3513–3521.

Jia L-G, Shimokawa K-I, Bjarnason JB, Fox JW. 1996. Snake venom metalloproteinases: Structure, function and relationship to the ADAMs family of proteins. *Toxicon 34*:1269–1276.

Jia L-G, Wang X-M, Shannon JD, Bjarnason JB, Fox JW. 1997. Function of disintegrin-like/cysteine-rich domains of atrolysin A. Inhibition of platelet aggregation by recombinant protein and peptide antagonists. *J Biol Chem 272*:13094–13102.

Jia L-G, Wang X-M, Shannon JD, Bjarnason JB, Fox JW. 2000. Inhibtion of platelet aggregation by the recombinant cysteine-rich domain of the hemorrhagic snake venom metalloproteinase, atrolysin a. *Arch Biochem Biophys 373*:281–286.

Kini RM, Evans HJ. 1992. Structural domains in venom proteins: Evidence that metalloproteinases and nonenzymatic platelet aggregation inhibitors (disintegrins) from snake venoms are derived by proteolysis from a common precursor. *Toxicon 30*:265–293.

Klaus W, Broger C, Gerber P, Senn H. 1993. Determination of the disulphide bonding pattern in proteins by local and global analysis of nuclear magnetic resonance data. Application to flavoridin. *J Mol Biol 232*:897–906.

Marcinkiewicz C, Calvete JJ, Marcinkiewicz MM, Raida M, Vijay-Kumar S, Huang Z, Lobb RR, Niewiarowski S. 1999a. EC3, a novel heterodimeric disintegrin from *Echis carinatus* venom, inhibits α4 and α5 integrins in an RGD-independent manner. *J Biol Chem 274*:12468–12473.

Marcinkiewicz C, Calvete JJ, Vijay-Kumar S, Marcinkiewicz MM, Raida M, Schick P, Lobb RR, Niewiarowski S. 1999b. Structural and functional characterization of EMF10, a heterodimeric disintegrin from *Eristocophis macmahoni* venom that selectively inhibits $\alpha5\beta1$ integrin. *Biochemistry 38*: 13302–13309.

Markland FS. 1998. Snake venoms and the hemostatic system. *Toxicon 36*: 1749–1800.

McLane MA, Marcinkiewicz C, Vijay-Kumar S, Wierzbicka-Patynowski I, Niewiarowski S. 1998. Viper venom disintegrins and related molecules. *Proc Soc Exp Biol Med 219*:109–119.

McLane MA, Vijay-Kumar S, Marcinkiewicz C, Calvete JJ, Niewiarowski S. 1996. Importance of the structure of the RGD-containing loop in the disintegrins echistatin and eristostatin for the recognition of $\alpha_{IIb}\beta_3$ and $\alpha_v\beta_3$ integrins. *FEBS Lett 391*:139–143.

Moss ML, Jin S-LC, Milla ME, Burkhart W, Carter HL, Chen W-J, Clay WC, Didsbury JR, Hassler D, Hoffman CR, et al. 1997. Cloning of disintegrin metalloproteinase that processes precursor tumor-necrosis factor-α. *Nature 385*:733–736.

Moura-da-Silva AM, Theakston RDG, Crampton JM. 1996. Evolution of disintegrin cysteine-rich and mammalian matrix-degrading metalloproteinases: Gene duplication and divergence of a common ancestor rather than convergent evolution. *J Mol Evol 43*:263–269.

Niewiarowski S, McLane MA, Kloczewiak M, Stewart GJ. 1994. Disintegrins and other naturally occurring antagonists of platelet fibrinogen receptor. *Semin Hematol 4*:289–300.

Paine MJI, Desmond HP, Theakston RDG, Crampton JM. 1992. Purification, cloning, and molecular characterization of a high molecular weight hemorrhagic metalloproteinase, jararhagin, from *Bothrops jararaca* venom. Insights into the disintegrin gene family. *J Biol Chem 267*:22869–22876.

Perry ACF, Jones R, Barker PJ, Hall L. 1992. A mammalian epididymal protein with remarkable sequence similarity to snake venom haemorrhagic peptides. *Biochem J 286*:671–675.

Pyluck A, Yuan R, Galligan E Jr, Primakoff P, Myles DG, Sampson NS. 1997. ECD peptidesinhibit *in vitro* fertilization in mice. *Bioorg Med Chem Lett 7*:1053–1058.

Rawlings ND, Barrett AJ. 1995. Evolutionary families of the metallopeptidases. *Methods Enzymol 248E*:183–228.

Shimokawa K, Jia L-G, Shannon JD, Fox JW. 1998. Isolation, sequence analysis, and biological activity of atrolysin E/D, the non-RGD disintegrin domain from *Crotalus atrox* venom. *Arch Biochem Biophys 354*:239–246.

Shimokawa K-I, Jia L-G, Wang X-M, Fox JW. 1996. Expression, activation, and processing of the recombinant snake venom metalloproteinase, pro-atrolysin E. *Arch Biochem Biophys 335*:283–294.

Shimokawa K-I, Shannon JD, Jia L-G, Fox JW. 1997. Sequence and biological activity of catrocollastatin-C: A disintegrin-like/cysteine-rich two domain protein from *Crotalus atrox* venom. *Arch Biochem Biophys 343*:35–43.

Smith KJ, Jaseja M, Lu X, Williams JA, Hyde EI, Trayer IP. 1996. Three-dimensional structure of the RGD-containing snake toxin albolabrin in solution, based on ^{1}H NMR spectroscopy and simulated annealing calculations. *Int J Peptide Protein Res 48*:220–228.

Stone AL, Kroeger M, Sang QXA. 1999. Structure–function analysis of the ADAM family of disintegrin-like and metalloproteinase-containing proteins. *J Protein Chem 18*:447–465.

Takeya H, Oda K, Miyata T, Omori-Satoh T, 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.

Tsai I-H, Wang Y-M, Lee Y-H. 1994. Characterization of a cDNA encoding the precursor of platelet aggregation inhibitor and metalloproteinase from *Trimeresurus mucrosquamatus* venom. *Biochim Biophys Acta 1200*:337–340.

Wolfsberg TG, White JM. 1996. ADAMs in fertilization and development. *Dev Biol 180*:389–401.

Zhang X-P, Kamata T, Yokoyama K, Puzon-McLaughlin W, Takada Y. 1998. Specific interaction of the recombinant disintegrin-like domain of MDC-15 (metargidin, ADAM-15) with integrin $\alpha_v\beta_3$. *J Biol Chem 273*:7345–7350.

Zhou Q, Smith JB, Grossman MH. 1995. Molecular cloning and expression of catrocollastatin, a snake-venom protein from *Crotalus atrox* (western diamondback rattlesnake) which inhibits platelet adhesion to collagen. *Biochem J 307*:411–417.