# Molecular cloning and expression of catrocollastatin, a snake-venom protein from *Crotalus atrox* (western diamondback rattlesnake) which inhibits platelet adhesion to collagen

Qing ZHOU,* J. Bryan SMITH* and Mark H. GROSSMAN†

*Department of Pharmacology, Temple University School of Medicine, Philadelphia, PA 19140, and †Department of Pediatrics, Temple University School of Medicine, St. Christopher's Hospital for Children, Philadelphia, PA 19134, U.S.A.

A 50 kDa protein that inhibits platelet adhesion to collagen has been isolated from snake venom of *Crotalus atrox* (western diamondback rattlesnake) and has been named 'catrocollastatin'. The cDNA cloning of catrocollastatin has been accomplished. A full-length cDNA of 2310 bp with an open reading frame between nucleotides 51 and 1880 was obtained. The deduced amino acid sequence consists of 609 amino acids. The cDNA-predicted amino acid sequence is highly similar to that of haemorrhagic metalloproteinase jararhagin from *Bothrops jararaca* venom, HR1B from *Trimeresurus flavoviridis*, Ht-e from *C. atrox* and trigramin from *T. gramineus*. Like jararhagin and HR1B, catrocollastatin is a multidomain molecule composed of an N-terminal domain, a metalloproteinase domain, a disintegrin-like domain and a cysteine-rich C-terminal domain. In the disintegrin-like domain, the frequently seen RGD (Arg-Gly-Asp) sequence is replaced by SECD (Ser-Glu-Cys-Asp). This cDNA was expressed in *Spodoptera frugiperda* (fall armyworm) (Sf9) insect cells using a baculovirus expression system. Like native catrocollastatin, the expressed protein is capable of selectively blocking collagen-induced platelet aggregation. This is the first full-length clone of a high-molecular-mass haemorrhagin to be expressed.

## INTRODUCTION

Among families of poisonous snakes, the Viperidae (vipers) and the Crotalidae (pit vipers) are those whose venoms are mainly composed of haemorrhagic toxins. The occurrence of haemorrhage is one of the most striking consequences of the parenteral injection of these venoms [1]. The toxic haematological components affect blood platelets, vessel wall, fibrinogen and other blood-coagulation factors [2]. The most intensively studied snake-venom haemorrhagic proteins are disintegrins and haemorrhagins.

Disintegrins are a family of low-molecular-mass cysteine-rich peptides containing the sequence RGD (Arg-Gly-Asp) isolated from the venom of various snakes [3,4], such as echistatin from *Echis carinatus* (saw scaled viper) [5], trigramin from *Trimeresurus gramineus* [6], halysin (also known as kistrin) from *Agkistrodon halys* (mamushi) [7] and trimucin from *T. mucrosquamatus* [8]. Both *in vivo* and *in vitro* experiments have demonstrated that disintegrins are potent inhibitors of platelet aggregation [3]. The sequence RGD is conserved in all of these peptides and is believed to specifically bind to the platelet surface fibrinogen receptor GP IIb/IIIa (glycoprotein IIb/IIIa, $\alpha_{IIb}\beta_3$) complex [9], which results in the inhibition of fibrinogen-dependent platelet aggregation. On the basis of this discovery, peptides have been designed to act as anti-thrombotic drugs which mimic disintegrins [10], some of which show very high anti-thrombotic potency [11] and high specificity for GP IIb/IIIa–fibrinogen interaction [12,13].

Haemorrhagins are snake-venom proteins belonging to a unique subfamily of metalloproteinases which contain a conserved $Zn^{2+}$-chelating sequence, H-E-X-X-H, but which have no other similarities to other metalloproteinases [14]. Haemorrhagins can be classified into two groups based on their molecular size. Small haemorrhagins have molecular masses of about 25 kDa. HT-2 from *Crotalus ruber ruber* (red diamond rattlesnake) venom [15] and H2-proteinase [16], Ht-b, Ht-c, Ht-d, and Ht-e from *C. atrox* [17] are among the small haemorrhagins that have been isolated and sequenced. The molecular masses of the large haemorrhagins range between 50 and 90 kDa. HR1B from *T. flavoviridis* venom [18], jararhagin from *Bothrops jararaca* venom [19] and Ht-a from *C. atrox* [20] are high-molecular-mass haemorrhagins that have been cloned and sequenced. They are mosaic proteins with a metalloproteinase domain similar to that of the small venom metalloproteinases, at the N-terminus, a disintegrin-like domain in the middle of the molecule and a cysteine-rich domain at the C-terminus. The metalloproteinases attack components of the basement membrane of the vessel wall and result in cutaneous or subcutaneous bleeding in snake-bite victims [2].

Injury to the endothelium, followed by platelet adhesion to subendothelial collagen with the release of platelet-derived growth factor, is believed to be one of the mechanisms of arteriosclerosis. Thus inhibition of adhesion is a possible route to prevention of thrombosis. In an attempt to identify an inhibitor of platelet–collagen adhesion, a number of snake venoms were screened [21] and venoms of *B. atrox*, *B. jararaca*, *A. blomhoffii* (copperhead or cottonmouth) and *C. basciliscus* (Mexican west-coast rattlesnake) were shown to have inhibitory activity. The active proteins from the four inhibitory snake venoms were purified by Sephadex G-100 chromatography and reverse-phase HPLC. On SDS/PAGE they all had an apparent molecular mass of 50 kDa. These proteins totally abolished platelet aggregation induced by collagen, but had no effect on platelet aggregation induced by ADP, 9,11-azoprostaglandin $H_2$ or platelet activating factor.

Abbreviations used: NBT, Nitroblue Tetrazolium chloride; BCIP, 5-bromo-4-chloroindol-3-yl phosphate; Sf9, *Spodoptera frugiperda* (fall armyworm) ovarian-derived.

† To whom correspondence and reprint requests should be addressed.

The nucleotide sequence reported in this paper has been submitted to the GenBank/EMBL Nucleotide Sequence Database under the accession number U21003.

In the present study, a 50 kDa snake-venom protein from *C. atrox* was isolated and purified. Bioassay showed the same inhibitory effect on platelet–collagen adhesion as the other 50 kDa proteins. It was named 'catrocollastatin' to indicate its source and function (the nomenclature is based on the method suggested by Pirkle and Marsh [22]). The cDNA cloning and functional expression of catrocollastatin were accomplished and are presented here. The primary structure of the peptide deduced from the cDNA sequence reveals that it belongs to a family of large metalloproteinases. This is the first-full length clone of a high-molecular-mass haemorrhagin to be expressed.

## EXPERIMENTAL

### Platelet adhesion and aggregation assays

The adhesion and aggregation assays were performed using human platelets labelled with [$^3$H]oleic acid as described previously [21]. Briefly, microtitre plates were coated with 1 $\mu$g of collagen per well. A 50 $\mu$l portion of a suspension of [$^3$H]oleic acid-labelled human platelets ($2.5 \times 10^8$ cells/ml) were added and incubated for 1 h at room temperature. Iloprost (a $PGI_2$ analogue) (28 nM) and Arg-Gly-Asp-Ser (100 $\mu$g/ml) were used to prevent further aggregation. Subsequently, the platelets remaining in suspension were removed and those adhering to the wells were solubilized with 2% (w/v) SDS; radioactivity was then determined by liquid-scintillation counting. For platelet-aggregation studies, the snake-venom protein fraction was added to platelet-rich plasma in an aggregometer 1 min before the addition of collagen, ADP or the thromboxane analogue U46619 and incubation was continued for 3 min at 37 °C.

### Purification of catrocollastatin

Venom was harvested from *C. atrox* glands (Biotoxins Inc., St. Cloud, FL, U.S.A.) by standard techniques. The venom was fractionated by Sephadex G-100 chromatography and the fractions analysed for their ability to inhibit platelet–collagen adhesion. The partially purified protein was further purified by reverse-phase HPLC on a $C_{18}$ column, where biological activity was eluted at approx. 60% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid. This anti-adhesive protein was subjected to SDS/PAGE after each step of purification to check the molecular mass and purity.

### Determination of peptide sequences and oligonucleotide probe synthesis

The purified protein was cleaved with CNBr in 5 M guanidinium chloride/70% (v/v) formic acid by overnight incubation at room temperature. Two cleavage fragments, peptide 1 and peptide 2, with the molecular mass of 8 and 13 kDa respectively, were obtained by HPLC of the CNBr digest and were sequenced. The 40 N-terminal residues of peptide 1 were:

YIHVALVGLEIWSNEDKITVKPEAGYTLNAFGEWRKT-DLL

The partial sequence of peptide 2 was:

VLQCYDLFGGDVVEAEDDCFERNQKGNYYGY

and a degenerate 48-mer oligonucleotide probe was prepared based on the underlined sequence:

5′-GAA(G)GAC(T)GAC(T)TGC(T)TTC(T)GAA(G)CGG(A)-AAC(T)CAA(G)AAA(G)GGA(G)AAC(T)TAT(C)TAT(C)TT-A(C)GGA(G)-3′

This probe was end-labelled to high specific radioactivity using $T_4$ polynucleotide kinase and [$\gamma$-$^{32}$P]ATP.

### Generation of polyclonal antiserum

Antiserum was raised in rabbits by injecting them subcutaneously with a mixture of the reduced (10 mM dithiothreitol) and non-reduced forms of purified 50 kDa protein emulsified in Freund's adjuvant.

### cDNA library construction

A custom cDNA library was constructed by the Stratagene Corporation (La Jolla, CA, U.S.A. ) from the *C. atrox* venom gland in the Uni-ZAP XR vector. The library contained $6.8 \times 10^6$ primary plaques, with an estimated titre of $1.2 \times 10^6$ plaque-forming units/ml and an average insert size of $> 400$ bp.

### Immunoscreening and isolation of catrocollastatin-plasmid DNAs

The library was screened with the antiserum raised as described above. Screenings were performed following the protocol described by Sambrook et al. [23]. To recover cloned plasmid DNAs, an aliquot of phage stock was mixed with host cells (XL1-Blue) and 1 $\mu$l of helper phage (R-408). After incubation at 37 °C for 3 h in medium, the mixture was heated at 70 °C for 20 min and centrifuged at 4000 ***g*** for 5 min. Aliquots of the supernatant were mixed with host cells and spread on to plates containing ampicillin and, after overnight incubation, single colonies appeared containing the cloned DNA insert in pBluescript SK(−). Double-stranded DNA was prepared using the Magic Minipreps kit from Promega Corporation (Madison, WI, U.S.A.).

### Southern blotting

An aliquot of each DNA preparation was digested with restriction enzymes *Eco*RI and *Xho*I and resolved on a 0.8% agarose gel and DNA was blotted on to a nitrocellulose membrane. The blot was prehybridized and hybridized at 60 °C with the $^{32}$P-end-labelled oligonucleotide 48-mer probe. After three 15 min washings with 2 × SSC (1 × SSC is 150 mM NaCl/50 mM sodium citrate) at 50 °C, the blot was exposed to X-ray film for autoradiography.

### Nucleotide sequencing

Cloned DNAs were sequenced by the dideoxynucleotide method of Sanger et al. [24]. The sequencing reactions were performed on double-stranded DNA using the Sequenase version of DNA polymerase from United States Biochemicals. Typical reactions gave readable sequences of 200–250 nucleotides on 6% (w/v) denaturing polyacrylamide gels detected by autoradiography. With the synthesis of new sequencing primers, additional sequences were obtained, and these were assembled into overlapping contiguous sequences using the Lasergene computer program from DNAStar. The assembled DNA sequences were compared with sequences in the GenBank database using the fasta program [25] of GCG (Genetics Computer Group; University of Wisconsin, Biotechnology Center, Madison, WI, U.S.A.).

### Expression of catrocollastatin cDNA with baculovirus expression system

*Spodoptera frugiperda* (Sf9) insect cells (Invitrogen Corp., San Diego, CA, U.S.A.) were grown in suspension with serum-free medium Sf900 (Gibco BRL, Grand Island, NY, U.S.A.) in cell-

culture flasks at 28 °C with shaking (125 rev./min). The catrocollastatin cDNA was cleaved out from plasmid pBlueScript and subcloned into the transfer vector pBlueBacIII. A 3 $\mu$g portion of recombinant transfer plasmid purified by CsCl-gradient centrifugation was co-transfected into $2 \times 10^6$ Sf9 cells along with 1 $\mu$g of wild-type *Autographa californica* multiple nuclear-polyhedrosis-virus viral DNA using the liposome procedure (Invitrogen Instruction Manual). *In vivo* homologous recombination in Sf9 cells between polyhedrin sequences in the wild-type viral DNA and the recombinant plasmid resulted in the generation of recombinant baculovirus. Recombinant baculoviruses were isolated and purified by plaque assay, then validated by PCR. Sf9 cells at a density of $2 \times 10^6$ cells/ml were inoculated with recombinant virus at a multiplicity of infection of about 5. The cells were allowed to grow continuously for 5 days. Samples were obtained from the cell culture every 24 h for time-course tests. Supernatants of cell-culture medium containing the recombinant catrocollastatin were collected by centrifugation at 150 *g* for 10 min. The supernatant was used for Western blotting and anti-aggregation tests.

### Western blotting

The supernatant of transfected Sf9 cell-culture medium, as well as native catrocollastatin, was subjected to SDS/PAGE. The proteins were transferred on to an Immobilin membrane (Millipore Corp., Bedford, MA, U.S.A.), probed with polyclonal antibodies raised against native catrocollastatin, treated with goat anti-rabbit IgG conjugated with alkaline phosphatase (Sigma, St. Louis, MO, U.S.A.) and, finally, revealed with Nitroblue Tetrazolium chloride (NBT; Sigma) and 5-bromo-4-chloroindol-3-yl phosphate (BCIP; Sigma).

## RESULTS

### Isolation, purification and characterization of catrocollastatin

The purified catrocollastatin migrated as a single band with an apparent molecular mass of 50 kDa when subjected to reducing PAGE (Figure 1), suggesting that it is composed of a single polypeptide chain. Catrocollastatin inhibited platelet adhesion to collagen in a concentration-dependent manner, as shown in Figure 2(b). At 5 $\mu$g/ml, the catrocollastatin totally abolished platelet aggregation induced by 1 $\mu$g/ml collagen, but had no effect on platelet aggregation induced by ADP or the thromboxane analogue U46619 (Figure 2a).

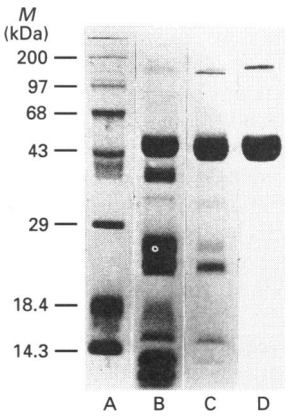



**Figure 1 Purification of catrocollastatin**

The Figure shows a reducing SDS/12%-PAGE gel after staining with Coomassie Blue. Lane A contained molecular-mass (*M*) standards. Venom was harvested from *C. atrox* glands (lane B). It was fractionated by Sephadex G-100 chromatography and analysed for its ability to inhibit platelet–collagen adhesion. The fractions with the above effect were collected (lane C) and were further purified by reverse-phase HPLC. The molecular mass of the purified anti-adhesive protein is 50 kDa and the protein was named 'catrocollastatin' (lane D).

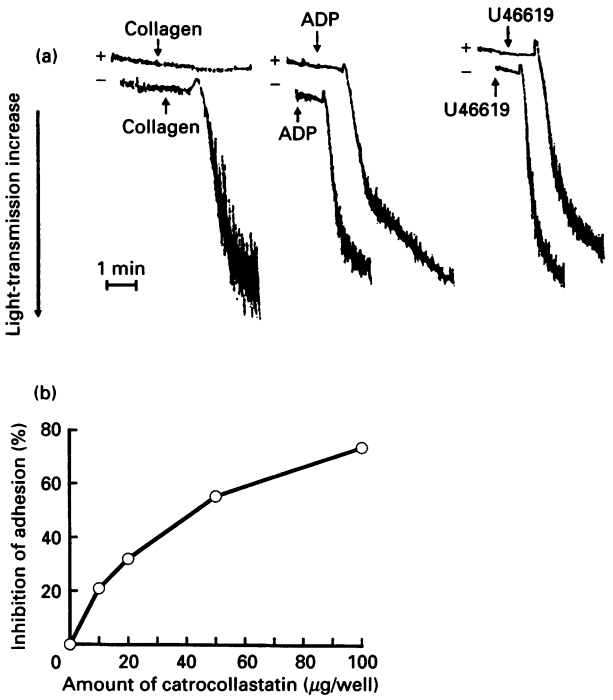
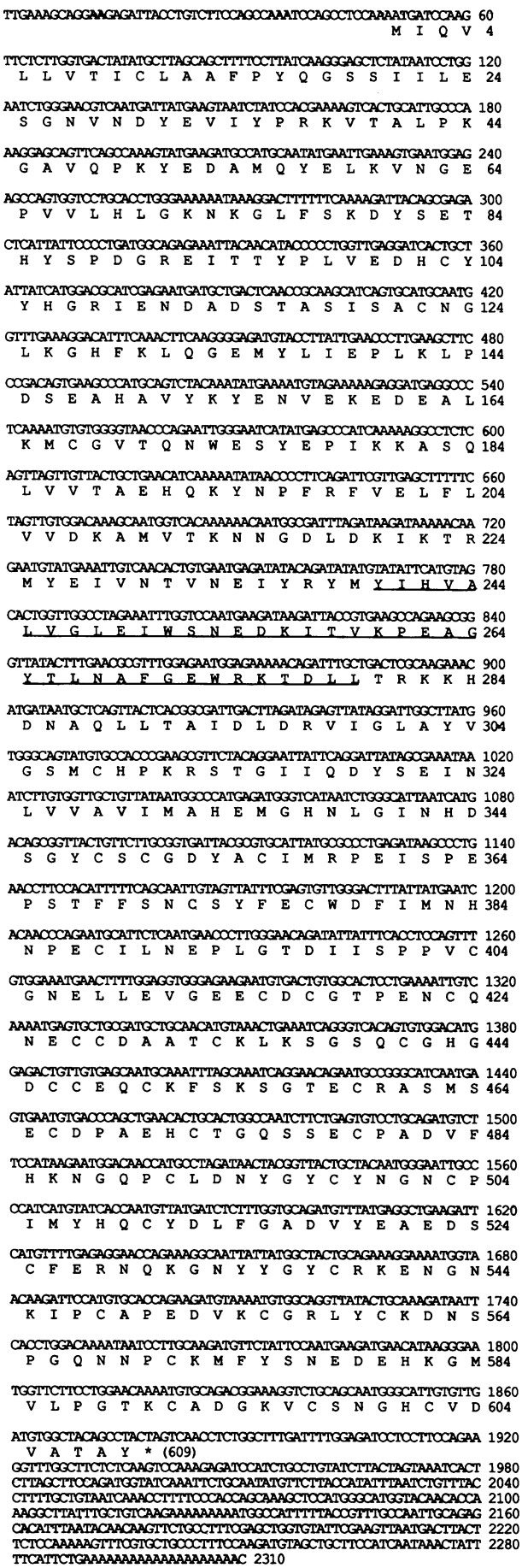



**Figure 2 Effects of catrocollastatin on platelet aggregation and adhesion**

(**a**) Effect of catrocollastatin (5 $\mu$g/ml) (indicated by the plus sign) on aggregation in human citrated platelet-rich plasma induced by collagen (1 $\mu$g/ml), ADP (5 $\mu$M) or U46619 (2 $\mu$M) versus the vehicle control (indicated by the minus sign). (**b**) Concentration–response curve for inhibition of platelet adhesion to collagen by catrocollastatin. The method is described in the Experimental section.

### Isolation, preliminary validation and sequencing of cDNA clones

A number of clones tested positive with the polyclonal antiserum probe in the primary screening. Five putative positive clones, designated Sn1–Sn5, were picked for further screening. Two more rounds of screening of these five clones resulted in purified positive clones. The authenticity of Sn2 was preliminarily validated by Southern-blot analysis. The $^{32}$P-end-labelled 48-mer oligonucleotide probe (based on a partial amino acid sequence of peptide 2) specifically hybridized with the 2.3 kb Sn2 insert (details not shown). The five clones, which varied in size from 1.6 to 2.3 kb, shared partial sequence identity. The Sn2 clone was a full-length clone and was completely sequenced. This sequence, confirmed by determining both strands, consisted of 2310 bases including a 50-nt 5′-end non-coding region, a 1830-nt open reading frame, and a 3′-end non-coding region with termination codon, AATAAA sequence and polyadenylation sites. Figure 3 shows the nucleotide sequence and the translated protein sequence. The cDNA deduced amino acids 240–279 are 100% identical with peptide 1. Sequence from amino acids 507–537, 26 out of 31 residues are identical with peptide 2. The divergence of the later peptide sequences may be due to the impurity of the peptide sample or amino-acid-sequencing errors.



## Western blot and activity test of baculovirus/Sf9 cell expressed recombinant catrocollastatin

Samples of supernatant of cell culture media were examined by SDS/PAGE followed by Western blotting. It showed a specific band of 50 kDa which co-migrated with the native catrocollastatin that appeared in significant amounts on the third day after infection. The intensity of this band increased in the following 2 days (Figure 4a). The inhibitory effect of the supernatant towards collagen-induced platelet aggregation appeared to correlate with the increasing amount of the 50 kDa band (Figure 4b). It appeared on the third day, became significant on the fourth day and completely inhibited collagen-induced aggregation on the fifth day. Breakdown products of expressed protein could be seen beginning on day 3 and accumulated on subsequent days. As the recombinant catrocollastatin was not isolated for the activity tests, determining the function of the breakdown products will require further investigation. Sf9-cell-expressed recombinant catrocollastatin showed the same inhibitory activity and specificity as the native protein (results not shown).

## DISCUSSION

We have cloned a full-length 2310-bp-long cDNA for catrocollastatin from a library constructed from *C. atrox* venom gland polyadenylated RNA (Figure 3). The deduced amino acid sequence consists of 609 amino acids with a calculated molecular mass of 67 kDa, larger than the 50 kDa determined for catrocollastatin by SDS/PAGE. A search of the GenBank database found similar nucleotide sequences in snake-venom proteins jararhagin from *B. jararaca* [19], Ht-a and Ht-e from *C. atrox* [20] and trigramin from *T. gramineus* [26]. They all share the consensus zinc-binding site H-E-X-X-H, which is seen in matrix metalloproteinases [14], and a disintegrin-like domain. This implies the presence of a signal peptide, an N-terminal domain, a metalloproteinase domain, a disintegrin-like domain and a C-terminal domain in the catrocollastatin sequence. This multidomain structure is in agreement with the common precursor model proposed by Kini and Evens [27]. Assuming the propeptide is cleaved by post-translational modification, the molecular mass of the processed protein would be 48.4 kDa, close to the apparent molecular mass of 50 kDa seen on SDS/PAGE.

The platelet-adhesion assay indicated that purified catrocollastatin inhibited platelet–collagen adhesion in a concentration-dependent manner (Figure 2b). We also demonstrated that catrocollastatin abolished platelet aggregation induced by collagen, but not by ADP or U46619 (Figure 2a). These effects are similar to those of the other 50 kDa snake-venom proteins we reported previously [21]. Recombinant protein expressed by Sf9 cells has the same anti-aggregation activity and specificity as the native protein. These results show correlation of the sequence obtained from cDNA cloning with biological activity.

The nucleotide sequence similarity in the metalloproteinase domain is 77.2% between catrocollastatin and jararhagin, 56.7% between catrocollastatin and Ht-a, and 60% between catrocollastatin and Ht-e. The zinc-binding site consensus sequence

**Figure 3 cDNA and deduced amino acid sequence of Sn2**

The nucleotide sequence is shown in the upper line and the predicted amino acid sequence is shown in the lower line of each row. The cDNA-deduced amino acids 240–279 are 100% identical with the partial sequence of the CNBr digest peptide of native catrocollastatin (underlined).

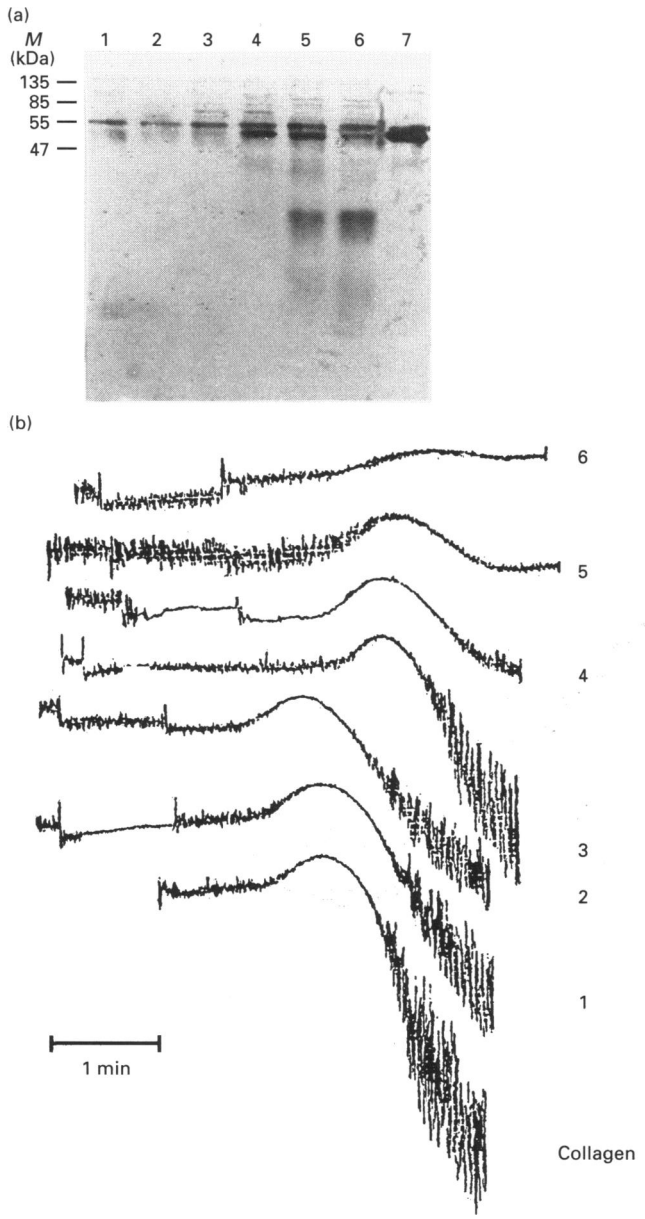



**Figure 4 Expression of catrocollastatin in Sf9 insect cells with baculovirus system: time courses**

(**a**) Western blot: SDS/PAGE of the supernatant of transfected Sf9-cell culture and native catrocollastatin were transferred on to membrane, probed with polyclonal antiserum probe, treated with goat anti-rabbit IgG conjugated with alkaline phosphatase, and proteins revealed with NBT and BCIP. Lane 1, supernatant of non-infected Sf9 cell culture. Lanes 2–6 show the supernatant of recombinant baculovirus-infected Sf9-cell cultures at days 1–5 respectively. Lane 7 is HPLC-purified native catrocollastatin. Abbreviation: *M*, molecular mass. (**b**) Anti-aggregation effect of expressed catrocollastatin (60 $\mu$l of the supernatant of Sf9-cell culture in 0.5 ml of platelet-rich plasma). The numbers correspond to those of the Western blot.

(H-E-X-X-H)-X-X-G-X-X-H [14] is conserved in all of these proteins (Figure 5). Recently, Hite and his co-workers proposed that the snake-venom metalloproteinases define a new family of zinc metalloproteinases [28]. Microscopic studies indicate that the high-molecular-mass haemorrhagic metalloproteinase, Ht-a from *C. atrox* and small molecules like Ht-e degrade the proteins of the basement membrane of blood vessels [29]. Kamiguti et al. [30] observed that systemic haemorrhage induced by subcutaneous injection of *B. jararaca* venom resulted from disruption of the collagenous basement membrane of the vascular endothelium. The substrate specificity of these venom toxins demonstrated that they degrade type IV collagen [31,32], but do not degrade collagen types I, II, and III [33]. These are similar to the substrate specificities of jararafibrases I and II from *B. jararaca* venom, which degrade type IV collagen, gelatin, laminin and fibronectin into smaller fragments, but degrade type I collagen only partially in a non-specific manner [34]. These findings suggest that snake-venom haemorrhagins may disrupt the pericellular basement membrane through proteolytic activity.

The sequences of disintegrin-like domains of catrocollastatin and jararhagin are identical, and the C-terminal domains show 95.9% similarity. Catrocollastatin and Ht-a have 67.4% similarity in the disintegrin domain and 59.5% similarity in the C-terminal domain. Neither of these disintegrin-like domains contains the RGD sequence (Figure 5). The structural difference between high- and low-molecular-mass haemorrhagins is that the high-molecular-mass ones have a disintegrin-like domain following the metalloproteinase domain. A comparison between Ht-e cDNA and that of the mature Ht-e protein indicates that the proprotein of this metalloproteinase contains a non-RGD containing disintegrin-like C-terminal region [20]. High-molecular-mass metalloproteinases HR1A, HR1B and Ht-a will autodigest when incubated at 37 °C with conformational perturbants such as SDS or in the absence of $Ca^{2+}$. Thus high-molecular-mass metalloproteinases are thought to be the precursors of low-molecular-mass haemorrhagins and disintegrin-like proteins [28,35]. However, we did not observe autodigestion of catrocollastatin under the same conditions (Q. Zhou and J. B. Smith, unpublished work).

It has been known for some time that the high-molecular-mass Ht-a has more than 10-fold higher a haemorrhagic activity than the smaller molecules Ht-b to Ht-e inclusive [17]. It was suggested that the disintegrin-like domain of the high-molecular-mass metalloproteinase may be important in synergistically stimulating haemorrhagic activity in combination with the N-terminal metalloproteinase domain [36]. The most recent site-directed-mutagenesis studies on kistrin, one of the simple disintegrins from venom protein, showed that the aspartate residue in RGD is critical for its high-affinity binding to GP IIb/IIIa. Mutations outside of the RGD region led to proteins indistinguishable from native kistrin [37]. Almost all the disintegrins lose their activity after being reduced. Direct chemical analysis of the disulphide-bridge pattern of disintegrins has recently been accomplished [38,39]. It has been found that two highly conserved disulphide bridges form a motif with respect to the RGD sequence. This observation is in agreement with the results of NMR studies [40, 41]. It has therefore been suggested that a favourable conformation of the RGD region alone is responsible for the high-affinity binding of disintegrin to GP IIb/IIIa. By comparing the similarity of disintegrins and jararhagin, Paine et al. [19] suggested that SECD (Ser-Glu-Cys-Asp) in jararhagin, which corresponds to RGD in disintegrins, is a crucial sequence (Figure 5). Both the disintegrin and disintegrin-like domains are rich in cysteine, and the cysteine pattern is highly conserved. The disulphide bonds in these large molecules could potentially modify the configuration of SECD or other corresponding sequences as they do in the RGD loop (Figure 5).

Our previous findings suggested that, unlike disintegrins, catrocollastatin and other 50 kDa inhibitory snake-venom proteins do not function by interaction with the fibrinogen receptor GP IIb/IIIa, as their inhibitory effect was demonstrated on platelet–collagen adhesion even in the presence of RGDS, an inhibitor of this receptor [21]. The cDNA-deduced amino acid

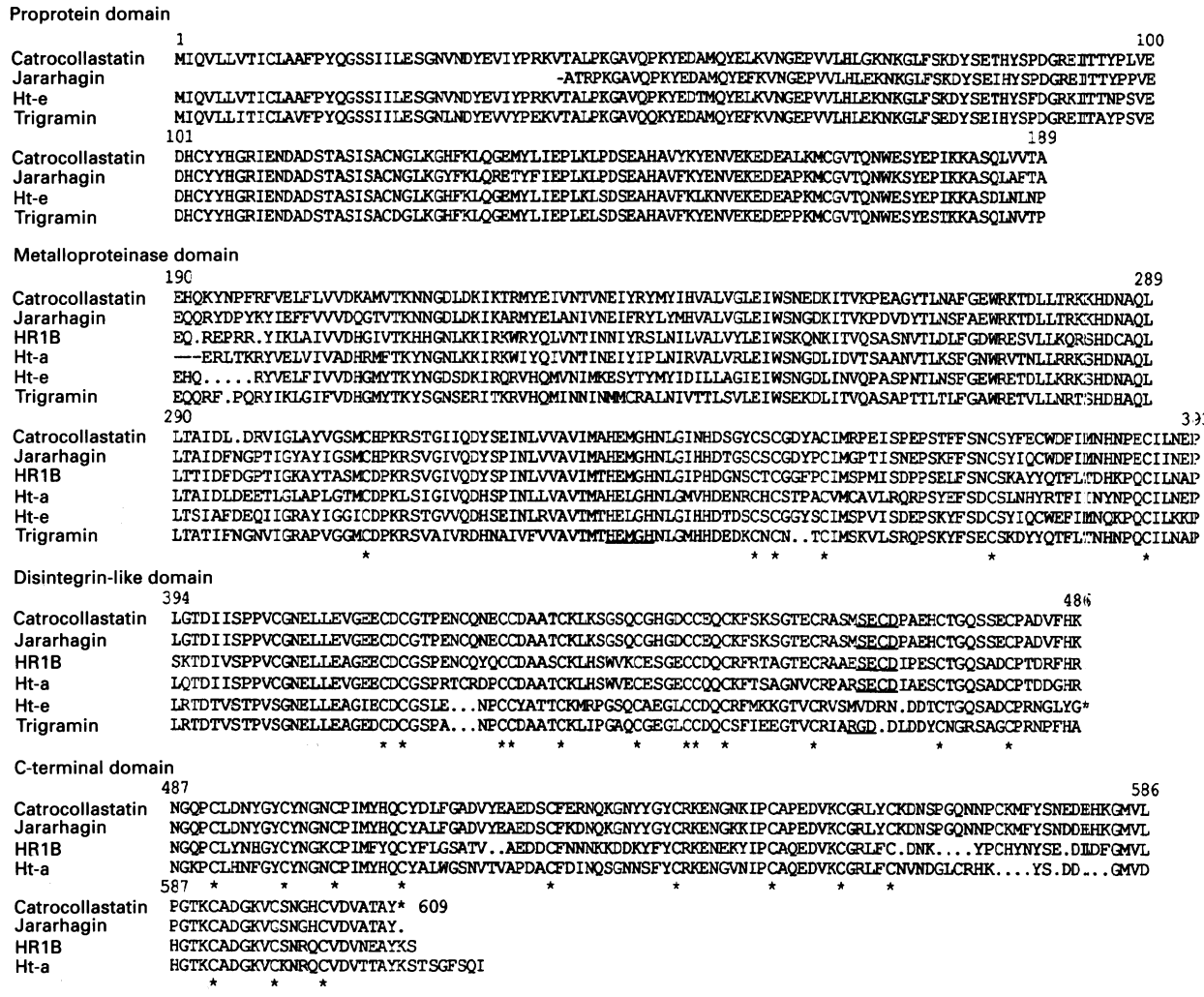



**Figure 5 Comparison of the cDNA-deduced amino acid sequence of catrocollastatin with members of the snake-venom metalloproteinases and disintegrins**

Comparisons are made by dividing the whole molecules into four domains: proprotein domain (1–189), metalloproteinase domain (190–393), disintegrin-like domain (394–486) and C-terminal domain (487–609). The zinc motif (HEXXH) is underlined. SECD sequences in catrocollastatin, jararhagin, HR1B and Ht-a, and RGD in trigramin, are underlined. Conserved cysteine residues are indicated with an asterisk.

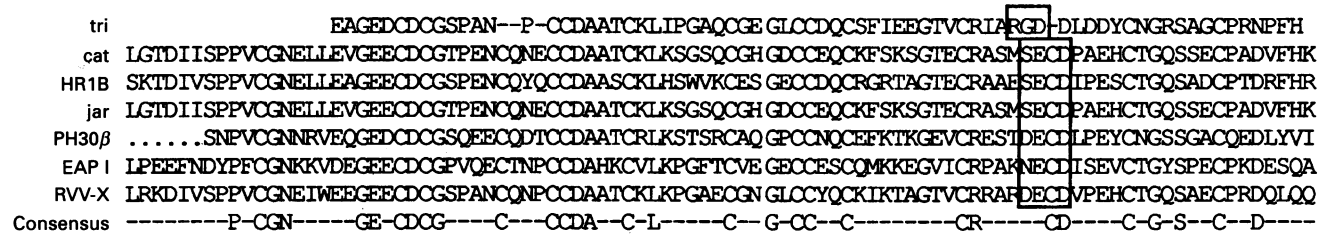



**Figure 6 Similarities between disintegrins**

Disintegrin-like domains of catrocollastatin (cat), HR1B, jararhagin (jar), coagulation-Factor-X-activating enzyme from Russell's-viper venom (RVV-X), monkey epididymal apical protein (EAP I) and guinea-pig sperm surface protein PH-30$\beta$ are compared. The RGD in trigramin (tri), SECD in catrocollastatin and the corresponding sequences in the rest of the proteins are outlined. The conserved sequences are shown on the bottom line.

sequence of the catrocollastatin shows that, as in HR1B and jararhagin, the disintegrin-like domain RGD sequence is replaced by SECD. Putative collagen receptors on the surface of platelets are not RGD-recognizing receptors [42]. Early studies by Staatz et al. [43] suggested that a tetrapeptide sequence, DGEA (Asp-Gly-Glu-Ala), in type I collagen was a collagen receptor GP Ia/IIa ($\alpha_2\beta_1$) ligand. A search of the Swiss-Prot database revealed that SECD has no similarity in any type of collagen. Coagulation

Factor X-activating enzyme from Russell's-viper (*Vipera russelli*) venom has a molecular structure and sequence similar to those of HR1B [44], jararhagin and catrocollastatin (Figure 6). The activating enzyme has DECD (Asp-Glu-Cys-Asp) in the positions corresponding to SECD in catrocollastatin, jararhagin and HR1B (Figure 6). A recently cloned ecacrin, a prothrombin activator in the venom of *Echis carinatus*, was also found to contain a metalloproteinase and disintegrin-like domain structure. However, it has RDD (Arg-Asp-Asp) in the place of RGD [45]. Surprisingly, this molecular pattern is not restricted to snake-venom metalloproteinases. A newly cloned mammalian epididymal apical protein (EAP I) shows an intriguing similarity to a variety of snake-venom haemorrhagic proteins, including HR1B [46]. In this protein, NECD (Asn-Glu-Cys-Asp) is substituted for RGD in disintegrin and SECD in catrocollastatin. PH-30$\beta$, a guinea-pig sperm surface protein involved in sperm–egg fusion distinct from the apical protein [47] also has a disintegrin-like domain with DECD (Asp-Glu-Cys-Asp) instead of RGD (Figure 6). Perry et al. [46] inferred that these crucial sequences might introduce the specific binding of sperm to integrin-like receptors on the egg. Recently, it has been observed that catrocollastatin binds collagen (Q. Zhou and J. B. Smith, unpublished work). We hypothesize that high-molecular-mass metalloproteinases of haemorrhagic snake venoms are targeted to a specific site on collagen by the sequence aligned with RGD in the disintegrin-like domain. By this means the disintegrin-like domain could not only restrict the specificity of the metalloproteinase, but also play a synergistic role by attaching the enzyme to the substrate, and thereby increase the local concentration of the enzyme. Further studies into the molecular structure and function of this protein using site-directed mutagenesis are in progress. These studies may lead to the development of a new class of anti-thrombotic drugs.

We are grateful to Dr. Lee-Yuan Liu-Chen for her helpful criticism, to Dr. Mary Selak for the work done at the initial stage of the study, and for the technical contributions of Carol Dangelmaier, Shaohu Sheng, and James B. Littrell. The study was supported in part by the American Heart Association and the Monsanto Co.

## REFERENCES

1 Ohsaka, A. (1979) in Handbook of Experimental Pharmacology, vol. 52 (Chen-Yuan Lee, ed.), p. 480, Springer-Verlag, New York
2 Meier, J. and Stocker, K. (1991) Crit. Rev. Toxicol. **21**, 171–182
3 Gould, R. J., Polokoff, M. A., Friedman, P. A., Huang, T.-F., Holt, J. C., Cook, J. J. and Niewiarowski, S. (1990) Proc. Soc. Exp. Biol. Med. **195**, 168–171
4 Dennis, M. S., Henzel, W., Pitti, R. M., Lipari, M. T., Napier, M.A, Deisher, T. A., Bunting, S. and Lazarus, R. A. (1989) Proc. Natl. Acad. Sci. U.S.A. **87**, 2471–2475
5 Gan, Z.-R, Gould, R. J., Jacobs, J. W., Friedman, P. A. and Polokoff, M. A. (1988) J. Biol. Chem. **263**, 19827–19832
6 Huang, T.-F., Holt, J. C., Kirby, E. P. and Niewiarowski, S. (1989) Biochemistry **28**, 661–666
7 Huang, T.-F., Liu, C.-Z., Ouyang, C. and Teng, C.-M. (1991) Biochem. Pharmacol. **42**, 1209–1219
8 Tsai, I.-H, Wang, Y.-M. and Lee, Y.-H. (1993) Thromb. Haemostasis **69**, 671 (abstr.)
9 Savage, B., Marzec, U. M., Chao, B. H., Harker, L. A., Maraganore, J. M. and Ruggeri, Z. M. (1990) J. Biol. Chem. **265**, 11766–11772
10 Nichols, A. J., Vasko, J., Koster, P., Smith, J., Barone, F., Nelson, A., Stadel, J., Powers, D., Rhodes, G., Miller-Stein, C., Bopana, V., Bennet, D., Romoff, T., Calvo, R., Ali, F., Sorenson, E. and Samanen, J. (1990) Eur. J. Pharmacol. **183**, 2019
11 Nichols, A. J., Ruffolo, R. R., Jr., Huffman, W. F., Poste, G. and Samanen, J. (1992) Trends Pharmacol. Sci. **13**, 413–417
12 Scarborough, R. M., Rose, J. W., Hus, M. A., Philips, D. R., Fried, V. A., Campbell, A. M., Nannizzi, L. and Charo, I. F. (1991) J. Biol. Chem. **266**, 9359–9362
13 Scarborough, R. M., Baughton, M. A., Teng, W., Rose, J. W., Philips, D. R., Nannizzi, L., Arfster, A., Campbell, A. M. and Charo, I. F. (1993) J. Biol. Chem. **268**, 1066–1073
14 Murphy, G. J. P., Murphy, F. and Reynolds, J. J. (1991) FEBS Lett. **289**, 4–7
15 Takeya, H., Onikura, A., Nikai, T., Sugihara, H. and Iwanaga, S. (1990) J. Biochem. (Tokyo) **108**, 711–719
16 Takeya, H., Arakawa, M., Miyata, T., Iwanaga, S. and Omori-Satoh, T. (1989) J. Biochem. (Tokyo) **106**, 151–157
17 Bjarnason, J.B. and Tu, A. T. (1978) Biochemistry **17**, 3395–3404
18 Takeya, H., Oda, K., Miyata, T., Omori-Satoh, T. and Iwanaga, S. (1990) J. Biol. Chem. **265**, 16068–16073
19 Paine, M. J. I., Desmond, H. P., Theakston, R. D. G. and Crampton, J. M. (1992) J. Biol. Chem. **267**, 22869–22876
20 Hite, L. A., Jia, L.-G, Bjarnason, J. B. and Fox, J. W. (1994) Arch. Biochem. Biophys. **308**, 182–191
21 Smith, J. B., Dangelmaier, C. and Selak, M. (1991) FEBS Lett. **283**, 307–310
22 Pirkle, H. and Marsh, N. (1991) Thromb. Haemostasis **66**, 264
23 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
24 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
25 Pearson, W. R. and Lipman, D. J. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 2444–2448
26 Neeper, M. P. and Jacobson, M. A. (1990) Nucleic Acids Res. **18**, 4255
27 Kini, R. M. and Evans, H. J. (1992) Toxicon **30**, 265–293
28 Hite, L. A., Fox, J. W. and Bjarnason, J. B. (1992) Biol. Chem. Hoppe-Seyler **373**, 381–385
29 Ownby, C. L., Bjarnason, J. B. and Tu, A. T. (1978) Am. J. Pathol. **93**, 201–218
30 Kamiguti, A. S., Theakstojn, R. D. G., Desmond, H. and Hutton, R. A. (1991) Toxicon **29**, 1097–1105
31 Baramova, E. N., Shannon, J. D., Fox, J. W. and Bjarnason, J. B. (1991) Biomed. Biochim. Acta **50**, 763–768
32 Shannon, J. D., Baramova, E. N., Bjarnason, J. B. and Fox, J. W. (1989) J. Biol. Chem. **264**, 11575–11583
33 Baramova, E. N., Shannon, J. D., Bjarnason, J. B. and Fox, J. W. (1989) Arch. Biochem. Biophys. **275**, 63–71
34 Maruyama, M., Sugiki, M., Yoshida, E., Shimaya, K. and Mihara, H. (1992) Toxicon **30**, 1387–1397
35 Takeya, H., Kubo, J., Omori-Satoh, T. and Iwanaga, S. (1991) Thromb. Haemostasis **65**, 842 (abstr.)
36 Takeya, H., Oda, K., Onikura, A., Miyata, T. and Iwanaga, S., Nikai, T., Sugihara, H. and Omori-Satoh, T. (1990) Jpn. J. Med. Sci. Biol. **43**, 252–253
37 Dennis, M. S., Carter, P. and Lazarus, R. A. (1993) Proteins **15**, 312–321
38 Calvete, J. J., Schafer, W., Soszka, T., Lu, W., Cook, J. J., Jameson, B. A. and Niewiarowski, S. (1991) Biochemistry **30**, 5225–5229
39 Calvete, J. J., Wang, Y., Mann, K., Schafer, W., Soszka, T., Niewiarowski, S. and Stewart, G. J. (1992) FEBS Lett. **309**, 316–320
40 Saudek, V., Atkinson, R. A. and Pelton, J. T. (1991) Biochemistry **30**, 7369–7372
41 Adler, M., Lazarus, R. A., Dennis, M. S. and Wagner, G. (1991) Science **253**, 445–448
42 D'sonza, S. E., Ginsberg, M. H. and Plow, E. F. (1991) Trends Biochem. Sci. **16**, 246–250
43 Staatz, W. D., Fok, K. F., Zutter, M. M., Adams, S. P., Rodriguez, B. A. and Santoro, S. A. (1991) J. Biol. Chem. **266**, 7363–7367
44 Takeya, H., Nishida, S., Miyata, T., Kawada, S, Saisaka, Y, Morita,T. and Iwanaga, S. (1992) J. Biol. Chem. **267**, 14109–14117
45 Nishida, S., Fujita, T., Kono, N., Atoda, H., Morita, T., Takeya, H., Kido, I., Paine, M. J. I., Kawabata, S. and Iwanaga, S. (1993) Thromb. Haemostasis **69**, 1047
46 Perry, A. C. F., Jones, R., Barker, P. J. and Hall, L. (1992) Biochem. J. **286**, 671–675
47 Blobel, C. P., Wolfsberg, T. G., Turck, C. W., Myles, D. G., Primakoff, P. and White, J. M. (1992) Nature (London) **356**, 248–252

Received 1 September 1994/22 November 1994; accepted 29 November 1994