Amino acid sequence and molecular modelling of glycoprotein IIb-IIIa and fibronectin receptor iso-antagonists from *Trimeresurus elegans* venom

Andrea SCALONI*, Elisabetta DI MARTINO†, Nadia MIRAGLIA*, Alessandra PELAGALLI†, Rossella DELLA MORTE†, Norma STAIANO† and Piero PUCCI‡§||

*Servizio di Spettrometria di Massa-I.A.B.B.A.M., Consiglio Nazionale delle Ricerche, via Pansini 5, 80131 Naples, Italy, †Dipartimento di Biochimica e Biotecnologie Mediche, and ‡Dipartimento di Chimica Organica e Biologica, Università di Napoli 'Federico II', Via Mezzocannone 16, 80134 Naples, Italy, and §CEINGE Biotecnologie Avanzate, via Pansini 5, 80131 Naples, Italy

Low-molecular-mass Arg-Gly-Asp (RGD)-containing polypeptides were isolated from the venom of *Trimeresurus elegans* by a simple two-step procedure consisting of membrane filtration and reverse-phase HPLC. A combination of electrospray MS, fast-atom bombardment MS and Edman degradation allowed us to ascertain the presence in the venom of different isoforms and to determine their primary structures. The amino acid sequences resembled the structure of elegantin, the only disintegrin previously reported from the *T. elegans* venom [Williams, Rucinski, Holt and Niewiarowski (1990) *Biochim. Biophys. Acta* **1039**, 81–89]. MS analyses indicated the occurrence of differential proteolytic processing at both the N-terminus and the C-terminus of the polypeptide chains. The amino acid sequence alignment of the elegantin isoforms with known components of the disintegrin

family demonstrated the complete conservation of the 12 cysteine residues involved in disulphide bridges. Molecular modelling of elegantins predicted an overall folding of these molecules quite similar to that reported for the kistrin solution structure. The newly identified polypeptide isoforms strongly inhibited ADP-induced aggregation in both human and canine platelet-rich plasma but showed a different species-dependent specificity. These molecules were also able to inhibit B16-BL6 murine melanoma cell adhesion to immobilized fibronectin. The comparison of the structures and biological activities of elegantin isoforms and kistrin allowed us to highlight some structural features that, in addition to the RGD locus, might be involved in the interaction of these snake-venom polypeptides with the integrin receptors on the platelet and cell surface.

INTRODUCTION

Cell–cell and cell–matrix interactions are mediated by cytoadhesive receptors expressed on the cell surface. These receptors, referred to as integrins, comprise a superfamily of transmembrane heterodimeric molecules, which includes the platelet membrane glycoprotein (GPIIb-IIIa), the vitronectin receptor on endothelial cells and fibronectin receptors on fibroblasts [1]. A number of adhesive proteins, such as fibronectin, vitronectin, osteopontin, collagens, laminin, fibrinogen, von Willebrand factor and complement component C3bi, contain the tripeptide sequence Arg-Gly-Asp (RGD), which represents a common cell-attachment recognition site [2].

A novel family of low-molecular-mass RGD-containing, cysteine-rich proteins, named disintegrins, have been isolated from the venom of various snakes in the Crotalidae and Viperidae [3–11] and from leeches [12–14]. These molecules were demonstrated to be potent inhibitors of platelet aggregation. Several experimental results indicate that disintegrins bind to the platelet GPIIb-IIIa, and thus they act by blocking the binding of adhesive proteins such as fibrinogen and von Willebrand factor [4,5,7,14,15]. However, other mechanisms, including fibrin(ogen)olysis [13,16] or thrombin [17] and factor Xa [18] inhibition, might be involved in mediating their effects on haemostasis. Disintegrins were also shown to inhibit the adhesion of human and murine melanoma cells to extracellular matrix components [19–23], suggesting their ability to bind other integrins on cell surface in addition to GPIIb-IIIa, which is found exclusively on platelets.

Disintegrins exert their biological activity by a common mechanism of action involving the RGD sequence, which is

present in all these polypeptides with the exception of barbourin [24], a KGD (Lys-Gly-Asp) peptide that selectively inhibits fibrinogen binding to platelet GPIIb-IIIa. However, it has been proposed that the specificity and/or affinity of a disintegrin for individual cell-surface receptors are determined by the spatial configuration of the RGD motif as well as by either the RGD flanking residues or other recognition sequences in the protein molecule [25,26].

This paper reports a simple purification procedure and the complete structural characterization of nine different RGD-containing peptide isoforms present in *Trimeresurus elegans* venom by integrating classical biochemical methodologies with advanced MS procedures. Results are also presented on the ability of these proteins to inhibit ADP-induced aggregation of human or canine platelet-rich plasma (PRP) and the attachment of B16-BL6 murine melanoma cells to immobilized fibronectin. The molecular model of the elegantin isoforms, constructed with the three-dimensional structure of kistrin as a template [27], allowed us to make some hypotheses on additional structural features other than the RGD motif, which might play a role in the interaction of these polypeptides with GPIIb-IIIa and fibronectin receptors on platelet and mouse melanoma cells respectively.

EXPERIMENTAL

Chemicals

Freeze-dried venom from *T. elegans*, dithiothreitol, EDTA, Tris/HCl, tosylphenylalanylchloromethyl trypsin (EC 3.4.21.4),

Abbreviations used: E1, elegantin 1; E2, elegantin 2; ES/MS, electrospray MS; FAB/MS, fast-atom bombardment MS; GPIIb-IIIa, glycoprotein IIb-IIIa; MEM, Eagle's minimal essential medium; PRP, platelet-rich plasma; TFA, trifluoroacetic acid.

|| To whom correspondence should be addressed at Servizio di Spettrometria di Massa, via Pansini 5, 80131 Naples, Italy.

human plasma fibronectin, BSA, non-essential amino acids and sodium pyruvate were purchased from Sigma (St. Louis, MO, U.S.A.). Endoproteinase Asp-N (EC 3.4.24.33) was obtained from Boehringer (Mannheim, Germany), HPLC solvents from Carlo-Erba (Milan, Italy), fetal calf serum from Hyclone Laboratories Inc. (Logan, UT, U.S.A.), and Eagle's minimal essential medium (MEM), vitamin solution and glutamine from Flow Laboratories (Irvine, Scotland, U.K.). The adhesive protein fibronectin appeared more than 95% homogeneous by SDS/PAGE analysis and Coomassie Blue staining.

Purification of elegantins from *T. elegans* venom

Elegantins were purified from crude venom from *T. elegans*. Freeze-dried venom (100 mg) was dissolved in 1 ml of 0.1% trifluoroacetic acid (TFA), stored on ice for 10 min and centrifuged at 17000 *g* for 5 min at 4 °C to remove any insoluble material. The supernatant was subjected to ultrafiltration through a Centricon 10 device (Amicon Inc., Beverly, MA, U.S.A.) in a refrigerated centrifuge. The filtrate was then centrifuged at 3000 *g* for 10 min at room temperature on a Centricon 3 device and the non-diffusible material was directly separated on a semi-preparative Vydac C18 column 218TP510 (250 mm × 10 mm; 5 µm, 300 pore size; Separations Group, Hesperia, CA, U.S.A.) eluted with a linear gradient of acetonitrile from 0% to 30% in 0.1% TFA over 60 min at a flow rate of 2.4 ml/min. Fractions were collected manually and tested for the inhibition of platelet aggregation. The active fractions were finally repurified by a further analytical HPLC step on a Vydac C18 column 218TP54 (250 mm × 4.6 mm; 5 µm, 300 pore size; The Separations Group). Proteins were dissolved in 0.1% TFA, loaded on to the column equilibrated with 0% or 10% acetonitrile in 0.1% TFA and eluted by means of a linear 0–30% or 10–40% gradient of acetonitrile in 0.1% TFA over 60 min at a flow rate of 1 ml/min.

Protein concentrations were determined either by amino acid analysis or by a microscale version of the Lowry procedure [28]. Peptide hydrolysates obtained by 6 M HCl hydrolysis in the vapour phase, at 110 °C for 48 h, were analysed with an Eppendorf-Biotronik LC300 analyser.

Amino acid sequence determination

Before sequence analyses, elegantins were reduced and carboxymethylated as previously reported [29,30]. Protein samples were freed from the excess of reagents by an HPLC desalting step on a C18 Vydac column as described above. Carboxymethylated elegantins were digested with tosylphenylalanylchloromethane-treated trypsin or with endoproteinase Asp-N in 0.4% ammonium bicarbonate, pH 8.5, at 37 °C for 6 and 18 h respectively, using an enzyme-to-substrate ratio of 1:50 (w/w). The resulting peptide mixtures were fractionated by analytical reverse-phase HPLC as described above. The N-terminal sequences of either alkylated samples of elegantins or HPLC-purified peptides were determined by using an Applied Biosystem 477A pulsed-liquid protein sequencer equipped with an Applied Biosystem 120A HPLC apparatus for phenylthiohydantoin-amino acid identification.

Electrospray (ES/MS) and fast-atom bombardment (FAB/MS) mass spectrometric analyses

HPLC-purified elegantins and individual peptides obtained after enzymic digestion were directly analysed by ES/MS with a VG Bio-Q triple quadrupole mass spectrometer (VG Micromass, Manchester, U.K.). Aliquots (10 µl) of protein and peptide

solutions were injected into the ion source at a flow rate of 10 µl/min; the mass spectrometer was scanned from *m/z* 600 to *m/z* 1800 at 10 s per scan. Mass calibration was performed by means of multiply charged ions from a separate injection of myoglobin (average molecular mass 16951.5 Da); all molecular masses are reported as average masses.

Peptide mixtures obtained from enzymic hydrolyses of both native and carboxymethylated elegantins were analysed by FAB/MS on a VG-ZAB 2SE double-focusing mass spectrometer, equipped with a caesium gun operating at 25 kV (2 µA) (VG Micromass). MS analyses were performed by loading 2 µl of sample solution onto a glycerol/thioglycerol-coated probe tip; spectra were recorded on ultraviolet-sensitive paper and counted manually. Assignments of the recorded mass values to individual peptides were performed on the basis of their molecular mass and confirmed by submitting the entire peptide mixture to a single step of manual Edman degradation and re-running the FAB spectrum of the truncated peptide mixtures as described previously [29,30].

Sequence alignment and model construction

The degree of overall sequence similarity of the elegantins was determined by calculating the average percentage of similar residues among all pairwise comparisons of the 20 aligned sequences of disintegrins and other RGD-containing peptides. In assessing the similarity of two residues, the following conservative substitutions were allowed: I, L, V, M; D, E; R, K, H; F, W, Y; T, S; whereas Q, N, A, G, P and C were considered to be unique.

A three-dimensional model of elegantin isoforms was constructed on the basis of the solution structure of kistrin [27] (PDB code 1KST) taken from the Protein Data Bank, Brookhaven National Laboratory (Upton, NY, U.S.A.) [31,32], and by using the sequence alignment previously described. The elegantin models were constructed with the Insight/Discover program version 2.5 (Biosym Technologies, San Diego, CA, U.S.A.) running on a Silicon Graphics Indigo2 workstation. Kistrin residues different from those of elegantins were replaced with the appropriate amino acid residues by using standard geometrical parameters provided in the Insight II software package while retaining the appropriate backbone conformations; residue numbering is based on the numbering system for elegantins. The N-terminal dipeptide and C-terminal tripeptide were not included in the model construction, because these residues are absent from the kistrin sequence.

Platelet aggregation

Blood was obtained from healthy donors who had not taken medications within the previous 2 weeks. Blood samples were collected in sodium citrate, pH 7.35 (0.01 M final concentration), and then centrifuged at 160 *g* for 10 min at room temperature to obtain PRP. The platelet count was adjusted to 250000/mm³. Platelet-poor plasma was prepared from the PRP by further centrifugation at 2000 *g* for 10 min. Test substances were added to PRP and allowed to incubate for 1 min, followed by the addition of ADP (10 µM final concentration) to initiate aggregation. Platelet aggregation was recorded in a Chronolog aggregometer. The change in light transmission was measured to the point where the tracing reached a plateau. The extent of inhibition of platelet aggregation was expressed as the percentage of the rate of aggregation observed in the absence of inhibitor. The IC₅₀ values, defined as the concentration of disintegrin required to inhibit 50% of the aggregation of PRP caused by 10 µM ADP, were determined from dose-dependence curves.

Canine PRP was prepared by the same procedure, by collecting the blood from dog carotid artery.

Cell culture and adhesion assay

B16-BL6 mouse melanoma cells were grown in MEM supplemented with 10% fetal calf serum, 1% non-essential amino acid solution, 2% vitamin solution, 1% sodium pyruvate solution and 1% glutamine. Cells were harvested for propagation or cell attachment study by treatment with 0.25% trypsin/0.02% EDTA in PBS, pH 7.2, washed with MEM and resuspended in complete MEM or in serum-free MEM respectively. Cell lines were maintained in culture for no longer than 2 weeks to ensure that their metastatic phenotype did not change as a result of prolonged passages *in vitro*.

Plastic tissue-culture dishes (Costar, Cambridge, MA, U.S.A.) with 96 wells were coated overnight at 4 °C by incubation with 0.1 ml of human plasma fibronectin (10 µg/ml) and were then treated with 1% BSA in water for 20 min, at 37 °C, for blocking free binding sites on the plastic. Wells coated with 1% BSA alone served as negative controls.

Freshly resuspended cells (10^5 cells per well) were plated on to coated wells and allowed to attach for 1 h at 37 °C. At the end of the attachment period, non-adherent cells were removed by gentle washing with PBS; adherent cells were fixed and stained for 3 min with a 0.5% Crystal Violet solution in 20% (v/v) methanol. After being washed extensively with distilled water, the plates were left to dry overnight at room temperature; colour yields were then measured with a Bio-Rad ELISA reader with a 570 nm filter. In this assay the colour intensity is proportional to the number of adherent cells. Approx. 75% of B16-BL6 cells initially added to the wells adhered to the fibronectin-coated dishes under our experimental conditions. To assess the inhibition of cell attachment by disintegrins, equal volumes of cell suspension were mixed with different dilutions of peptide and added immediately to the wells.

RESULTS

Purification and determination of primary structure

Elegantins were purified to homogeneity from the soluble portion of the freeze-dried venom of *T. elegans* with a simple two-step procedure consisting of membrane ultrafiltration and reverse-phase HPLC. Their inhibitory activity towards platelet aggregation was not assayed in the crude material owing to the presence of protein components able to stimulate platelet aggregation. After elimination of these species by 10 kDa cut-off membrane filtration, the biological activity of elegantins was tested by measuring the inhibition of ADP-induced platelet aggregation *in vitro*. Centrifugation through a 3 kDa cut-off membrane of the filtrate was performed, and the non-diffusible material was directly subjected to semipreparative reverse-phase HPLC (Figure 1, top panel). Two fractions accounting for most of the inhibitory activity, designated elegantin 1 (E1) and elegantin 2 (E2), were detected and characterized further: 1 g of lyophilized venom yielded 4.8 mg of purified elegantins.

E1 was further purified on an analytical column with a shallower initial gradient (Figure 1, middle panel). The active fraction, indicated as peak 1, was directly analysed by ES/MS, showing the presence of two major components with molecular masses of 7880.2 ± 0.8 and 7678.6 ± 0.4 Da respectively, and a minor species whose molecular mass was measured as 7591.9 ± 0.4 Da. On the basis of its accurate molecular mass, one compound was tentatively identified as the already known elegantin [here named isoform 1a (E1a)] previously described [7]

Figure 1 HPLC purification of elegantin isoforms

Top panel: semipreparative reverse-phase HPLC of ultrafiltrated venom from *T. elegans*. The active fractions are indicated as E1 and E2. The column was eluted with a linear gradient from 0% to 50% acetonitrile in 0.1% TFA over 60 min. Middle panel: analytical HPLC separation of fraction E1 with a linear gradient from 0% to 30% acetonitrile in 0.1% TFA over 60 min. The active fraction is indicated as peak 1. Peaks eluted early corresponded to non-peptidic material. Bottom panel: analytical HPLC separation of fraction E2 with a linear gradient from 10% to 40% acetonitrile in 0.1% TFA over 60 min. The resulting active fractions are indicated by 1 and 2.

Table 1 ES/MS analysis of the nine elegantin isoforms purified from *T. elegans*

Elegantin isoform	Molecular mass (Da)	
	Expected	Experimental
E1a (1–73)	7879.7	7880.2 ± 0.8
E1b (3–73)	7679.5	7678.6 ± 0.4
E1c (3–72)	7592.4	7591.9 ± 0.4
E2a (1–73)	7819.5	7818.8 ± 0.4
E2b (2–73)	7690.4	7689.5 ± 0.7
E2c (3–73)	7619.3	7618.2 ± 0.8
E2d (1–72)	7748.5	7748.5 ± 0.4
E2e (2–72)	7619.3	7618.8 ± 0.4
E2f (3–72)	7548.2	7547.7 ± 0.3

from *T. elegans*. The other components corresponded to truncated forms of E1a that lacked the N-terminal dipeptide Glu-Ala [isoform 1b (E1b); theoretical mass 7679.5 Da] and the N-terminal dipeptide plus the C-terminal Ser residue [isoform 1c (E1c); theoretical mass 7592.4 Da] respectively (Table 1). These assignments were confirmed by amino acid sequencing of the reduced and alkylated proteins and by FAB/MS analysis of their respective tryptic digests, with no detection of amino acid substitutions.

E2, the most abundant low-molecular-mass component in the venom, was similarly resolved into two active fractions by the use of reverse-phase HPLC chromatography on an analytical column (Figure 1, bottom panel, peaks 1 and 2). These purified species were directly analysed by ES/MS, showing that both components consisted of three different molecular species whose mass differences suggested the occurrence of truncated isoforms.

Aliquots of both fraction 1 and fraction 2 were then reduced, alkylated with iodoacetic acid and subjected to automated N-terminal sequence determination, leading to the identification of the first 30 amino acid residues. The complete amino acid sequence of both proteins was determined by sequencing the HPLC-purified peptides originating from the tryptic and Asp-N digestions and confirmed by FAB/MS analyses of the same peptide mixtures.

The primary structure of the elegantin corresponding to peak 1 (Figure 1, bottom panel) [named isoform 2a (E2a)] is shown in Figure 2; on the basis of this sequence, the nature of the three components detected in the ES/MS analysis was easily identified. In fact, the theoretical mass of E2a is 7819.5 Da, which agrees very well with that determined for the main component (7818.8 ± 0.4 Da) (Table 1). The remaining two species resulted from progressive amino acid removal from the N-terminus of

isoform 2b (*des*-Glu, theoretical mass 7690.4 Da) and isoform 2c (*des*-Glu-Ala; theoretical mass 7619.3 Da) respectively.

The amino acid sequence of the elegantin corresponding to peak 2 (Figure 1, bottom panel) was found to be identical with that reported in Figure 2 except for the loss of the C-terminal Ala residue. Similarly the ES/MS analysis showed the presence of three components originating from N-terminal processing. The corresponding isoforms were named E2d, E2e and E2f respectively. Table 1 summarizes the MS data obtained from the analysis of the nine iso-inhibitors of platelet aggregation, with the theoretical masses calculated on the basis of their primary structures. The determination of the accurate molecular masses of the native molecules before and after alkylation with iodoacetic acid demonstrates the occurrence of six disulphide bridges in each isoform (results not shown).

ES/MS analysis quickly allowed the evaluation of the relative ratios of the isoforms depending on the different batches of crude starting material. Peptides considered to be homogeneous by MS analysis were tested for biological activity.

Sequence alignment and model construction

The sequence alignment of elegantins with representative RGD-containing bioactive peptides is shown in Figure 3. Isoforms E1a and E2a display 79.4% similarity with each other, and share 54.8% identity and 63.0% similarity with respect to the 20 disintegrins presented in the figure considered as a group. Lower values were found for elegantins when compared with other RGD-containing peptides isolated from leech. E1a and E2a contain all of the 40 invariant residues found in the disintegrin family, including all the 12 cysteine residues involved in disulphide bridges [33,34]. These residues generally constitute the compact core of the molecule, as clearly shown by the kistrin NMR structure [27]. The three-dimensional structure of flavo-ridin has been determined by NMR spectroscopy and found to resemble that of kistrin [35].

The primary structures of E1a and E2a present 76.5% and 70.6% identity and 80.1% and 75.0% similarity, and the spacing of the 12 cysteine residues is conserved, with respect to that of kistrin. Thus we assumed that kistrin and elegantins adopt a similar folding of the peptide chain. The kistrin solution structure was modified by replacing its sequence with that of elegantin isoforms, skipping the N- and C-terminal regions. We analysed the local environments around each amino acid residue to ascertain whether these substitutions were producing some sterically forbidden interactions. Most of the changes involved amino acids with side chains pointing outwards from the molecular surface. Thus the replacement with Glu-4, Gly-9, Ala-30, Asp-31, Asp-36, Arg-39 and Asn-69 in both isoforms, and Gly-70 in E1a or Gly-64 and Pro-70 in E2a, seemed not to influence the global folding of the molecules.

Figure 2 Amino acid sequence of E2a (Eleg 2a)

The primary structure of E2a was established by amino acid sequencing of the intact carboxymethylated protein and of the purified peptides resulting from trypsin (T) and endoproteinase Asp-N (D) digestions and confirmed by the respective FAB/MS analyses. Nt, N-terminal sequence.

	1	10	20	30	40	50	60	70
	o o#o#	o oo##oooo#ooooooo#	oo## o# o	#o +++	o # o o	#oo		
Elegantin 2a	EAGEECDGSPENFCCDAATCKLRPGAQCADGLCCDQCRFIEEGIIICRRARGDDLDDYCNGISGDCPRNPFHA							
Elegantin 2b	-----							
Elegantin 2f	-----							
Elegantin 1a	-----				KKKRT	NP-R-T-Q-A	GLYS	
Elegantin 1b	-----				KKKRT	NP-R-T-Q-A	GLYS	
Elegantin 1c	-----				KKKRT	NP-R-T-Q-A	GLY	
Trigramin A	---D---A-----I---GE---				S---TV--I---		R-AG	
Batroxostatin	---T-----				E---KGA-K---	NP-R-T-Q-A	R-	
Kistrin	-K---S-----				GE---E-K-SRA-K---	IP--MP-R-T-Q-A	YH	
Albolabrin	---D---A-----L---				GE---	S-MKK-T---	AG---L--	
Flavordin	-----S-----					KKKTG---I---	FP-R-T-L-N---	WNDL
Barbourin	-----					MKK-TV--V-K--WN--	T-T-Q-A---	GLYG
Applagin	-----				E---	K-MK--TV--	VN---	AG---
Molossin	---I-----					KK-K---	NP-R-T-Q-A---	R---
Tergeminin	-----A-----					MKK-TV--V---	WN--T-T-Q-A---	GLYG
Rhodostomin	-K---S-----				GE---E-K-SRA-K---	IP--MP-R-T-Q-A	YH	
Viridin	-----A-----					KK-K---	NP-R-T-Q-A---	R---
Jararacin	-----T-G-----				E---	KGA-K---	NP-R-T-Q-A---	R---
Durissin	-----A-----					KK-TV--P---	WN--T-T-Q-A---	GLYG
Basilicin	-----A-----				E---	KK-K---	NP-R-T-Q-A---	H---
Cotiarin	-----A-----				E---	KGA-K---	NP-R-T-Q-A---	R---
Lachesin	-----A-A-----				E---	KK-K---	NP-R-T-Q-A---	GYYG
Cereberin	-----A-----				E---	KK-K---	NP-R-T-Q-A---	R---
Lutosin	-----A-----					KK-TV--V---	WN--T-T-Q-A---	GLYG
Cerastin	-----T-----					MKK-TV--V---	WN--T-T-Q-A---	GLYG
Crotatroxin	-----A-----					KK-TV--P---	WN--T-T-Q-A---	GLYG
Echistatin A1	-----			E-ES-P--RN-K-LK--T--K-----	M-----	KTC-----	HKGPAT	

Figure 3 Comparison of the amino acid sequences of snake venom integrin antagonists

Primary structures of disintegrins isolated from *T. elegans* as determined in this study and known members of the snake venom disintegrin family [3–11]. The molecules are arranged by length, with the tripeptide sequence RGD, identical cysteine residues and conserved amino acids highlighted (+, # and O respectively). Amino acid substitutions are reported; identical residues are indicated with —.

Figure 4 Proposed model of the α -carbon co-ordinates for E2a

This stereochemical view of the molecule was constructed on the basis of the solution structure of kistrin [27] as described in the Experimental section. The polypeptide backbone corresponding to amino acid substitutions with respect to kistrin is numbered and is shown in light grey; the cysteine residues involved in the disulphide bridges are emboldened; the RGD motif is highlighted.

In contrast, amino acid substitutions in the region of residues 39–46 seemed to produce some distortions of the polypeptide conformation, because Ser-41 and Lys-45 in kistrin are pointing towards the interior of the protein structure. In isoform E1a the presence of a Thr residue at position 45 instead of Lys determined the formation of a cavity; this space was filled with the side chain of the lysine that replaces the Ser residue at position 41. Approximate hand-modelling of isoform E2a in this region, on

the basis of local staggered conformations and van der Waals and charge interactions, indicates that the cavity is filled with the hydrophobic side chains of Ile-41 and Ile-45 replacing Ser-41 and Lys-45. The presence of a Tyr residue at position 58 in E2a, close to these two hydrophobic amino acids, could stabilize this interaction.

The amino acids present in the region of residues 48–55 embedding the RGD motif participate in a β -hairpin structure

Figure 5 Dose-dependent inhibition of platelet aggregation by E1a (○), E2a (●) and kistrin (▲)

ADP-induced platelet aggregation was measured in (A) human and (B) canine PRP. Each data point is the average for three experiments performed in duplicate (S.E.M. less than 10%). The relative IC₅₀ values are tabulated in the inset.

that protrudes extensively from the molecular core. Despite the amino acid substitutions observed, the residues present in the polypeptide chain seemed compatible with an overall conservation of a β -hairpin structure, because no sterically unfavourable interactions were observed. Figure 4 shows the stereographical view of the α -carbon atoms, the six disulphide bridges and the RGD loop in E2a.

Inhibition of platelet aggregation

Elegantin isoforms were assessed *in vitro* on human and canine PRP for inhibition of platelet aggregation in response to a standard concentration of ADP (10 μ M). Figure 5 shows the percentage inhibition of platelet aggregation as a function of disintegrin concentration; the activity of kistrin is also reported for comparison. The IC₅₀ values of elegantin isoforms as determined by the dose-effect curves are tabulated in the insets. Both isoforms E1a and E2a were strong inhibitors of ADP-induced platelet aggregation in either human or canine PRP. However, E1a was more active than E2a towards human platelets, whereas E2a showed greater inhibitory activity towards ADP-induced aggregation of canine platelets. The C-terminal truncated isoforms of E1a and E2a exhibited inhibitory activity towards aggregation in both human and canine PRP identical with those of the parent peptides (results not shown). Kistrin was more effective than elegantin isoforms in inhibiting platelet aggregation in both human and canine PRP.

Figure 6 Inhibition of B16-BL6 mouse melanoma cell attachment to fibronectin by E1a (○), E2a (●) and kistrin (▲)

Different amounts of inhibitor were added to the cells immediately before their being plated in the wells of plastic dishes coated with 10 μ g/ml of human plasma fibronectin. Plates were incubated at 37 °C for 1 h and then adherent cells were quantified as described in the Experimental section. Each data point represents the mean of 10 values (S.E.M. less than 5%).

Inhibition of cell adhesion to fibronectin

E1a and E2a were also evaluated for their ability to interact with the fibronectin receptor(s) on the cell membrane, by examining the attachment of B16-BL6 mouse melanoma cells to fibronectin. The effect of kistrin was also assessed. Both isoforms strongly inhibited B16-BL6 cell attachment to fibronectin-coated dishes, in a dose-dependent manner (Figure 6). E1a was more effective than E2a in inhibiting cell adhesion to fibronectin: 2.5 nmol/ml of E1a produced an inhibition of 60% in cell attachment to fibronectin, whereas the same concentration of E2a produced an inhibition of 42%. The inhibitory activity of kistrin was less than those of both elegantin isoforms: at the maximum tested dose (2.5 nmol/ml), kistrin showed a percentage of inhibition of approx. 35%, 1.7-fold and 1.2-fold lower than that caused by E1a and E2a respectively. The inhibitory activity on cell adhesion to the extracellular matrix protein was not due to a cytotoxic effect of disintegrins, because the viability of cells, as measured by the Trypan Blue exclusion method, was always 95% after the exposure of B16-BL6 cells to the highest tested dose of each peptide for 10 min, 1 h and 24 h.

The inhibition of B16-BL6 cell adhesion to fibronectin by E1a and E2a and kistrin was fully reversible: it was, in fact, completely removed by washing the cells preincubated with 2.5 nmol/ml of disintegrin for 10 min at 37 °C under a CO₂ atmosphere. Elegantin isoforms seemed to behave as competitive inhibitors of cell adhesion, because a decrease in the percentage inhibition was observed by increasing the amount of coating substrate (results not shown).

DISCUSSION

Elegantins have been purified from the venom of *T. elegans* and characterized as potent inhibitors of platelet aggregation. The procedure described here allowed the rapid execution of reverse-phase HPLC of the crude venom, avoiding the need for other preliminary chromatographic steps. ES/MS analysis of the purified active fractions indicated the presence of different components co-eluting in the same, apparently homogeneous,

chromatographic peak. These species were accurately characterized for their molecular masses and subjected to peptide sequencing. Their complete amino acid sequences were confirmed by FAB/MS. The combined use of classical protein chemistry procedures and MS methods allowed us to demonstrate the presence in the *T. elegans* venom of a whole family of bioactive peptides belonging to the disintegrin family, rather than a single component as previously described [7]. The two major elegantin isoforms were named E1a and E2a: the former corresponds to the protein previously reported, whereas the latter is a new member of the family of RGD-containing peptides. The structure of the other seven isoforms was also determined. They correspond to N- and C-terminal truncated forms of E1a and E2a. All the molecules contain 12 cysteine residues (which are highly conserved in all medium-size disintegrins [4]) organized in six disulphide bridges, and the cell-adhesion recognition site RGD tripeptide [1,2]. Isoform E2a is the most acidic disintegrin so far discovered.

Elegantins bear high sequence similarity to many of the venom inhibitors from the Crotalidae and Viperidae but very low similarity to RGD-containing peptides from leeches and the Elapidae. Some of these disintegrins present different truncated forms originated by proteolytic processing at both the N- and C-termini of the peptide chain, in a similar manner to elegantins [3,5,15] (Figure 3). This might be due to the action of different proteases on the disintegrin precursors. The existence in the venom of various snakes of several proteolytic enzymes interacting with components of the haemostatic system has been previously reported [36]. However, it cannot be excluded that the different disintegrin variants reflect the existence of a family of multiple genes or might result from the contributions of different snakes to the pooled venom.

Different disintegrin isoforms have been purified and characterized from *Echis carinatus*, *Bitis arietans* and *Trimeresurus gramineus* [5,15,37]. The occurrence of isoforms of different components, including phospholipases A1 and disintegrins, in the venom of the same viper has been described [38]. This molecular variety might represent a selective advantage for the viper that feeds on different animal species and needs to defend itself against different predators [37–39].

The proposed models for E1a and E2a predict an overall conservation of the three-dimensional arrangement of the molecule with respect to the solution structure of kistrin [27]. Although the positioning of the 12 Cys residues is retained in all disintegrins, a different pattern of disulphide bridges has been proposed for albolabrin from those for kistrin and flavoridin [34]. Thus there is the possibility of different S–S pairings for elegantins, although most of the conclusions reported here would not be affected. Calculations suggest that the energies of different combinations of disulphide bonds are approximately equivalent in the disintegrins. Moreover the alternative bridging pattern would not affect the conformation of the RGD-containing loop.

Both isoforms present the RGD sequence at the tip of a β -hairpin loop, and most of the changes in their primary structures are well accommodated into the structure of kistrin without any severely unfavourable steric effect. However, the effect of amino acid substitutions near the RGD motif must be emphasized. NMR studies showed that in the kistrin structure Glu-12 is involved in ion-pairing with Lys-45. This amino acid is replaced by Thr-45 and Ile-45 in E1a and E2a respectively; as a result, Glu-12 seemed to be involved in a salt bridge with Arg-49 in both elegantin models. These hypothetical interactions might induce a different flexibility in the loops containing the RGD motif, thus affecting the interaction with different receptor types. Furthermore, the amino acid substitutions at the positions flanking the

RGD motif in both elegantin isoforms progressively reduce the hydrophobicity of this region passing from kistrin to E1a and E2a:

	49	55
Kistrin	-IP-RGD-MP-	
E1a	-RA-RGD-NP-	
E2a	-RA-RGD-DL-	

The relevance of the hydrophobic nature of the residues present around the RGD loop has been reported [3,25]. In the light of these considerations, the observed biological properties can be rationalized in terms of structural differences.

E1a was more active than E2a in the inhibition of ADP-induced aggregation in human PRP. These findings agree well with previous studies showing that large hydrophobic side chains in the X position of the RGD_X sequence are optimal for high-affinity interactions with human platelet GPIIb-IIIa and possibly other integrins [3,25,26]. Thus the higher activity of E1a than E2a could be due in part to the presence of Asn-54 instead of Asp-54 after the RGD sequence in E1a and E2a respectively. Kistrin, which has a Met residue at position 54, showed a 1.5-fold and 3-fold higher potency than E1a and E2a respectively. Moreover the reduced hydrophobicity of the region flanking the RGD motif in both elegantins could explain their lower biological activity with respect to kistrin. These results are consistent with the observations of Lu et al. [40], who described a higher activity for kistrin than elegantin in inhibiting the interaction of human platelet GPIIb-IIIa with immobilized fibrinogen.

A different pattern of inhibitory activity was observed when canine PRP was employed in the test system. Isoform E2a was 2.2-fold more effective than E1a, whereas kistrin showed again a higher inhibitory effect towards canine platelet aggregation than both elegantin isoforms. These results demonstrate that elegantin isoforms present different specificities towards the platelet GPIIb-IIIa from different mammalian species and support the idea of an evolutionary molecular heterogeneity for the disintegrins directed to an optimal interaction with different enemies [37]. It has been suggested that the tripeptide sequences -IEE- and -LEE- at positions 41–43 of trigramin a and echistatin γ respectively, could be involved in their interaction with platelets [37]. The sequence -IEE- is present in E2a, whereas E1a contains a stretch of basic residues (-KKK-), thus causing a difference of local net charge on the protein surface that could affect the interaction of the two isoforms with the phospholipid bilayer. The truncated isoforms of both E1a and E2a exhibited comparable biological activities to those of the parent peptides when assayed in either human or canine PRP (results not shown), thus suggesting that the C-terminal regions of disintegrins do not affect their selectivity and specificity in receptor recognition.

The ability of E1a and E2a to compete with fibronectin for binding to its receptor(s) on B16-BL6 mouse melanoma cells was also evaluated. Both isoforms were strong, non-cytotoxic, reversible, competitive, dose-dependent inhibitors of B16-BL6 cell attachment to fibronectin, suggesting that elegantin isoforms interact with fibronectin receptor(s) on the cell surface. Comparison of their activity with that of kistrin showed that the elegantins are more effective in inhibiting cell attachment to immobilized fibronectin. E1a, containing the sequence RGDN at positions 51–54, was more active than E2a, which contains the tetrapeptide RGDD. These results are in agreement with previous observations showing that RGDNP-containing peptides are more potent inhibitors of fibronectin receptors $\alpha_v\beta_3$ and $\alpha_5\beta_1$ than peptides containing the RGD_X sequence [3].

In conclusion, disintegrins seem to be able to exert a variety of possible effects on cell–cell and cell–matrix binding and signalling. The results of this study confirm the growing evidence that the specificity and the affinity of a snake-venom peptide for individual receptors on the cell surface should not be ascribed only to the conformational state of the RGD sequence in its molecule: the presence of secondary binding determinants on the protein has also to be considered. The evaluation of the structural features of these compounds together with their functional characterization could provide a novel approach to the design of integrin antagonists with efficacy for the therapy of various human diseases, including thrombosis, tumour invasion and metastasis.

We thank Professor D. Barra (Dipartimento di Scienze Biochimiche, Università 'La Sapienza' di Roma, Rome, Italy) for the amino acid analyses, and V. Pavone (Dipartimento di Chimica, Università di Napoli, Naples, Italy) for helpful discussion about the molecular modelling work and for careful reading of the manuscript. This work was supported by grants from M.U.R.S.T., project B/2180.

REFERENCES

- Hynes, R. O. (1987) *Cell* **48**, 549–554
- Ruoslahti, E. and Pierschbacher, M. (1987) *Science* **238**, 491–497
- Scarborough, R. M., Rose, J. W., Naughton, M. A., Phillips, D. R., Nannizzi, L., Arfsten, A., Campbell, A. M. and Charo, I. F. (1993) *J. Biol. Chem.* **268**, 1058–1065
- 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
- Dennis, M. S., Henzel, W. J., 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
- Gan, Z. R., Gould, R. J., Jacobs, J. W., Friedman, P. A. and Polokoff, M. A. (1988) *J. Biol. Chem.* **263**, 19827–19832
- Williams, J., Rucinski, B., Holt, J. and Niewiarowski, S. (1990) *Biochim. Biophys. Acta* **1039**, 81–89
- Williams, J. A., Lu, X., Rahman, S., Keating, C. and Kakkar, V. (1992) *Biochem. Soc. Trans.* **21**, 73S
- Huang, T. F., Holt, J. C., Lukasiewicz, H. and Niewiarowski, S. (1987) *J. Biol. Chem.* **262**, 16157–16163
- Chao, B. H., Jakubowski, J. A., Savage, B., Chow, E. P., Marzec, U. M., Harker, L. A. and Maraganore, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8050–8054
- Yasuda, T., Gold, H. K., Leinbach, R. C., Yoita, H., Fallon, J. T., Guerrero, L., Napier, M. A., Bunting, S. and Collen, D. (1991) *Circulation* **83**, 1038–1047
- Mazur, P., Henzel, W. J., Seymour, J. L. and Lazarus, R. A. (1991) *Eur. J. Biochem.* **202**, 1073–1082
- Sawyer, R. T. (1986) in *Leech Biology and Behaviour*, vol. 2, pp. 467–523, Clarendon Press, Oxford
- Seymour, J. L., Henzel, W. J., Nevins, B., Stults, J. and Lazarus, R. A. (1990) *J. Biol. Chem.* **265**, 10143–10147
- Shebuski, R. J., Ramijit, D. R., Bencen, G. H. and Polokoff, M. A. (1989) *J. Biol. Chem.* **264**, 21550–21556
- Malinconico, S. M., Katz, J. B. and Budzynski, A. (1994) *J. Lab. Clin. Med.* **103**, 44–58
- Markwardt, F. (1988) in *Hemostasis and Animal Venoms* (Pirkle, H. and Markland, F. G., eds.), pp. 255–269, Marcel Dekker, New York
- Waxman, L., Smith, D. E., Arcuri, K. E. and Vlasuk, G. P. (1990) *Science* **248**, 593–596
- Knudsen, K. A., Tuszyński, G. P., Huang, T.-F. and Niewiarowski, S. (1988) *Exp. Cell Res.* **179**, 42–49
- Soszka, T., Knudsen, K. A., Beviglia, L., Rossi, C., Poggi, A. and Niewiarowski, S. (1991) *Exp. Cell Res.* **196**, 6–12
- Staiano, N., Villani, G. R. D., Di Martino, E., Squillacioti, C., Vuotto, P. and Di Natale, P. (1995) *Biochem. Mol. Biol. Int.* **35**, 11–19
- Sheu, J. R., Lin, C. H., Cheng, J. L., Teng, C. M. and Huang, T.-F. (1992) *Jpn. J. Cancer Res.* **83**, 885–893
- Trikha, M., De Clerck, Y. and Markland, F. (1994) *Cancer Res.* **54**, 4993–4998
- Scarborough, R. M., Rose, J. W., Hsu, M. A., Phillips, D. R., Fried, V. A., Campbell, A. M., Nannizzi, L. and Charo, I. F. (1991) *J. Biol. Chem.* **266**, 9359–9362
- Fauchere, J. L., Morris, A. D., Thuriereau, C., Simonet, S., Verbeuren, T. J. and Kieffer, N. (1993) *Int. J. Pept. Protein Res.* **42**, 440–444
- Dennis, M. S., Carter, P. and Lazarus, R. A. (1993) *Proteins Struct. Funct. Genet.* **15**, 312–321
- Adler, M., Lazarus, R. A., Dennis, M. and Wagner, G. (1991) *Science* **253**, 445–448
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275
- Pucci, P., Carestia, C., Fioretti, G., Mastrobuoni, A. M. and Pagano, L. (1985) *Biochem. Biophys. Res. Commun.* **130**, 84–90
- Nitti, G., Orrù, S., Bloch, C., Mohry, L., Marino, G. and Pucci, P. (1995) *Eur. J. Biochem.* **228**, 250–256
- Abola, E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. and Weng, J. (1987) in *Data Commission of the International Union of Crystallography* (Allen, F. H. and Sievers, R., eds.), pp. 107–132, Plenum, Bonn
- Bernstein, F., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.* **112**, 535–542
- Gray, W. R. (1993) *Protein Sci.* **2**, 1749–1755
- Calvete, J. J., Schafer, W., Soszka, T., Lu, W., Cook, J. J., Jameson, B. A. and Niewiarowski, S. (1991) *Biochemistry* **30**, 5225–5229
- Klaus, W., Broger, C., Gerber, P. and Senn, H. (1993) *J. Mol. Biol.* **232**, 897–906
- Markland, F. S. (1983) *J. Toxicol.* **2**, 119–160
- Chen, Y.-L., Huang, T.-F., Chen, S.-W. and Tsai, I.-H. (1995) *Biochem. J.* **305**, 513–520
- Nakashima, K. I., Ogawa, T., Oda, N., Hattori, M., Sakaki, H. and Ohno, M. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5964–5968
- Daltry, J. C., Wuster, W. and Thorpe, R. S. (1996) *Nature (London)* **379**, 537–540
- Lu, X., Williams, J. A., Deadman, J. J., Salmon, G. P., Kakkar, V. V., Wilkinson, J. M., Baruk, D., Anthi, K. S. and Rahman, S. (1994) *Biochem. J.* **304**, 929–936